Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 62 (2014) 576 – 582

Energy Procedia

6th International Conference on Sustainability in Energy and Buildings, SEB-14

Hand Gesture Recognition and Interface via a Depth Imaging Sensor for Smart Home Appliances

Dong-Luong Dinh[a], Jeong Tai Kim[b], and Tae-Seong Kim[c,]*

[a]*Department of Computer Engineering, Kyung Hee University, Yongin, South Korea*
[b]*Department of Architectural Engineering, Kyung Hee University, Yongin, South Korea*
[c]*Department of Biomedical Engineering, Kyung Hee University, Yongin, South Korea,*

Abstract

This paper presents a novel hand gesture interface system via a depth imaging sensor for appliances control in smart home environments. To control appliances with hand gestures, we recognize the hand parts in a depth hand silhouette and generate control commands. In our methodologies, we first create a database (DB) of synthetic hand depth silhouettes and their corresponding hand parts-labelled maps and then train a random forests (RFs) classifier with the DB. Via the trained RFs, we recognize the hand parts in a depth silhouette. Then based on the information of the recognized hand parts, control commands are finally generated according to our implemented control interface. By testing our system on real hand gestures, we have obtained an average recognition rate of 98.50 % from five different subjects. With the presented interface system, users can control smart home appliances such as TV, fan, lighting, doors and change channels, temperature, and volume by just hand gestures.

Selection and peer-review under responsibility of KES International

Keywords: Smart home; Hand gesture recognition; Appliances control interface;

* Corresponding author. Tel.: +82-31-201-3731; fax: +82-31-201-3666.
E-mail address: tskim@khu.ac.kr

Selection and peer-review under responsibility of KES International
doi:10.1016/j.egypro.2014.12.419

1. Introduction

In recent years, intelligent human computer interaction (HCI) in smart home environments is getting a lot of attentions from researchers in many fields such as architectural, electrical, computer, and biomedical engineering. In particular, a hand gesture-based human and machine interface is a more intuitive, natural, and intelligent way than the traditional interface methodologies using keyboard, mouse, touchscreen, or remote control devices, since users can interface with household appliances with just their hand gestures. Potential applications of such an intelligent HCI based on hand gesture recognition include home entertainments [1], home appliances control [2], home healthcare systems [3, 4], etc. Among these applications, the appliances control for smart home is one of the important applications for its daily usage: household appliances such as TV, radio, fans, and doors can be controlled by just hand gestures along with changing channels, temperature, and volume [5, 6].

Earlier works proposed for hand gesture recognition to control household appliances can be divided into two approaches: one is inertial sensor-based and the other vision-based approaches [6]. In the first approach, one or an array of inertial sensors is used. They are, in general, tracked to get their positions, velocity, and acceleration. Then some motion features are used for hand gesture recognition to control home appliances such as radio, TV, and lightings [7-9]. However, the requirement of the sensor devices like a remote controller makes this kind approach unnatural and inconvenient in spite of its high sensitivity. To overcome the limitations of the sensor-based approach, the second approach uses imaging sensors or cameras to get gesture commands from colour, shapes, orientations, positions, contours, or motions features of a hand or body [10]. For instance, in [11], the authors developed an assisting system for the elderly and handicapped to open or close household appliances such as TV, lamps, and curtains with just hand gestures. In this work, a pointing gesture, which was captured on three cameras, was recognized by getting its position and direction features. In [12], the authors proposed a simple method to control air conditions remotely using four hand gestures. By using two cameras, their hand gestures were recognized by detecting the right or left hand gestures using some colour image-based detection algorithms. In general, the use of colour images in gesture recognition is sensitive to noise, lighting conditions, cluttered backgrounds, and especially occlusions. Recently, with the introduction of new depth imaging sensors which provide pixel-by-pixel distance images, a new brand of hand gesture recognition has been introduced. For instance, in [13], the authors used some extracted features from depth images which reflect the hand contour information to recognize some static hand poses. Then, dynamic hand gestures were recognized by analyzing a consecutive sequence of these static poses. Seven dynamic hand gestures were recognized and used for household appliances control. However, this study was not robust enough since only the hand shape information was used without discerning each hand part.

In this study, we present a novel hand gesture recognition and HCI system which recognizes each hand part in a hand depth silhouette and generate control commands to control appliances in smart home environments. The main advantage of our proposed approach is that the state of each finger is directly identified by recognizing the hand parts and then hand gestures are recognized based on the state of each finger. We have tested and validated our system on real data. Our experimental tests achieved 98.50 % in recognition of hand gestures with five subjects. Finally we have implemented and tested our HCI interface through which one can control home appliances: smart home appliances can be turned on and off; channels and volumes can be changed with just simple hand gestures.

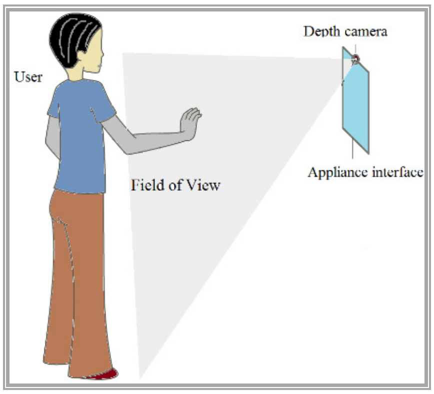


Fig 1. The setting of our proposed hand gesture-based interface system

2. Hand gesture-based interface system

Fig. 1 shows the setting of our hand gesture recognition and HCI system for appliances control in smart home environments. The system consists of two main parts: a depth camera which is used to get hand depth silhouettes and an appliances control interface which is used to give instructions to appliances. To make a user friendly interface, our hand gesture interface system allows users to interface with the appliances by understanding dynamic hand gestures which are recognized from the hand poses and their movements as described in Tables 1 and 2.

Table 1. Four basic types of hand poses

Pose ID	1	2	3	4
Function	Ready	Selection/Enter	Exit	Move menu
Hand Poses				

Table 2. Types of static and dynamic hand gestures

Gesture types	Functionality of the system	Gesture types	Functionality of the system
1	Recognize Pose 1 to ready the interface system	4	Recognize and track Pose 1 as a hand mouse
2	Recognize Pose 2 to make a selection (i.e., enter)	5	Recognize and track Pose 4 to move the menu right or left
3	Recognize Pose 3 to exit the system	6	Recognize and track Pose 4 to turn up or down volume

3. Methodology

Overall process of our proposed system for hand gesture recognition, shown in Fig. 2, consists of two main parts: in the first part of hand parts recognition, a synthetic hand database (DB), which contains more than thousands of pairs of depth maps and their corresponding hand parts-labelled maps, was generated. Then, the DB was used in training RFs. In the recognition stage, a depth image was first captured from a depth camera and then a hand depth silhouette was extracted by removing the background. Next, the hand parts of a depth silhouette were recognized using the trained RFs. In the second part of hand gesture recognition, a set of features was extracted from the labelled hand parts. Finally, based on the extracted features, hand gestures were recognized by our rules, generating interface commands.

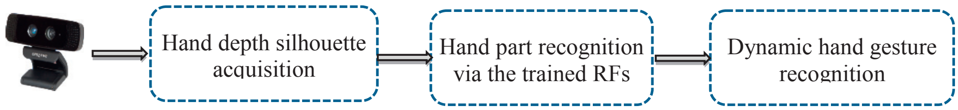


Fig 2. The framework of our hand gesture recognition system.

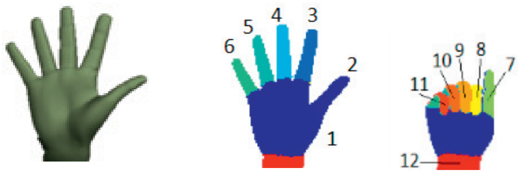
3.1. Hand depth silhouette acquisition

In our work, we used a creative interactive gesture camera [14]. The depth imaging parameters were set with the image size of 240 x 320, and frame rate of 30 fps. The hand parts were captured in the field of view of 70 degrees. Hand depth silhouettes were extracted with the background removal technique in [14].

3.2. Hand parts recognition

a. Synthetic hand DB generation

We created the DB with a synthetic hand model using 3Ds Max, a commercial 3D graphic package [15]. To identify hand parts, twelve labels were assigned to each hand model as shown in Fig. 3. The five fingers including the thumb, index, middle, ring, and pinkie fingers, were represented by ten corresponding labels including the five front and five back sides of the fingers. The front parts were coded with the indices of 2, 3, 4, 5 and 6. Likewise, the five back sides were coded with the indices of 7, 8, 9, 10 and 11, respectively. The images in the DB had a size of 320 x 240 with 16-bit depth values.



(a) (b)

Fig 3. Hand model: (a) 3D hand model in 3Ds-max and (b) Hand parts with twelve labels

b. Hand part recognition via RFs

In training RFs, we used an ensemble of 21 decision trees. The maximum depth of trees was 20. Each tree in RFs was trained with different pixels sampled randomly from the DB. A subset of 500 training sample pixels was drawn randomly from each synthetic depth silhouette. A sample pixel was extracted as done in [16], to obtain 200 candidate features.

To recognize hand parts of each hand depth silhouette, all pixels of each hand depth silhouette were classified by the trained RFs to assign a corresponding label out of the 12 indices. A centroid point was withdrawn from each recognized hand part, representing each hand part.

3.3. Dynamic hand gesture recognition

a. Hand poses recognition

From the recognized hand parts, we extracted a set of features. In our labelling, each finger was represented by two different labels: one label for its front side corresponding to the open state of the finger and another for its back side corresponding to the close state of the finger. From the information of the recognized hand parts, we identify the open or close states of each finger. The states of the five labelled fingers were identified and saved as features, namely f_{Thumb}, f_{Index}, f_{Middle}, f_{Ring}, and f_{Pinky} respectively.

$$f_{Fingers}(i) = \begin{cases} 1: Open & for\, Label = 2,3,4,5, or\ 6 \\ 0: Close & for\ Label = 7,8,9,10, or\ 11 \end{cases}$$

Table 3. Recognition rules of the four basic hand poses based on the states of the five fingers

f_{Thumb}	f_{Index}	f_{Middle}	f_{Ring}	f_{Pinky}	Pose type
0	1	0	0	0	1
1	1	0	0	0	2
0	0	0	0	0	3
1	1	1	1	1	4

0=Close and 1=Open.

To recognize four basic hand poses, we derived a set of recognition rules. The set of five features from the states of all fingers was used to decode the meaning of the four hand poses. The derived recognition rules are given in Table 3.

b. Dynamic hand gesture recognition

To understand dynamic hand gestures with the recognized hand poses as explained in Table 2, after recognizing the hand poses, we tracked their positions. To understand Gestures 1, 2, and 3, our system recognizes Pose 1, 2, and 3, respectively. Then Gesture 4 can be understood by recognizing Pose 1 and then tracking the centroid point position of the index finger in the *x* and *y* dimension which gets mapped on the interface screen: it acts as a hand mouse. To understand Gesture 5 which is used to move the menu to the right or left, Pose 4 is recognized and the centroid point position of the palm is tracked in the x dimension between two consecutive frames including the previous and current frames. By dividing the frame window into three sub-areas as presented in Fig. 4, if the tracked point of the current frame is moved from the area of current or previous pages to the area of next page, the screen menu slides to the right. Likewise, if the tracked point of the current frame is moved from the area of current or next pages to the area of previous page the screen menu slides to the left. To understand Gesture 6 to turn up or down volume, Pose 4 is recognized and then the centroid point position of the palm is tracked in the y dimension. The difference between the tracked points position of two consecutive frames is used to turn up or down the volume.

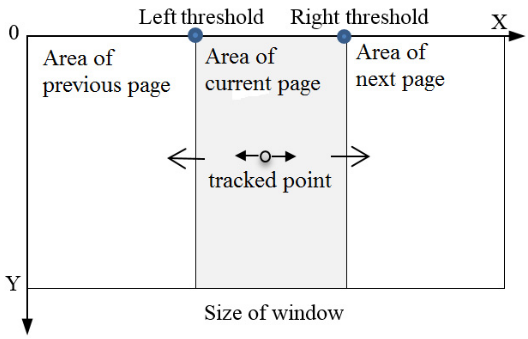


Fig 4. Three divided areas in the screen window to recognize dynamic hand gestures (i.e., Gesture type 5, 6, and 7)

4. Experimental Results and Demonstrations

4.1. Results of hand parts recognition

To evaluate our hand parts recognition quantitatively, we tested on a set of 500 hand depth silhouettes containing various poses over the four hand poses. The average recognition rate of the hand parts was 96.90 %. Then, we assessed the hand parts recognition on real data qualitatively, since the ground truth labels are not available. We only performed visual inspections on the recognized hand parts. A representative set of the recognized hand parts are shown in Fig. 5.

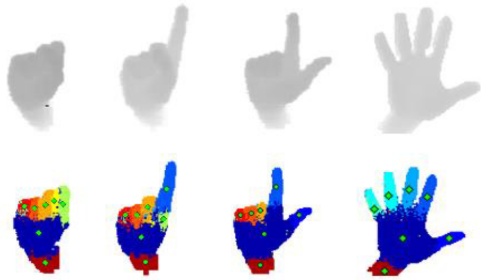


Fig 5. A set of representative results of the recognized hand parts via the trained RFs on real data.

4.2. Results of hand pose recognition

To test our proposed hand poses recognition methodology, a set of hand depth silhouettes was acquired from five different subjects. Each subject was asked to make 40 hand poses. Table 4 shows the recognition results of the four hand poses in a form of confusion matrix. The mean recognition rate of 98.50 % was achieved.

Table 4. The confusion matrix of the four hand poses recognition using our proposed system

Pose types	1	2	3	4
1	**100**			
2		**100**		
3		2	**98**	
4	1	1	2	**96**

4.3. Graphic user interfaces (GUIs) for demonstrations

We designed and implemented a GUI interface of our hand gesture interface system as presented in Figs. 6 and 7. How to use of our system can be explained by the following two examples. In example 1, to open and select any channel for TV, one should use Gesture 1 to open (i.e., activate) the system, use Gesture 4 as a hand mouse to select the TV icon on the GUI screen in Fig. 6 (a), and use Gesture 2 to open the TV. Then the TV channels as shown in Fig. 6 (b) opens for a selection of a channel. To browse the channel menus, one use Gesture 5 to slide the channel pages to the right (i.e., the next page) or left (i.e., the previous page), use Gestures 4 and 2 to select a channel, and finally use Gesture 3 to get back to the main GUI menu as shown in Fig 6 (a). In example 2, to control volume as well as change temperature, intensity of lightings, or speed of fans, one should use Gesture 1 to open the system, use Gesture 4 as a hand mouse to select the volume icon on the GUI screen, and then use Gesture 2 to open the volume screen as shown in Fig. 6 (c). To change volume, use Gesture 6 to turn up or down volume or use Gestures 4 and 2 to select the fixed minimum (Min), middle (Mid), or maximum (Max) volume levels. To get back to the main GUIs, one can use Gesture 3.

Fig 6. Illustrated gesture-based GUIs of appliance control system
The GUIs for (a) appliances selection, (b) TV channel selection, and (c) volume control.

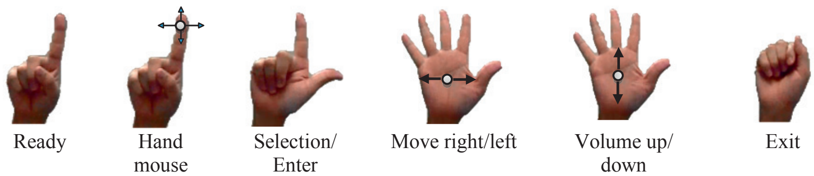


Fig 7. A hand gestures sequence and their tracked points used in our system

5. Conclusions

In this work, we have presented a novel hand gesture recognition system for appliance control in smart home using the labelled hand parts via the trained RFs from a hand depth silhouette. We have achieved the mean recognition rate of 98.50 % over the four hand gestures from five subjects. Our proposed hand gesture recognition method should be useful in automation applications for appliance control in smart home environment.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2008-0061908).

References

[1] Reifinger S, Wallhoff F, Ablassmeier M, Poitschke T, Rigoll G. Static and dynamic hand-gesture recognition for augmented reality applications. In HCI Intelligent Multimodal Interaction Environments, LNCS 4552; 2007. p. 728–737.

[2] Shao J. Smart Home Appliance Control. In 2008 IEEE Industry Applications Society Annual Meeting; 2008. p. 1–6.

[3] Bien ZZ, Park KH, Jung JW, Do JH. Intention reading is essential in human-friendly interfaces for the elderly and the handicapped. IEEE Transactions on Industrial Electronics; 2005. 52(6): 1500–1505.

[4] Levin-Sagi M, Pasher E, Carlsson V, Klug T, Ziegert T, Zinnen AA. Comprehensive Human Factors Analysis of Wearable Computers Supporting a Hospital Ward Round. In Proc. of the 4th International Forum on Applied Wearable Computing (IFAWC); 2007. p. 728–737.

[5] Shimada A, Yamashita T, Taniguchi R. Hand gesture based TV control system- Towards both user- & machine-friendly gesture applications. 19th Korea-Japan Joint Workshop on Frontiers of Computer Vision, (FCV); 2013. p. 121–126.

[6] Bhuiyan M, Picking R. Gesture-controlled user interfaces, what have we done and what's next. In 5th Collaborative Research Symposium on Security, E-Learning, Internet and Networking; 2009. p. 59–60.

[7] Mäntyjärvi J, Kela J, Korpipää P, Kallio S. Enabling fast and effortless customisation in accelerometer based gesture interaction. In Proceedings of the 3rd international conference on Mobile and ubiquitous multimedia; 2004. p. 25–31.

[8] Wan S, Hung TN. Human computer interaction using hand gesture, Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 2008.

[9] Ng WL, Ng CK, Noordin NK, Ali BM. Gesture based automating household appliances. 14th International Conference on HCI, Techniques and Environments; 2011. p. 285-293.

[10] Murthy GRS, Jadon RS. A review of vision based hand gestures recognition. International Journal of Information Technology and Knowledge Management; 2009. 2(2): 405–410.

[11] Bien ZZ, Do JH, Kim JB, Stefanov D, and Park KH. User-Friendly Interaction/Interface Control of Intelligent Home for Movement-Disabled People. In Proceedings of the 10th International Conference on Human-Computer Interaction; 2003.

[12] Cho JK, Park DR, Kim YH. A method of remote control for home appliance using free hand gesture. In IEEE International Conference on Consumer Electronics (ICCE); 2012. p. 293–294.

[13] Wu CH, Lin CH. Depth-based hand gesture recognition for home appliance control. In IEEE 17th International Symposium on Consumer Electronics (ISCE); 2013. p. 279–280.

[14] www.intel.com/software/perceptua.

[15] Autodesk 3Ds MAX, 2012.

[16] Shotton J, Sharp T, Kipman A, Fitzgibbon A, Finocchio M, Blake A, Moore R. Real-time human pose recognition in parts from single depth images. Communications of the ACM; 2012. 56(1):116-124.