Journal of Computers Vol. 27, No. 4, 2016, pp. 161-169
doi:10.3966/199115592016122704013

3D Arm Movement Recognition Using Kinect and Its Application in Interaction Design

Yu-Xiang Zhao[1*], Min-Chueh Kan[1]

[1] Department of Computer Science and Information Engineering, National Quemoy University
Kinmen 892, Taiwan, ROC
yxzhao@nqu.edu.tw

Received 20 January 2016; Revised 15 February 2016; Accepted 15 February 2016

Abstract. This paper proposed a Kinect based 3D arm movement recognition algorithm and applied this algorithm to an interactive water screen. In the proposed algorithm, a Kinect device was employed to track and capture the trajectories of user 3D arm movement. The 3D arm movement recognition was achieved using a moving average filter and projection filter to smooth the coordinate data in order to enhance the system's recognition efficiency and accuracy. Eight hand movement trajectories were established in this paper. The results indicated that the recognition efficiency was higher than 90%, both overall and for individual subjects. In order to test the proposed algorithm, we designed an interactive water screen and used the recognition result to control the display of the interactive water screen. The interactive water screen was designed by combining an Arduino microcontroller board with an electronic water valve. Finally, the experimental results show the effective integration between of the proposed algorithm and interactive water screen.

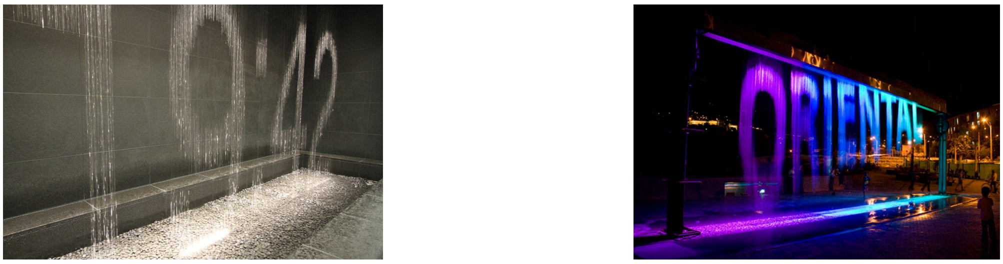
Keywords: arm movement recognition, interactive water screen, Kinect

1 Introduction

Motion sensing systems have become widely used by gaming enthusiasts in recent years. Such systems can detect body gestures to execute specific functions on a game console according to specific gestures, enabling users to control the system through gestures rather than conventional game controllers. This process is similar to communicating with other people through gestures. Accordingly, this study adopted a motion sensing system to recognize hand movement trajectories, enabling users to interact with the water screen through this motion sensing system. Currently, the two most widely used motion sensing systems for gaming are the Nintendo Wii Remote and Microsoft Kinect [1-3]. Because using the Nintendo Wii Remote would require the user to operate a hand-held game controller, Microsoft Kinect was selected as the interactive device for the water screen in this study. Current Kinect-based interactive systems can be characterized into two major groups based on their method of command recognition. For the first group, the Kinect device enables system interaction by obtaining and processing environmental depth images (grayscale), analyzing environmental dynamics, and recognizing complex hand gestures [4-6]. The second group of systems have been developed using the Kinect Software Development Kit or OpenNI, an open source software project, to collect and analyze skeletal tracking data, which are subsequently used to identify specific gestures and execute relevant functions [7-9].

Water screen is a type of installation art that is common in landscaping, where the time difference of dripping water is manipulated to display information in open spaces. Although the timeframe in which the water drop falls is short, water screens are effective for conveying information by utilizing a user's persistence of vision. Current water screen technologies have led to the development of various types of resentation method, such as color transformation and image projection [10-11] (see Fig. 1). Despite this

* Corresponding Author



(a) water clock [10]　　(b) water text [11]

Fig. 1. Water screen examples

diversity, water screens lack interactivity. If water screens were designed to enable viewers to interact with them, their applicability could be enhanced considerably.

In this paper, we proposed a Kinect based 3D arm movement recognition algorithm to recognize the hand gestures and applied this algorithm to an interactive water screen. In the proposed Kinect based 3D arm movement recognition algorithm, a Kinect device was employed to track and capture user hand gestures. The data were then converted into trajectory patterns through movement recognition system. Subsequently, the maps were then used as input data to control the water screen presentation (see Fig. 2). The aforementioned method is expected to effectively produce an interactive water screen design.

Fig. 2. Interactive water screen system.

2 3D Arm Movement Recognition

The Kinect device was used to acquire skeletal tracking data. Hand movements were tracked using coordinate data, which were used to determine the trajectory of the user's hand movements [12-13]. Fig. 3 describes the movement recognition procedure.

2.1 Obtaining hand movement trajectories

The Kinect device was used to record and store hand movement trajectories. To obtain 3-D coordinates (see Fig. 4(a)) for the movement trajectory of a user drawing an M-shape, the Kinect device renders color images based on the RGB color model (2-D images on an X–Y matrix), and then uses infrared elements to render the depth information of the image (Z axis).

2.2 Averaging hand movement traject ories

A moving average filter was used to smooth the tracked trajectories, because locally averaging sample data values can eliminate sudden changes in the data. A 5-point moving average filter was employed to optimize the movement trajectories twice, as shown in Fig. 4(b) and Fig. 4(c).

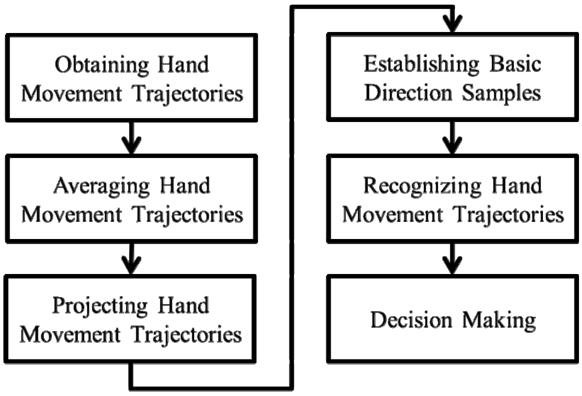
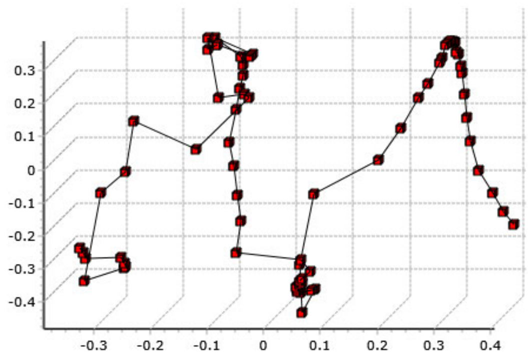


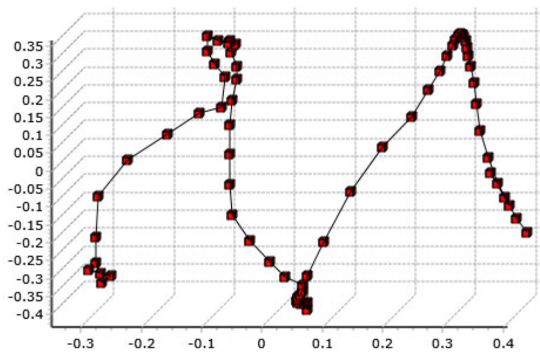
Fig. 3. Hand movement recognition procedure

2.3 Projecting hand movement trajectories

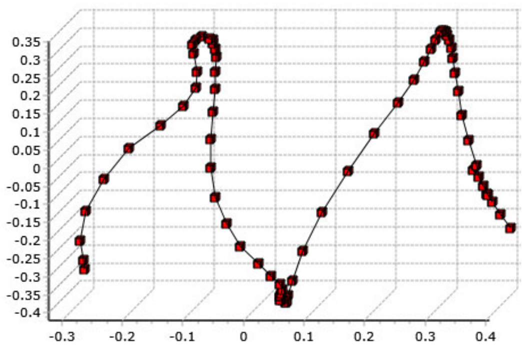
Hand movement trajectories are typically distributed along a plane. Therefore, the spatial problems associated with movement trajectories were simplified into planar problems to reduce the computational complexity, and enhance the computational speed. The hand movement trajectories were projected by observing the individual variance in the three dimensions, where the dimension with the lowest variance was the one with the smallest change (i.e., the one with the least data). The dimension with the lowest variance was eliminated to reduce the number of dimensions, the results of which are depicted in Fig. 4(d).



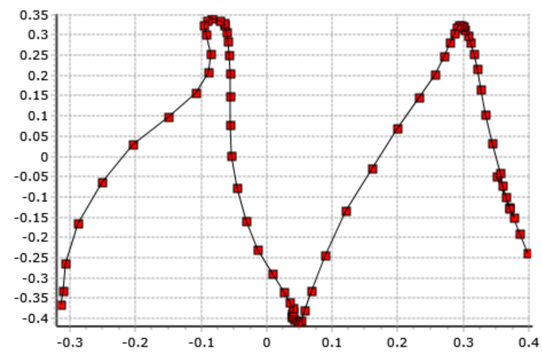
(a) Original movement trajectory



(b) Trajectory following initial smoothing



(c) Trajectory following second smoothing



(d) Projected trajectory

Fig. 4. Hand movement trajectories

2.4 Establishing basic direction samples

Fig. 5 shows eight basic unit vectors for describing hand movement trajectories. Eight trajectory samples were used in this study. Table 1 illustrates the association between the trajectory samples and the basic unit vectors.

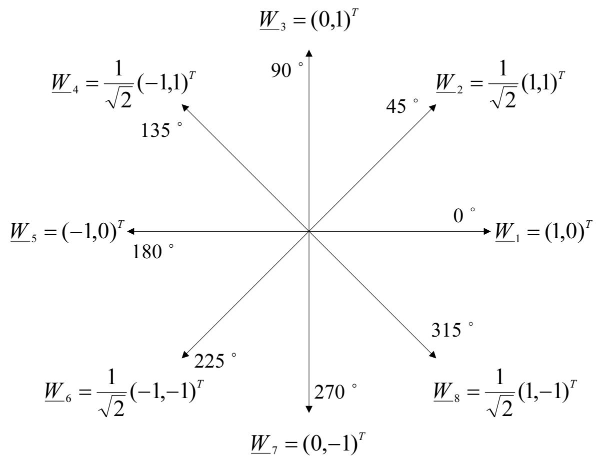
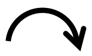
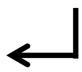
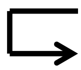


Fig. 5. Eight basic direction vectors

Table 1. The eight typical hand movement trajectories

Type	Trajectory	Template
1		$\underline{w}_1$
2		$\{\underline{w}_2, \underline{w}_1, \underline{w}_8\}$
3		$\{\underline{w}_7, \underline{w}_5\}$
4		$\{\underline{w}_8, \underline{w}_2\}$
5		$\{\underline{w}_5, \underline{w}_7, \underline{w}_1\}$
6		$\{\underline{w}_4, \underline{w}_6, \underline{w}_4, \underline{w}_6\}$
7		$\{\underline{w}_8, \underline{w}_5, \underline{w}_8\}$
8		$\{\underline{w}_1, \underline{w}_2, \ldots, \underline{w}_6, \underline{w}_7\}$ $\{\underline{w}_2, \underline{w}_3, \ldots, \underline{w}_7, \underline{w}_1\}$ ⋮

2.5 Recognizing hand movement trajectories

When an unknown arm movement is to be classified, the sample sequence is compared with each template and a measure of similarity (distance) between them is computed. To be precise, let $\underline{x}_i'$ be the 2D smoothed sample vector of the unknown arm movement with total $N+1$ sample vectors. First, we compute $\underline{x}_i$ in the following way:

$$\underline{x}_i = \underline{x}_{i+1}' - \underline{x}_i' \qquad i = 1, 2, \ldots, N \tag{1}$$

The vector $\underline{x}_i$ is then normalized to have unit length. Suppose the *k*th typical arm movement in the vocabulay is consisted of D_k directions, $\underline{w}_{d_1^k}, \underline{w}_{d_2^k}, \ldots \underline{w}_{d_{D_k}^k}$ where d_j^k is an integer between 1 and 8. In addition, these directions are also assumed to appear in the template in this order. Then the similarity between the unknown arm movement and *k*th typical arm movement is computed as follows:

$$S_k = \sum_{i=1}^{N_{d_1}^k} Out_{d_1^k}(\underline{x}_i) + \sum_{i=N_{d_1}^k+1}^{N_{d_1}^k+N_{d_2}^k} Out_{d_2^k}(\underline{x}_i) + \cdots + \sum_{i=N_{d_1}^k+N_{d_2}^k+\cdots+N_{d_{D_{k-1}}}^k+1}^{N} Out_{d_{D_k}^k}(\underline{x}_i) \quad \textbf{(2)}$$

where $N_{d_j}^k$ denotes the number of patterns that satisfy the following conditions:

$$Out_{d_{j-1}^k}(\underline{x}_{N_{d_1}^k+N_{d_2}^k+\cdots+N_{d_{j-1}}^k}) < Out_{d_j^k}(\underline{x}_{N_{d_1}^k+N_{d_2}^k+\cdots+N_{d_{j-1}}^k+1}) \quad \textbf{(3)}$$

and

$$Out_{d_j^k}(\underline{x}_{N_{d_1}^k+N_{d_2}^k+\cdots+N_{d_j}^k}) < Out_{d_{j+1}^k}(\underline{x}_{N_{d_1}^k+N_{d_2}^k+\cdots+N_{d_j}^k+1}) \quad \textbf{(4)}$$

2.6 Decision making

Finally, the unknown arm movement is classified to be the arm movement with the largest similarity in the vocabulary. That is, the unknown arm movement will be classified as the k^* th arm movement in the vocabulary if the following condition is satisfied:

$$S_{k^*} \geq S_k \qquad for \quad k \neq k^* \quad \textbf{(5)}$$

3 Interactive Water Screen System

Interactive water screen system is an application of our proposed 3D arm movement recognition algorithm. The system comprises two components: namely, a motion sensing system and a water screen system. The motion sensing system captures and tracks user 3D arm movements, which were processed on a computer and then displayed on the interactive water screen (see Fig. 6).

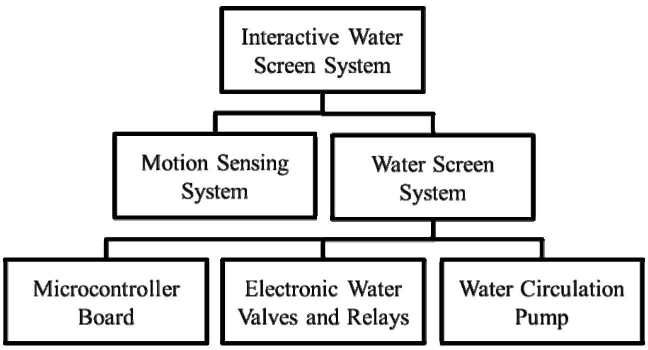


Fig. 6. System framework of the interactive water screen system

3.1 The motion sensing system

Microsoft Kinect was launched in November 2010. The Kinect device is a novel type of motion sensing game controller that enables users to control games by performing specific body gestures [14] (see Fig. 7). The Kinect is a right-handed coordinate system consisting of a red-green-blue (RGB) camera and a depth sensor composed of infrared emitter and receptor. The RGB camera is positioned at the center of the front panel of the Kinect device, with an infrared emitter and depth sensor located on either side of the camera. This array constitutes the time-of-flight (TOF) sensory system within the Kinect device, which is the primary system for sensing depth. The device also features an array of four microphones located across the bottom of the front panel. The mounting component enables users to tilt the device within a range of 54.

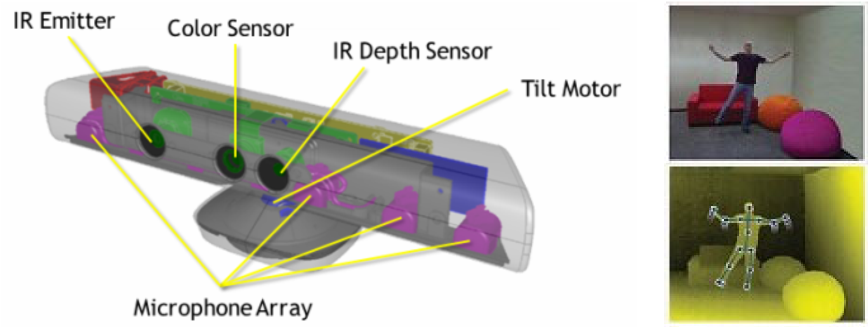


Fig. 7. Microsoft Kinect motion sensing device [14]

3.2 Water screen system

Fig. 8 shows the schematic diagram of the proposed water screen system. The water screen system comprises three modules: namely, a microcontroller board, an electronic water valves and relays, and a water circulation pump. The microcontroller's onboard software triggers the relays, causing the electronic water valve to adjust the time at which water flow is switched on and off to control the display resolution of the water screen. The circulation pump then controls water circulation to maintain stable water pressure.

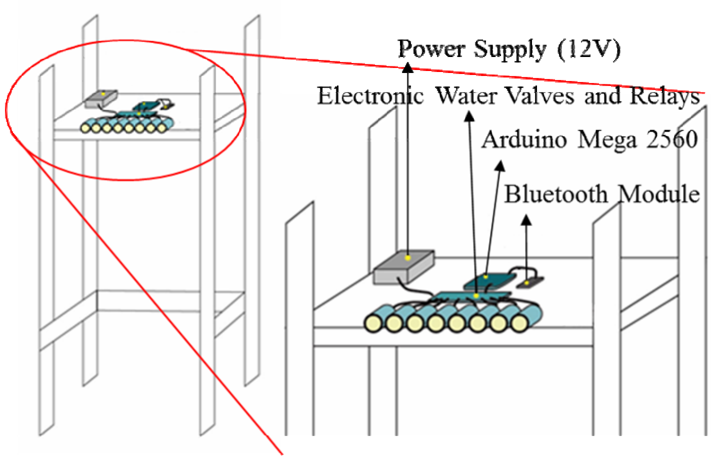
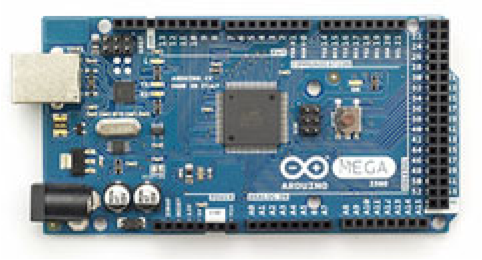


Fig. 8. Hardware configuration of the water screen system

The Arduino Mega 2560 microcontroller board was used as the core computing platform [15], as shown in Fig. 9(a). It is an open-source single-chip microcontroller that can be connected to a computer via USB. The microcontroller can be programmed using open source software, and an Atmel AVR microcontroller can be used to develop a simple user interface. In the present study, the Arduino microcontroller supplied power to the relays to control the electronic water valve, the image of which is shown in Fig. 9(b). An external power supply was required because of the electronic water valve's high power requirements. Fig. 10 shows a photograph of the complete water screen system.

4 Experimental Results

Five men and five women were recruited to test the recognition accuracy of the projected hand movement trajectories by recording their hand movements 10 times (80 recordings per subject). The entire databased consisted of 800 recordings. Table 2 shows the test results.



(a) Arduino Mega 2560 Microcontroller Board [15]

(b) Electronic water valve

Fig. 9. System hardware components

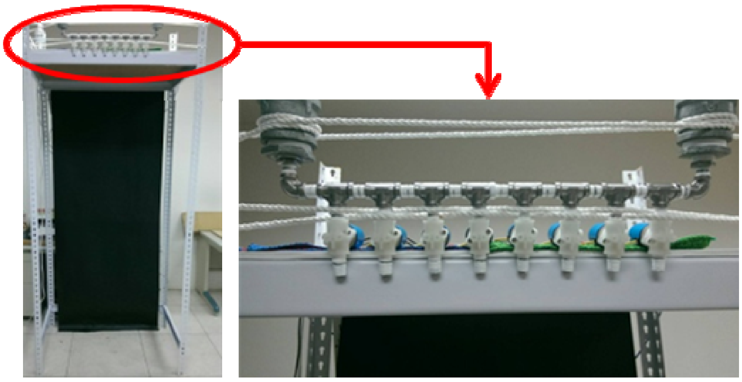


Fig. 10. Water screen system prototype

Table 2. The experimental results

U \ T	1	2	3	4	5	6	7	8	Ave.
1	80	80	100	100	100	100	100	100	95
2	90	80	90	90	90	90	100	100	91.25
3	100	90	80	100	100	90	100	100	95
4	90	90	90	90	90	100	80	100	91.25
5	100	100	90	80	90	100	100	100	95
6	90	80	90	90	100	100	100	100	93.75
7	100	90	80	90	80	80	80	100	87.5
8	100	80	100	90	90	80	80	100	90
9	100	90	100	80	100	90	100	100	95
10	100	90	100	90	100	90	100	100	96.25
Average	95	87	92	90	94	92	94	100	93

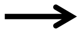
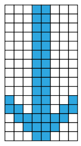
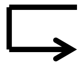
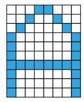
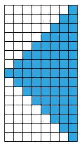
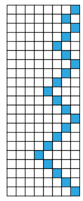
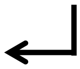
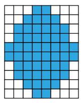
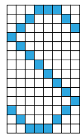
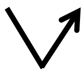
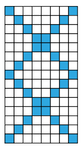
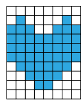
Note. * T: trajectory; U: user, and the measurements with percentage as the unit.

Finally, identified hand movement trajectories were used to produce trajectory patterns for the interactive water screen (Table 3). The interactive water screen then displayed the corresponding trajectory patterns according to the recognized hand movement trajectory.

5 Conclusion

This paper proposed a Kinect based 3D arm movement recognition algorithm and applied this algorithm to an interactive water screen. The Microsoft Kinect was employed to obtain coordinate data of user hand movements. To enhance the recognition efficiency and accuracy, the coordinate data were smoothed using a moving average filter and projection filter. Eight hand movement trajectories were established in the present study. The experimental results indicated that the recognition efficiency was higher than 90%, both overall and for individual subjects. Incorporating the proposed recognition method

Table 3. The trajectory patterns for the interactive water screen.

Trajectory	Pattern	Trajectory	Pattern
			
			
			
			

into interactive water screens can enhance user-device interactivity.

References

[1] J. Giles, Inside the race to hack the kinect, The New Scientist 208(2789)(2010) 22-23.

[2] J.C. Lee, Hacking the nintendo wii remote, IEEE Pervasive Computing 3(2008) 39-45.

[3] R. Francese, I. Passero, G. Tortora, Wiimote and kinect: gestural user interfaces add a natural third dimension to hci, in: Proc. of the International Working Conference on Advanced Visual Interfaces, 2012.

[4] A. Wilson, H. Benko, Combining multiple depth cameras and projectors for interactions on, above and between surfaces, in: Proc. of the 23nd Annual ACM Symposium on User Interface Software and Technology, 2010.

[5] O. Schlatter, B. Migge, A. Kunz, User-aware content orientation on interactive tabletop surfaces, in: Proc. of the 2012 International Conference on Cyberworlds, 2012.

[6] G. Gallagher, CSAIL brings minority report' technology alive, MIT CSAIL news, 〈http://www.csail.mit. edu/ node/1362〉

[7] F. Kammergruber, A. Ebner, P.W. Günthner, Navigation in virtual reality using Microsoft Kinect, in Proc. of 12th International Conference on Construction Application of Virtual Reality, 2012.

[8] T. Sakai, W. Fujimura, S. Robert, T. Kosaka, A. Shirai, AccuMotion: intuitive recognition algorithm for new interactions and experiences for the post-PC era, in: Proc. of the 14th Virtual Reality International Conference, 2012.

[9] D.Q. Freitas, A.E.F.D. Gama, L. Figueiredo, T.M. Chaves, D. Marques-Oliveira, V. Teichrieb, C. Araújo, Development and evaluation of a kinect based motor rehabilitation game, in: Proc. of Brazilian Symposium on Computer Games and Digital Entertainment, 2012.

[10] Koei Industry, 〈http://www.koeiaquatec.co.jp〉.

[11] Digital Water Curtain, 〈http://www.digitalwatercurtain.com〉.

[12] M.C. Su, Y.X. Zhao, H. Huang, H.F. Chen, A fuzzy rule-based approach to recognizing 3D arm movements, IEEE Trans. on Neural Systems and Rehabilitation Engineering 9(2)(2001) 191-201.

[13] Y.X. Zhao, A portable speaking-aid system for Taiwanese sign language, [Master thesis] New Taipei: Tamkang University, 2002.

[14] An Overview of The Sensor, 〈http://www.kinectingforwindows.com/ 2012/09/05/an-overview-of-the-kinect-sensor/〉.

[15] ARDUINO, 〈http://arduino.cc.〉.