Users’ Preferences for Map Navigation Gestures

Y. Y. Pang, N. A. Ismail

Abstract—Map is a powerful and convenient tool in helping us to navigate to different places, but the use of indirect devices often makes its usage cumbersome. This study intends to propose a new map navigation dialogue that uses hand gesture. A set of dialogue was developed from users’ perspective to provide users complete freedom for panning, zooming, rotate, tilt and find direction operations. A participatory design experiment was involved here where one hand gesture and two hand gesture dialogues had been analysed in the forms of hand gestures to develop a set of usable dialogues. The major finding was that users prefer one-hand gesture compared to two-hand gesture in map navigation.

Keywords—Hand gesture, map navigation, participatory design, intuitive interaction.

I. INTRODUCTION

MAPS have become a major part of our lives in helping us to navigate to different locations with ease and reliability. These maps are readily available through websites and applications, but as technology advances, user expectation has also escalated particularly in regards to the abilities of the systems. Traditional input tools, such as keyboard and mouse, are no longer sufficient to meet these demands even with adequate knowledge on the application’s functionality [1]-[3]. Research in the field of human-computer interaction has shown that the use of hand gesture improve the intuitive, natural, and ergonomic interface design [4].

Another challenge lies in the physical and conceptual complexity of the map navigation interface. Most systems designed by experts have neglected how difficult the interface is for people who lack these experiences [1]. There are also researchers focusing on algorithms to improve the accuracy, speed, and robustness of the system without considering the human aspect [3]. These algorithms are only improving the aspects of hand gesture recognitions, not ergonomics to enhance user navigation performance. Human factors should be taken into consideration in designing an interface for map navigation.

The appropriate design solutions to design problems of the participatory design in methodologies are one of the concerns in designing hand gesture interface [5]. Nevertheless, it has to be emphasized that there is no perfect design which is able to fulfil each user’s requirement [4]; every interface has its own behaviour and needs a custom-designed or standardised dialogue to make it intuitive and easy to use for specific tasks [6].

Aside from the little studies published on user-centred interfaces map navigation design, there are apparently insufficient investigations on the modalities preferred by users based on delivered performance. It will be interesting to explore the design and operations of the interface on the usage context and map navigation domain [7]. This has been partly retrieved from the encouraging result from semi-structured interviews in this study. With this, a usable hand-defined hand gesture dialogue for map navigation applications has been developed using the participatory design. This paper presents investigation on user preferred hand in map navigation interfaces, participatory design methodology of hand gesture interaction in map navigation interfaces, and user-defined dialogue for map navigation.

II. RELATED RESEARCH

A. Interactive Map Navigation

Interactive maps allow users to change the map viewpoint alternatively to meet their personal needs and retrieve further information [8]. As such, the usability of maps navigation interfaces depends heavily on their interaction modality and operation. Map interfaces operations are very important in designing the context of use where the operations directly affect the performance of map navigations. Meng [9] has investigated the possible operations of navigation interfaces as listed in Table I. This operation list is more suitable for advance users in the aspect of complexity and redundancy of the operations. For example, it requires the user to study the different characteristic of the maps so that the visualisation parameter can be altered to perform the action. This intention seldom arises in novice users because they only need to view the map, not study the map's characteristic. Kim et al. [10] proposed another version of operations which covered few simple operations including pan, rotate, zoom, tilt, and play tour. When select the operations, the scope can be narrowed down to reduce complexity and save time. Therefore, it is more suitable for novice users.

B. Study of Participatory Design

Nielsen et al. [4] proposed an approach to determine the human gestures through participatory experiment. The gesture sets were tested and claimed to have improved, though limited discussion and evaluation have been reported regarding its subsequently adapted by other researchers to fulfill the requirements for designing gesture applications. For instance, Wobbrock et al. [11] adopted the participatory design approach and proposed a tabletop hand gesture for surface computing and successfully elicited a set of user-defined gestures. They classified the surface gestures into form, nature, binding, and flow. Besides that, they extended their

Y. Y Pang is with the Universiti Teknologi Malaysia, Skudai 81300 Malaysia (corresponding author to provide phone: 601-233-21215; e-mail: pangeeong@gmail.com).
N. A. Ismail was with the Universiti Teknologi Malaysia, Skudai 81300 Malaysia (e-mail: azman@utm.my).
previous work to formulate an agreement score to determine
the degree of consensus for the gesture performed by the
participants [12]. Based on the highest agreement score from
the participants, a set of hand gesture commands was
determined and designed to control the surface computer. The
result only showed that the design gestures were better in
terms of behavioural aspects with no evaluation reported on
the task performance.

TABLE I
THE POSSIBLE OPERATIONS OF INTERACTION BETWEEN THE USER AND MAP

<table>
<thead>
<tr>
<th>Operation</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panning</td>
<td>The user may stepwise move the map towards different directions.</td>
</tr>
<tr>
<td>Zoom</td>
<td>The user may enlarge or reduce the display window without content changes.</td>
</tr>
<tr>
<td>Zooming with</td>
<td>When the user enlarges or reduces the display window, a new level of detail</td>
</tr>
<tr>
<td>LoD</td>
<td>will be rendered which is either pre-calculated or generated in real time.</td>
</tr>
<tr>
<td>Hiding and</td>
<td>The user may visually hide or highlight certain objects or object classes.</td>
</tr>
<tr>
<td>Revealing</td>
<td></td>
</tr>
<tr>
<td>Switching</td>
<td>The user may choose different complementary presentation styles.</td>
</tr>
<tr>
<td>Focusing</td>
<td>The user may click at a certain object and retrieve its detailed information</td>
</tr>
<tr>
<td></td>
<td>stored in a database.</td>
</tr>
<tr>
<td>Tuning of</td>
<td>The user may change his viewing angle and graphic variables in terms of</td>
</tr>
<tr>
<td>visualisation</td>
<td>colour texture, symbol size, figure-ground contrast, and etc. within the</td>
</tr>
<tr>
<td>parameters</td>
<td>allowable value ranges.</td>
</tr>
<tr>
<td>Dialogue</td>
<td>The user may activate a dialogue window and input his personal data.</td>
</tr>
<tr>
<td>Query</td>
<td>The user may search for certain object or object classes by giving one or</td>
</tr>
<tr>
<td></td>
<td>many criteria.</td>
</tr>
</tbody>
</table>

The participatory design aspect also works in three-
dimensional gestures of smart-home control. Kühnel et al. [13]
enacted extensive work on the human psycho-physiological
aspect of smart-home control by adapting the approach of
Wobbrock et al. [11] and Nielsen et al. [4] to apply in a mobile
device. The smart-home control was able to control television,
radio, lighting, and blinds by using seven gestures. The
performance test showed that the gesture commands work
well, but the commands are only limited to mobile phone users
and are not supported in multimodal interfaces.

Several researchers also claimed that the systems using the
participatory design approach are more intuitive and natural to
users compared with the application design from designer's
perspective [14]. Nacenta et al. [14] mainly focused on the
comparison between the memorability of the predesigned
gesture sets, user-defined gesture sets, and random gesture
sets. The predesigned gesture sets were created by designers
for a particular application. From the user-centered
experiment, the user-defined gesture sets were analysed and
designed while random gesture sets were assigned the
corresponding action from the database of gestures. In their
findings, they discovered that the user-defined gestures are the
easiest to remember.

Recent studies showed that the participatory design
attracted more research attention in designing the gesture
system. Participants who collaborate and contribute in the
participatory design process tend to give more articulate and
creative process [15], but there are limited researches focusing
on designing applications based on the participatory design
aspect. It is largely an underexplored domain.

C. Hand Gesture Classification

The use of gesture is a common way of communication in
human conversation and is also meaningful during the
communication between human and computer. In particular,
hand gestures can be classified to compromise the context of
use. Poggi [16] and [17] investigated the gestures and it
relations with speech. Though not comprehensive, their
classification schemes complement each other. For example,
Poggi [16] discussed the deictics and iconics that can
complement the manipulative, semaphoric, and conversational
gesture of Quek et al. [17]. Karam and Schraefel [18]
proposed a more comprehensive gesture classification based
on human computer interaction that has become the
fundamental of researches in the domain of hand gesture
analysis. However, as argued by Aigner et al. [19], it still fails
to cover the distinction between iconic and pantomimic
gesture. Hence, the work has been extended to the semaphoric
class of Karam and Schraefel [18] gesture taxonomy. In their
gesture classification scheme, they have refined details in
gesture types including pointing, semaphoric, pantomimic,
iconic, and manipulation; these are the basics of hand gestures
classification in this paper. Table II summarizes the
classification scheme in this research.

III. OVERVIEW OF THE PROPOSED METHOD

A. Interview

The preliminary requirements are to gather the required
context of use and operations of the map navigation
applications that are adequate as input for multimodal
interaction dialogues. In the interview, the interviewees
were prompted with questions in a structured manner while their
feedbacks were recorded to investigate their favourite
functions of using maps. The related operations of the map
navigation application were provided after the relevant data
had been identified through response analysis to initialise a set
of operations. The data were formulated to become
meaningful tasks and then assigned to users during the
subsequent participative design stage.

B. Participatory Design

This stage uses the participatory design approach to collect
hand gesture dialogues, which has been done through
conducting user experience design in an inductive approach. A
group of participants were tested and guided through scenarios
in a set up room (with camera surveillance, monitor, and
microphone). They were asked to imagine themselves
communicating with the computer that could understand their
command when performing the previously identified
operations from the previous stage. In order to observe the
participants' natural and intuitive behaviour while navigating
through the maps, the gaps of technical feedback were
removed from the dialogue between user and device to avoid
in-depth thinking [11], [20]. This also meant that recognition
response was not expected when the participants were
demonstrating their hand gesture dialogue. The participants
were also advised to assume that the system was able to understand their dialogue.

### TABLE II
**CLASSIFICATION SCHEME FOR HAND GESTURE ON HUMAN COMPUTER INTERACTIONS**

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deictic</td>
<td>Deictic is used to indicate spatial location of an object. It may either be performed by pointing a stretched index finger or with multiple fingers such as the combination of index and middle finger. Manipulative gestures apply the tight relationship between the movements of the gesturing hand and the object being manipulated to the command object. Visual feedback for controlling the targets is typically shown in manipulation gesture as a guide of being manipulated.</td>
</tr>
<tr>
<td>Semaphoric</td>
<td>Semaphoric gestures are signalling systems which convey specific meanings of the gesturing hand poses and movements. The gesture may be any unrelated hand gestures which convey meanings. Semaphoric gestures can be further categorized into static poses, dynamic movements, and strokes. Static poses are static hand poses such as stretching all the fingers to mean “stop”. The dynamic movements of semaphoric can refer to information delivered through the temporal. An example would be to twist or turn the wrist like opening a faucet gesture, which means “rotate”. Stokes gesture resembling hand flicks which have similarity with dynamic movements. Stokes gesture may be familiar with slide gestures that are mostly applied in smartphones. For example, a single hand swinging to sideward means “next”. Pantomimic gestures are actions to depict or imitate the real world objects using hand gesture. Example of pantomimic would be imaginary of capturing photograph using both “L” hand shape. The process can be coded as one pantomimic gesture such as vertical both “L” hand shape in the opposite position, then moving closer, and releasing back to the first action.</td>
</tr>
<tr>
<td>Pantomimic</td>
<td>Pantomimic gestures are actions to depict or imitate the real world objects by performing specific sizes, pattern, and movements. Static iconics are demonstrated by a static hand pose and do not necessarily reflect the real world objects. Dynamic iconics are used to depict the shapes; for instance, drawing a circle may mean “rotate”.</td>
</tr>
<tr>
<td>Iconic</td>
<td>Iconic gestures deliver the information of the objects by conveying specific sizes, pattern, and movements. Static iconics convey the corresponding meanings of the gesturing hand poses and dynamic iconics can be shown to manipulate the command object. Visual feedback for controlling the targets is typically shown in manipulation gesture as a guide of being manipulated.</td>
</tr>
</tbody>
</table>

**C. Participants**

Before starting, each participant was asked to fill in a questionnaire on their background and preference on technology devices. Since most participants were young adults (aged 14 to 39), they were expected to possess fundamental cognition knowledge on how to use gesture-based devices and might perform gestures similar to those devices. This age group also has the highest smartphone ownership among all ages [13].

**D. Procedures of Experiment**

The hand gesture activity perform by participant was recorded by a video recorder. In this experiment, it was used to analyse the hand gestures dialogues. A series of videos that illustrated a scenario of tasks that should be carried out was played to the participants. The experiment was conducted in a private room as in Fig. 1 to avoid any unexpected disturbance. Only the participant and observer were allowed to stay in the room once the experiment started.

The purpose of the study and procedures were explained to the participant beforehand. Within the group, factorial design was applied where each participant was given two conditions that were independent variables. The first involved interaction with hand gesture using one hand and the second was about hand gesture interaction with two hands. They were requested to develop and demonstrate self-dialogue set for map navigation. A total of five tasks were divided into 21 subtasks; think aloud protocol was applied to the participants and video recorded. They were provided subjective preference ratings for each task. The experiment was conducted in the following order:

**Step 1.** The researcher briefed and asked the participant to the pre-test questionnaire.

**Step 2.** The participants were given a document and information sheet describing each task listed into two categories.

**Step 3.** The participants were instructed to design their self-defined hand gesture dialogue.

**Step 4.** After designing the overall set of dialogue for every task, the participants were requested to demonstrate each dialogue for five times while thinking aloud to represent each task. For example, when the participant intended to zoom into the map, he or she might say, “I am zooming into the map to view the dept.” with a parallel demonstration of the hand gesture.

**Step 5.** After completing each task, the participants were asked to rate the dialogue using a 7-point Likert scale concerning the goodness and ease of use of the hand gesture.

**Step 6.** During the debrief section, the participants were asked to suggest other tasks where the dialogue would be beneficial and rate the number of preferred hand gesture in a post-questionnaire sheet.

**E. Definition of Dialogue Sets**

This stage focused on developing usable vocabulary for the operations of map navigations after the multimodal interaction vocabulary had been collected in the previous stage. The usable vocabulary sets denote meaningful commands that can be used to navigate maps.

Based on the agreement score, A, of Wobbrock et al. [11],
the match hand gestures of the participants can be translated into reliability. The mathematical equation for agreement score is:

\[ At = \sum_{Pm \in Pt} \left( \frac{Pm}{|Pt|} \right)^2 \]  

where \( At \) is the agreement score of a task in a single number and level of consensus among the participants; \( t \) is a task in all tasks; \( Pt \) is the set of proposed hand gestures for \( t \); and \( Pm \) is a subset of matched hand gesture from \( Pt \). The range of \( A \) is \([0, 1]\).  

IV. RESULTS AND ANALYSIS

A. Result of Preliminary Requirements Gathering

There were 12 voluntary participants with average age of 24 (standard deviation, \( SD = 2.79 \)) who took part in this interview. In the interviewees' satisfaction feedback on the map navigation operations (see Fig. 2), six interviewees were satisfied with the current map navigation operations. The number of interviewees who were slightly satisfied, slightly dissatisfied, and had a neutral stand were two each. Since majority of them were satisfied with the current map navigation operations, no modification was carried out. An interviewee was concerned about the complexity of the map navigation, so slight adjustment on the selection of operations was made. Result showed that most of the operations mentioned were similar with Kim, et al. [10]. The information from the participants was further analysed. With this, the operations that had been tailored for the map navigations were justified (see Table III).

B. Results Analysis of Participatory Experiment Design

The participatory experiment design had 13 voluntary participants, nine males and four females, with an average age of 28 (\( SD = 2 \)). Participants with less experience on gesture-based technology devices were included in the experiment. The intention was to balance data collection from the natural and intuitiveness of participants. There were 286 dialogues (13 participants x 11 tasks x 2 x categories) collected from the study. The dialogues were analysed and qualitative observations were used. The observed hand gestures performed by the participants were classified based on the classification scheme in Table II and as depicted in Fig. 3. The combination of more than one type of element was found in each task. For example, it was noticed that participants would first point at the location to zoom and then stretch their hands to further zoom in. The gestures used here were deictic gesture (applied when pointing at the location) then followed by manipulating gesture (when stretching the hand). However, the majority of gestures presented in the tasks were of a single gesture type only since multiple gestures might be confusing and complex.

C. Results Analysis and Defining of Dialogues Sets

The agreement score for the hand gesture performed by all participants is depicted in Fig. 4. The average agreement score for one-hand gestures and two-hand gestures were 0.41 and 0.26 respectively.

Every time after the participants had expressed the dialogues for each task, they were requested to rate it on two Likert scale for one-hand and two other Likert scales for two-hands. The scales were from strongly disagree (-3) to neutral (0) to strongly agree (3). The major items for rating were: (1) “The hand gesture I chose is a good match for the intended purpose” and (2) “The hand gesture I picked is easy to perform”. These subjective ratings were then compared and it was discovered that there was a significantly positive correlation between the good match of the dialogues and its ease to perform (\( r_s = 0.791; p < 0.01 \)). Simply to say, the participants felt that the better the dialogue could match the function, the easier it was to perform the hand gesture.

During the debriefing session, the participants were asked for their suggestions on improvement and to rate their personal preference on the number of hand that should be used for map navigation on a 5 point scale from neutral (0) to strongly prefer (5). A Wilcoxon Signed Ranks Test was conducted to compare these subjective ratings, and it was found that one-hand gesture was more preferred (\( Z = -2.887; p < 0.005 \)). The participants found it easier to relate and use one-hand gesture. This personal preference has been reflected in the rating of good match (\( Z = -4.656; p < 0.001 \)) of vocabulary and its ease to perform (\( Z = -4.843; p < 0.001 \)) between the one-hand and two-hand gestures in map navigation.

### TABLE III

**SUMMARY OF MAP NAVIGATION TASKS**

<table>
<thead>
<tr>
<th>Operation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Get the direction from a location to another location.</td>
</tr>
<tr>
<td>Zoom in</td>
<td>Zoom the viewer in more details.</td>
</tr>
<tr>
<td>Zoom out</td>
<td>Zoom the viewer in less detail.</td>
</tr>
<tr>
<td>Pan left</td>
<td>Shifts the viewer to the left.</td>
</tr>
<tr>
<td>Pan right</td>
<td>Shifts the viewer to the right.</td>
</tr>
<tr>
<td>Pan up</td>
<td>Shifts the viewer to the above.</td>
</tr>
<tr>
<td>Pan down</td>
<td>Shifts the viewer to the below.</td>
</tr>
<tr>
<td>Rotate left</td>
<td>Rotates the view counter-clockwise.</td>
</tr>
<tr>
<td>Rotate right</td>
<td>Rotates the view clockwise.</td>
</tr>
<tr>
<td>Tilt up</td>
<td>Incline the view of the map horizontally.</td>
</tr>
<tr>
<td>Tilt down</td>
<td>Incline the view of map “top-down”</td>
</tr>
</tbody>
</table>
In this experiment, the dichotomous referents including zoom in/zoom out, pan up/pan down, pan left/pan right, rotate left/rotate right, and tilt up/tilt down had fulfilled the reversible gestures [11].

Table IV shows the proposed dialogues. This formalised and consistent set of dialogues has been defined based on all collected hand gesture interactions from the previous stages and the participants’ preference on one-hand gesture.

**V. SUMMARY AND DISCUSSION**

In this paper, the methodology for surface computing interaction from Nielsen, et al. [4] and Wobbrock, et al. [11] and the methodology for air hand gestures from Aigner, et al. [19] were adopted to design the map navigation interface in a user-centered approach. The findings suggested that the usage of one hand gesture interaction is preferable by users compare with two hand gesture for map navigation. The experimental design of this research has outlined the methodology to collect the data from the participants and analysed the attributes from the observation to the independent variable.

The user experience gained from the participatory design is distinct from ordinary interface experts. For example, designers who design the interface based on their expertise go through the iteration process of user testing for their created interface. The user-centered approach which develops the interface from the user can cut down the cost and ineffectiveness of the former approach. However, the experimenters must ensure that the participants’ perceptions are held constant across the groups throughout all phases and conditions of the experiments. This step is to lower the risk of human errors that may affect the precision of results collected.

The statistical analysis showed that the users prefer one-hand dialogue that is easy to perform and related to the real world environment. Defining usable dialogue from user-defined is new strategies in design applications. In future, we will examine the effects of the usable dialogue sets on humans’ performance. The participatory design not only increases the user experience, it is also beneficial for expert designs. Besides that, cultural differences may be one of the limitations in this study Abadi and Peng [21].

Overall, we have contributed to the development of user-defined dialogue sets for map navigation interface where most other researches are focused on expert designing dialogues sets. We also introduced information about how to extract users define dialogue from user-centered approach under each of processes that we proposed, which can benefits to interfaces design other than map navigation applications. This study leads us toward the future direction of interface design based on user center to suit to the nature of interaction.
## Table IV: Definition of Usable Vocabulary for Map Navigation Interfaces

<table>
<thead>
<tr>
<th>Operation</th>
<th>Description</th>
<th>Dialogue sets sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deictic</td>
<td>Direction: point at origin, move and point destination</td>
<td></td>
</tr>
<tr>
<td>Zoom</td>
<td>Manipulation gesture is selected to use in zoom task. Pinch fingers to represent “zoom out” and splay fingers to represent “zoom in”.</td>
<td></td>
</tr>
<tr>
<td>Rotate</td>
<td>Semaphorics dynamic gesture is used to design rotating tasks. Twist the wrist to represent rotating.</td>
<td></td>
</tr>
<tr>
<td>Pan</td>
<td>The semaphoric strokes like flick are applied in designing the pan task.</td>
<td></td>
</tr>
<tr>
<td>Tilt</td>
<td>The manipulation gesture is adopted to represent tilt. For example, scroll hand backward to tilt up.</td>
<td></td>
</tr>
</tbody>
</table>

## Acknowledgment

This work was supported by the Research University Grant (Vote No. Q.J130000.2501.09H09), Universiti Teknologi Malaysia.

## References


