2023-2024

Exploring Hand Gesture as a Touchscreen Alternative for Public Kiosks

2023-2024

Du Mengwei, Wu Jiajing

1. Introduction

1.1 Context

Public kiosks, commonly found in high-traffic areas like airports, shopping malls, and museums, traditionally utilize touchscreens for user interactions such as information access, ticketing, and check-ins.

1.2 Challenges

However, touchscreen interactions in kiosks pose several challenges.

Firstly, it compromises hygiene as physical contact increases the risk of germ transmission, a significant concern during a pandemic.

The height at which some kiosks are installed can limit accessibility for wheelchair users, making it difficult for them to reach and effectively use the touchscreen.

Furthermore, touchscreens often require frequent maintenance as they are prone to becoming unresponsive or malfunctioning after extensive use. This high maintenance can lead to increased operational costs and potential downtimes.

1.3 Objective

Our objective in this project is to propose a new in-air gesture that could potentially improve the current status quo of touch interfaces.

2. Approach

2.1 Related Works

"Television Control by Hand Gestures"[1]
System developed by William T. Freeman and Craig D. Weissman, demonstrates control of a television using simple hand gestures. It reduces the need for physical controllers or complex gesture memorization and showcases a similar application of intuitive, gesture-based interaction.

"A Virtual Dance Floor Game using Computer Vision"[2]
Designed by Daniel Brehme and his team, Cam2Dance leverages computer vision to detect and analyze the player's foot movements in real-time, reducing dependence on markers attached to the player's shoes. This advancement introduces an alternative mode of interaction beyond hand gestures, underscoring the versatility of human-computer interaction.

2.2 Proposed Solution

As a solution to the challenges posed by traditional touchscreens, we propose two ways of interactions: hand gestures and hand movement tracking. The second interaction would function similarly to touchscreens but operate in-air.

These solutions were chosen because they do not have the wheelchair accessibility limitations associated with foot movement tracking. They also significantly address hygiene, accessibility, and high maintenance issues posed by traditional touchscreens.

3. Methodology

3.1 System Design

1. Hand Gesture System

For each possible action to control kiosks, we designed corresponding hand gestures.

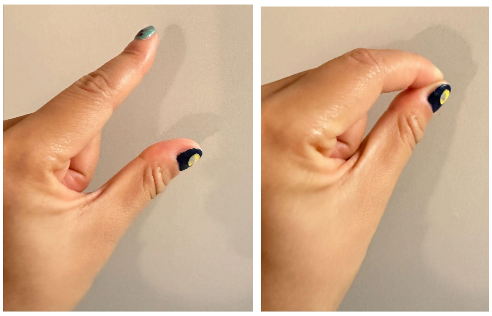
1) Navigation Gestures

- **Thumb_Down**: Move to the previous item or menu or move to the previous page.

- **Thumb_Up**: Move to the next item or menu or move to the next page.

2) Selection and Confirmation Gestures

- **Pinch Close**: Select the highlighted item.



- **Victory**: Confirm an action (e.g., placing an order, submitting input).

3) Cancel/Go Back

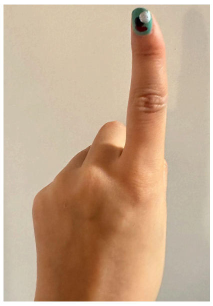
- **Fist**: Cancel the current action or return to the previous menu.

4) Assistance/Help

- **Open Palm**: a)At the beginning, wake up the system. b) Call for assistance or open the help menu.

2. Hand Movement Tracking

This method is similar to the touchscreen control but in a contactless way.
Move Cursor:



Move the finger to the appropriate position and keep pointing up, then the cursor will be moved to the position of the top of the finger.

Click:

Show the "**Victory**" gesture to click (confirm).

4. Implementation

Our solution mainly consisted of two parts: gesture detection and action execution.

4.1 Gesture Detection

This part is implemented by python using OpenCV especially **mediapipe** and **pynput** libraries. And **mediapipe** is for detecting gestures, **pynput** is for mouse simulation.

In our code, we load an existing model which includes detection of six gestures: fist, open palm, pointing up, thumb up, thumb down, victory. Based on this, we additionally implemented the detection of pinch gestures. This model can be downloaded by gesture_recognizer.task.

4.2 Action Execution

Since the actions are executed with the same amount of time in the kiosks system, we do not focus on this implementation of this part, and we use textual output instead.

For example, when it detects the "Victory" gesture, it outputs "Confirm".

4.3 Pipeline

Step1: Open camera, and get frames from camera video.
Step2: Process each frame to detect gestures.
Step3: Output the corresponding command of the gesture。

5. Results

5.1 Experimental Protocol

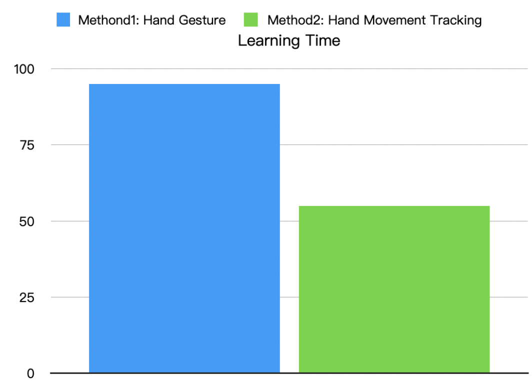
To evaluate our system, we invited 15 participants (2 children, 5 youths, 6 adults, and 2 elderly individuals) to test it and fill out a user satisfaction questionnaire.

The test was divided into two sessions:

1. **Learning Session**:
   In this session, participants are asked to learn and remember to use the gestures with our guidance.
2. **Operating Session**:
   In this session, participants were allowed to operate freely. Participants displayed different gestures in front of the camera and read the corresponding command outputs to check if it is correct (if the needs were met).

5.2 Quantitative Result

Learning Time:
We calculated the average of the learning time for each participant and plotted the graphs.



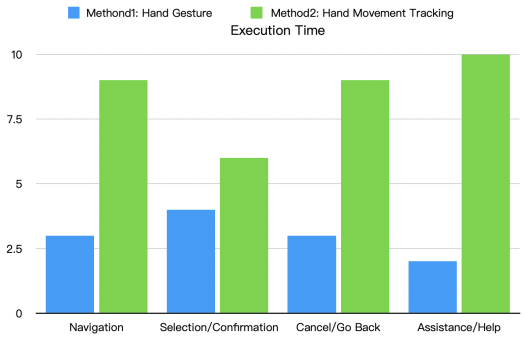
From the figure, we can see that Method1 took longer to learn than Method2, but the difference was small. And overall, the average learning time for both methods does not exceed 100 seconds.

Efficiency:

We calculated the average of the execution time of each gesture for each participant to evaluate efficiency and plotted the graphs.

For the Hand gesture method, we count the time from displaying the gesture to recognizing the gesture and outputting the corresponding command as the execution time.

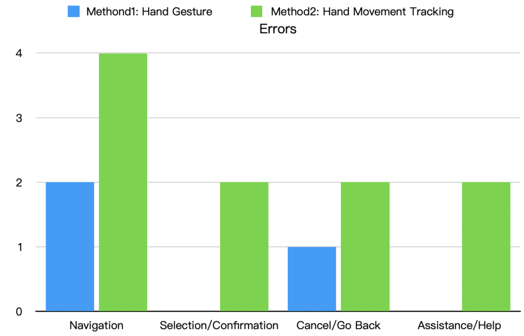
For the hand movement tracking method, we count the execution time from moving the finger to the corresponding position to recognizing the gesture and outputting the corresponding command.



From this figure, we can see that overall, method1 is more efficient than method2. And both of the two methods have a good efficiency with their execution time being no more than 10 seconds.

Accuracy:

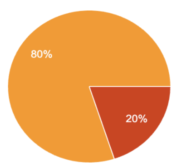
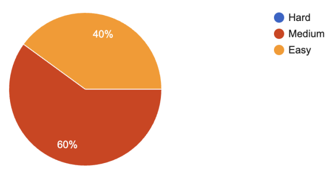
We asked participants to repeat each gesture six times during the test（totally 6*15= 90 times per gesture) and counted the errors of the corresponding gesture execution results. We calculate the number of incorrect executions /total executions per gesture to evaluate accuracy.



From this figure, we can see both of the two methods have a high accuracy, especially method1 in which some gestures have no errors for execution.

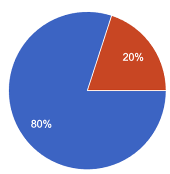
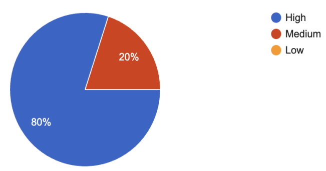
5.3 Qualitative Result

Qn1: How did you find the learning curve of the interaction?

	
Hand Gesture	Hand Movement Tracking

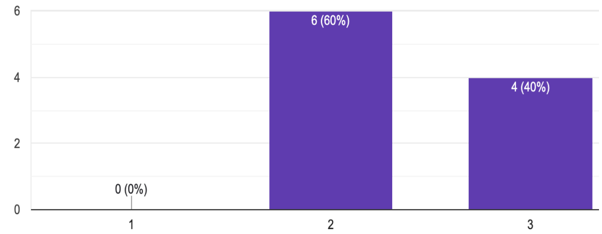
Based on the survey, we discovered that more people(80%) think hand gestures have the easiest learning curve compared to that of hand movement tracking(40%).

Qn2: How accessible do you think the system would be for users with mobility impairments, such as wheelchair users?

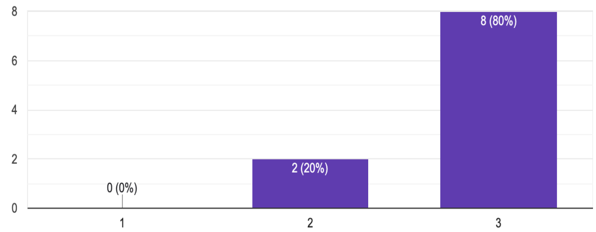
	
Hand Gesture	Hand Movement Tracking

Both interactions have the same level of accessibility based on the participants rating.

Qn3: How intuitive do you find the interaction?



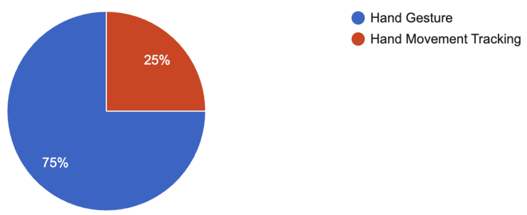
Hand Movement Tracking



Hand Gesture

Based on the survey, 80% of participants found hand gestures to be the most intuitive, compared to 40% for hand movement tracking.

Qn4: Which is your preferred method?



75% of the participants preferred hand gestures over hand movement tracking, as they find hand gestures easier to learn and more intuitive, highlighting the potential of hand gesture control as a viable alternative to traditional touchscreens.

6. Discussion and Conclusion

The results of the quantitative evaluation shows that hand gestures are a better method compared to hand movement tracking, with high efficiency and high accuracy after learning the gestures system under guidance. Also, despite the requirement to learn the gesture system, the gesture learning time is manageable and acceptable.

The results of the qualitative evaluation clearly indicate that hand gestures are a more effective and user-friendly interaction method compared to hand movement tracking. A significant majority of participants (80%) found hand gestures easier to learn and more intuitive, reflecting their natural and seamless integration into user interactions. Additionally, 75% of participants preferred hand gestures, reinforcing their potential as a viable alternative to traditional touchscreens. While both methods were rated equally accessible for users with mobility impairments, the overall preference for hand gestures suggests that they offer a more intuitive and efficient solution for public kiosk interactions. These findings emphasize the suitability of hand gesture control in addressing the limitations of traditional touchscreens, such as accessibility challenges and usability concerns, making it a promising option for enhancing user experience in public environments.

7. References

[1] Freeman, William T., and Craig D. Weissman. "Television control by hand gestures." In *Proc. of Intl. Workshop on Automatic Face and Gesture Recognition*, pp. 179-183. 1995.

[2] Brehme, Daniel, Fabian Graf, Frederik Jochum, and Ioannis Mihailidis. "A virtual dance floor game using computer vision." (2006): 71-78.