International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: http://www.tandfonline.com/loi/hihc20

User-Defined Gestures for Gestural Interaction: Extending from Hands to Other Body Parts

Zhen Chen, Xiaochi Ma, Zeya Peng, Ying Zhou, Mengge Yao, Zheng Ma, Ci Wang, Zaifeng Gao & Mowei Shen

To cite this article: Zhen Chen, Xiaochi Ma, Zeya Peng, Ying Zhou, Mengge Yao, Zheng Ma, Ci Wang, Zaifeng Gao & Mowei Shen (2018) User-Defined Gestures for Gestural Interaction: Extending from Hands to Other Body Parts, International Journal of Human–Computer Interaction, 34:3, 238-250, DOI: 10.1080/10447318.2017.1342943

To link to this article: https://doi.org/10.1080/10447318.2017.1342943

Accepted author version posted online: 23 Jun 2017.
Published online: 30 Aug 2017.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
2018, VOL. 34, NO. 3, 238–250
https://doi.org/10.1080/10447318.2017.1342943

User-Defined Gestures for Gestural Interaction: Extending from Hands to Other Body Parts

Zhen Chen, Xiaochi Ma, Zeya Peng, Ying Zhou, Mengge Yao, Zheng Ma, Ci Wang, Zaifeng Gao, and Mowei Shen

Department of Psychology and Behavioral Sciences, Zhejiang University

ABSTRACT

Most gestural interaction studies on gesture elicitation have focused on hand gestures, and few have considered the involvement of other body parts. Moreover, most of the relevant studies used the frequency of the proposed gesture as the main index, and the participants were not familiar with the design space. In this study, we developed a gesture set that includes hand and non-hand gestures by combining the indices of gesture frequency, subjective ratings, and physiological risk ratings. We first collected candidate gestures in Experiment 1 through a user-defined method by requiring participants to perform gestures of their choice for 15 most commonly used commands, without any body part limitations. In Experiment 2, a new group of participants evaluated the representative gestures obtained in Experiment 1. We finally obtained a gesture set that included gestures made with the hands and other body parts. Three user characteristics were exhibited in this set: a preference for one-handed movements, a preference for gestures with social meaning, and a preference for dynamic gestures over static gestures.

1. Introduction

Body gestures in human–computer interaction (HCI) allow users to interact with a device in an intuitive and easy manner (Norman, 2010; Pang et al., 2014; Tian, Lyu, Zhang, Ren, & Wang, 2016). For instance, we can play a Kinect-based computer game by freely using our hands, arms, and legs. Compared with traditional interaction devices (e.g., mouse and keyboard), body gesture interaction is more natural and comfortable for users (Nacenta, Kamber, Qiang, & Kristensson, 2013; Pereira, Wachs, Park, & Rempel, 2015) and is less constrained (Pantic, Pentland, Nijholt, & Huang, 2007). Body gesture interaction is even considered to be the next generation of the computer mouse (Norman, 2010; Pang et al., 2014).

Although body gesture interaction is a natural means of interaction, gestures in HCI at present are unnatural and are difficult to learn and memorize (e.g., Liang, 2013; Norman, 2010; Norman & Nielsen, 2010; Pang et al., 2014; Zhao et al., 2014). To a large extent, these problems occur because the gestures have been defined by designers without fully considering the compatibility between gestures and commands. It has been well acknowledged that designers and users typically have distinct mental models of interaction (Wickens, Hollands, Banbury, & Parasuraman, 2015). Users' knowledge of the world (including perception of physical laws, experience with HCI, and customs) has a comprehensive impact on their preferences for body gestures for interaction. Therefore, to obtain high degree of user experience, it is important to take users' mental models into consideration when developing body gestures for gestural interaction (Wickens et al., 2015).

At present, one of the most effective ways of generating user-friendly body gesture sets is via a user-defined approach. In particular, researchers first gathered a gesture set by using certain visual effects to explain to the user the wanted effect within the system, and then requiring the user to perform at least one gesture that should trigger this effect. Finally, researchers determined gesture candidates by checking all user-proposed gestures for a specific effect, and calculating an agreement score according to how often the same gestures were proposed (Nielsen, Störring, Moeslund, & Granum, 2003; Wobbrock, Aung, Rothrock, & Myers, 2005; Wobbrock, Morris, & Wilson, 2009). This approach was first proposed by Wobbrock et al. (2009) to generate a 2D hand gesture set for touchscreens and was later extended to generate 3D gesture sets for environments such as smart homes (Choi, Kwon, Lee, Lee, & Chung, 2012; Kühnel et al., 2011), and for virtual reality (Wu, 2013) and augmented reality (Piumsomboon, Clark, Billinghurst, & Cockburn, 2013) systems. With these gesture sets, users consistently reported that their experience significantly improved in terms of compatibility, memorability, learnability, and comfort (e.g., Nacenta et al., 2013; Vatavu & Zaiti, 2014; Wu, 2013), resulting in high preference levels (Vatavu & Zaiti, 2014). However, past gesture-elicitation studies have predominantly focused on hand gestures, ignoring the involvement of other body parts in gestural interaction (e.g., Baudel & Beaudouin-Lafon, 1993; Bowman & Wingrave, 2001; De La Barré, Chojecki, Leiner, Mühlbach, & Ruschin, 2009; Kühnel et al., 2011; Nielsen et al.,

CONTACT Mowei Shen mwshen@zju.edu.cn; Zaifeng Gao zaifengg@zju.edu.cn Department of Psychology and Behavioral Sciences, Xixi Campus, Zhejiang University, Hangzhou 310007, P.R. China.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc

2003; Pereira et al., 2015; Piumsomboon et al., 2013; Vatavu, 2012; Vatavu & Zaiti, 2014; Vogel & Balakrishnan, 2005; Wright, Lin, O'Neill, Cosker, & Johnson, 2011; Wu, 2013; Wu & Wang, 2012; Zaiți, Pentiuc, & Vatavu, 2015). One key reason for this was the relatively small size of typical interactive displays (e.g., Microsoft Surface). Hand gestures require less space and can be easily and flexibly performed to meet different purposes; hence, they are particularly suitable for small-screen displays. Further, involving gestures of other body parts usually requires a large space and a large-screen display and consumes more physiological energy than hand gestures, causing physical fatigue after long-term tasks.

However, we believe that it is important to extend our exploration from hand gestures to non-hand gestures. First, we often use hands as well as other body parts to express an intention or fulfill an aim in our daily lives. Ignoring these non-hands gestures, therefore, may lead to the neglect of certain valuable means of communication. Moreover, interacting with a computer without body part constraints can increase the level of immersion for short-term tasks (e.g., exercise gaming). Second, owing to the rapid development of ubiquitous computing and smart homes, large-screen displays have become widely available; this has provided users with more opportunities to use non-hand gestures. Meanwhile, current motion-capture technology can successfully and efficiently track whole-body movement (e.g., Microsoft^{TM} Kinect), allowing users to interact with the computer in a more complete manner. Third, relying heavily on hand gestures may increase the memory load of users due to similarities among the gestures (e.g., Alvarez & Cavanagh, 2004; Gao et al., 2009); incorporating gestures of other body parts will increase gesture variety and hence reduce the potential memory load. Finally, yet critically, gestural interaction aims at making the users feel more "natural" during gestural interaction. Constraining the interacting mode to hand gestures will inevitably lead to an unnatural feeling under certain circumstances, particularly considering that non-hand gestures are effective for relaxing (e.g., playing games). Further, incorporating gestures of other body parts into the interaction does not exclude hand gestures; instead, it means that the participants can choose the most suitable gesture to fulfill the task. In this manner, users can take advantage of both hand and non-hand gestures. Therefore, it is necessary to consider the movement of other body parts when developing a gesture set for gestural interaction. The few studies that have considered the contribution of non-hand body parts to gestural interaction explored certain specific circumstances, such as intense gameplay and TV (Dim, Silpasuwanchai, Sarcar, & Ren, 2016; Silpasuwanchai & Ren, 2015). To the best of our knowledge, no study has attempted to offer a gesture set that can be used in various interactive situations in which all body parts are considered. Through the user-defined method, in this study, we provide such a solution for the most frequently used gestural interaction commands. Critically, to find the most suitable gestures for a specific command, the participants were told that they could make gestures using any body part. In this case, we can directly examine user preferences and determine the necessity of considering all body parts in future gestural interaction.

Although the user-defined method has several advantages with regard to finding gestures that match users' mental models, there are two drawbacks to it. First, most studies on the user-defined method use the frequency of the proposed gestures as the main index (taking the gesture with the highest frequency as the aimed gesture), without considering the participants' subjective assessments (e.g., Kühnel et al., 2011; Piumsomboon et al., 2013; Vatavu, 2012; Vatavu & Zaiti, 2014; Wright et al., 2011; Wu, 2013). As a result, the gesture set obtained may be not the users' favorite, or the most suitable, as corroborated by Choi et al. (2012). They generated a gesture set for a smart home via the user-defined method with the frequency of gestures as the main index and then asked the same group of participants to perform a gesture-command match task using previously generated gestures. It was found that 65% of the high-scoring gestures from the initial experiment were altered in the second experiment, the agreement score of which was also significantly higher. To compensate for such defects, users' subjective assessments have been considered in recent gesture-elicitation studies (e.g., Kühnel et al., 2011; Piumsomboon et al., 2013; Vatavu & Zaiti, 2014; Wu, 2013). However, they were only used as a supplementary index to support the user-defined gesture set, rather than as a core index. Pereira et al. (2015) took objective, subjective, and biomechanical factors into consideration for the first time in revealing optimal gesture-command matches in a 3D hand gesture-elicitation study; this was achieved by using frequency, subjective ratings, and posture risk rating as core indices. However, several limitations persist in this method, particularly with regard to subjective ratings; for each command, each user only assessed two of his/her proposed gestures without learning gestures proposed by other users. Consequently, it is possible that the gestures proposed by a few users actually recorded higher gesture-command compatibility, as Choi et al. (2012) pointed out. Another drawback of the user-defined method is that users often found it difficult to create more gestures for target commands, particularly when they were not familiar with the design space. To address this issue, researchers recently suggested that in addition to the user-defined procedure (e.g., Dim et al., 2016; Silpasuwanchai & Ren, 2015), it is necessary to further show users a pre-defined list of representative gestures and require them to select the most suitable one.

In this study, we aimed to generate a gesture set for the most frequently used gestural interaction commands, while taking non-hand gestures into consideration. We achieved this by combining gesture frequency, subjective ratings, and physiological risk ratings as core indices. We first collected candidate gestures in Experiment 1 through the user-defined method, in which we required the participants to propose gestures without body part limitations. In Experiment 2, we further recruited a group of participants to subjectively assess all representative gestures (i.e., those proposed by more than two people) obtained in Experiment 1. Based on the gesture sets obtained in Experiments 1 and 2, we proposed a gesture set for commonly used gestural interaction commands.

2. Experiment 1: Generating a gesture set throughthe user-defined method

2.1. Method

Participants

Twenty-two undergraduates (12 male) were recruited from Zhejiang University. All participants had normal color vision and normal or corrected-to-normal visual acuity. Participants were 20.73 (SD = 1.42) years old on average and were all right-handed. This research complied with the American Psychological Association Code of Ethics and was approved by the Research Ethics Board of the Department of Psychology, Zhejiang University. Informed consent was obtained from each participant. Eleven participants had never used a Kinect before, and the remaining participants had Kinect experience of no more than 14 hrs.

Tasks

To determine the most frequently used gestural interaction commands and guarantee the feasibility of gestures for gestural interaction, we first listed all the commands used in past 3D gestural studies, resulting in a total of 45 commands (removing tasks related to specific contexts, such as TV and games). Then, four expert users of Kinect-based interaction, who were very familiar with the Xbox 360 (gesturally interacting with Xbox 360 at least four hours per day (focusing on gaming, but also enjoying other Xbox functions) and five days per week, with total experience of at least 200 hrs), ranked the commands by their importance and frequency in gestural interaction. We selected the commands with importance scores above 3 (of at most 5), and those with "being-explored" frequencies of greater than two in previous elicitation studies; we removed the commands that were used with GUI interfaces but were obsolete in the gestural interface. This procedure finally led to 15 representative commands for further exploration (see Table 1). To further examine the representativeness of the chosen commands, we (tried our best) surveyed 31 people (through an online questionnaire) familiar with gestural interactions within almost 2 months. Most of them agreed that the selected commands were important

Table 1. Fifteen typical commands used in current study.

Command	Description
Accept	Accept a suggestion/command offered by the system.
Reject	Reject a suggestion/command offered by the system.
Gesture on	System starts to capture gestures.
Gesture off	System stops capturing gestures.
Open	Open a file or a program.
Close	Exit from an on-going process.
Volume up/down	Increase/decrease the volume of system.
Zoom in	Zoom into a certain area of the interface.
Zoom out	Zoom out of a certain area of the interface.
Back	Return to a previous interface.
Delete	Delete a file.
Help	Open FAQ to get help.
Menu access	Access the main menu.
Move	Move an icon from one position to another.
Pan	Display the follow-up list.

(M = 3.53, SD = 1.42; full score was five), and no extra commands were added.

Experimental setup

An overview of the experimental setup is shown in Figure 1. The participants were instructed to stand 160 cm from a Microsoft Kinect 1.0 sensor and a Kinect 2.0 sensor, which were placed in front of a 48-inch LCD screen. The Kinect 1.0 was used to track and record the skeletal data (see 2.1.5), while the Kinect 2.0 was used to videotape the participants' proposed gestures for offline checking when analyzing the oral descriptions of the gestures. A computer 100 cm from the left-rear of the participants was connected to the RGB camera of the Kinect sensors, such that the users' actions could be tracked and videotaped during the experiment. Following previous user-defined studies (e.g., Pereira et al., 2015; Wobbrock et al., 2009), we showed each task by displaying before-and-after images on the monitor. These images were screen snapshots of the interface on an Xbox 360 and helped participants understand the task commands. Two experimenters sat 100 cm from the left-rear of the participants: one experimenter was in charge of running the experiment and the gestural video recordings, while the other observed the gestures performed and wrote down participants' oral descriptions of the gestures. The overall setup offered enough space for the participants to perform any

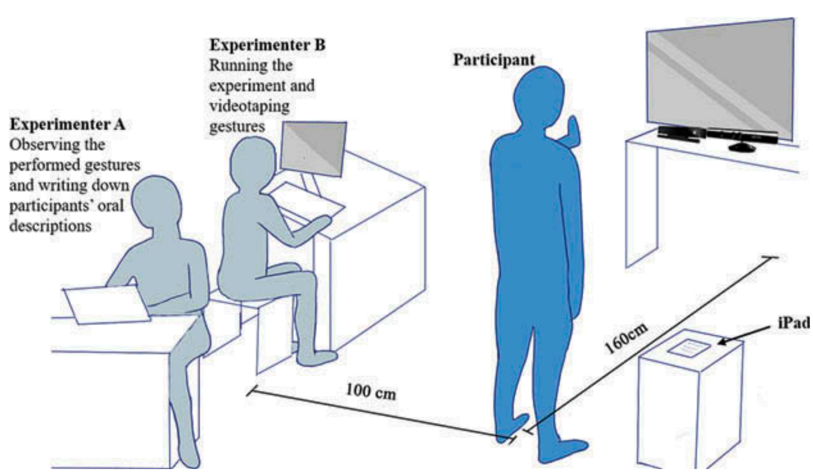

Figure 1. Setup used in Experiment 1.

intended gesture, and the participants did not hit any elements/experimenters during the experiment.

During the formal experiment, instructions for the task to be conducted were first presented on the screen; the experimenter also explained instructions to the participants to ensure that they had been fully understood. Additionally, participants were informed that the gestures they proposed would be repeatedly used in future gestural interaction applications. A task interface was then presented on the screen; this interface displayed images from before and after the task being conducted, with the before-image right above the after-image. An extra task description (in Chinese) was added to the left of the images to inform the participants about the command. Participants were required to make at least two gestures for each task to avoid legacy bias, which refers to a trend wherein "users' gesture proposals are often biased by their experience with prior interfaces and technologies, particularly the WIMP (Windows, icons, menus, and pointing) interfaces that have been standard on traditional PCs for the past two decades" (Page 42, Morris et al., 2014). They were instructed to think aloud while performing the corresponding gestures in turn without a time constraint and to confirm the start and end of their gestures. Once all gestures had been completed, the participants were required to conduct a subjective assessment (see section 2.1.3) for each proposed gesture and finally rank all proposed gestures according to their preferences (from best to worst). The four subjective assessment items were displayed simultaneously on an iPad. The 15 tasks were presented randomly, and the entire experiment was video recorded. Before the formal experiment, participants carried out two trials to familiarize themselves with the procedure. The entire experiment lasted approximately 45 minutes.

Subjective ratings

On the basis of past research (Igbaria & Nachman, 1990; Kirakowski, Claridge, & Whitehand, 1998; Lewis, 1995, 2002; Lund, 2001; Pereira et al., 2015; Sharp, 2003; Wobbrock et al., 2009), we collected the subjective ratings of the proposed gestures from five aspects: learnability, match, effort, subjective fatigue, and preference. Following the completion of a gesture, participants were shown a seven-point Likert scale (1 = completely disagree, 7 = completely agree) to evaluate the first four aspects in turn: "the gesture I performed can be learned quickly," "the gesture I performed is easy to perform," "the gesture I performed is a good match for its purpose," and "the gesture I performed is tiring." Once all gestures for a task were completed, the participants were required to rank the proposed gestures according to their preferences (from best to worst).

Physiological risk rating

Because this study allowed participants to make full-body gestures, we used Rapid Entire Body Assessment (REBA; Hignett & McAtamney, 2000) to evaluate the potential physiological risk embedded in the proposed posture. REBA considers the postures of the trunk, neck, legs, upper arms, lower arms, and wrists. Although there is apparently some overlap between the contents of REBA and the subjective ratings of fatigue, they reflect different dimensions of physical movement. In particular, REBA considers the posture risk of work-related musculoskeletal disorders following the repeated performance of a gesture for a relatively long period, whereas subjective fatigue measures how a user feels immediately after performing a gesture. Indeed, our subsequent analysis from Experiment 1 confirmed that there was no significant correlation between REBA scores and subjective fatigue ratings (r = .176, p = .472). When using REBA, it is important to accurately measure the highest degree of movement of each body part involved in a gesture. We achieved this by analyzing users' skeletal data via a second Kinect 1.0, which was placed near the Kinect 2.0. It has been shown that Kinect is a fast and reliable motion-capture system in tracking human movements of the trunk, neck, shoulders, elbows, legs, and arms (e.g., Choppin, Lane, & Wheat, 2014; Clark, Pua, Bryant, & Hunt, 2013; Diego-Mas & Alcaide-Marzal, 2014; Galna et al., 2014; Schmitz, Ye, Shapiro, Yang, & Noehren, 2014; Van Diest et al., 2014) and hence can be used to measure workload (Diego-Mas & Alcaide-Marzal, 2014; Patrizi, Pennestri, & Valentini, 2016). Moreover, Kinect data are more accurate when the tracked body orientations and positions are from a user who is facing the sensor and performing simple, low speed motions (Choppin et al., 2014; Diego-Mas & Alcaide-Marzal, 2014), which was met by the current gesture elicitation situation. Using a homemade script written in C#, we searched for and recorded the highest degree of movement for each body part. It is worth noting that the movement of the wrist belongs to fine movement, and Kinect does not work well under this situation (Galna et al., 2014). We overcame this issue by looking into the video recorded by the Kinect 2.0 and computing the moving degree manually. On the basis of these results, we obtained the REBA scores.

2.2. Data analysis

The amplitude and position of gestures differ from person to person, leading to large variations in the proposed gestures. Consequently, gesture classification was required before further analysis. Two experts (the second and the fifth authors) first conducted gesture classification independently, based on videotapes and participants' descriptions of the gestures. Gestures were considered to be the same if they contained the same or similar motions or postures. Motion mainly comprised the direction of joint motion (according to the oral descriptions of the participants, there were mainly three moving directions: vertical, horizontal, and diagonal) and synchronized body-part movements. Posture referred to the end postures for all involved body parts. Additionally, in some cases, participants made essentially the same gesture for a specific command yet used different hands or feet; we categorized these gestures as the same type. The classification consistency between the two experts was 93.23%. Any divergence in classification between them was resolved by asking a third expert (the first author). We finally obtained 710 distinct gestures in total, nine of which were removed because either the participants confused the meaning of the task, or the gestures were difficult to recognize in practical usage. This resulted in 701 valid gestures.

To evaluate the degree of gestural consensus among the participants, we calculated the gesture agreement rate (AR) for a command using the formula of Vatavu and Wobbrock (2015):

$$\mathrm{AR(r)} = \frac{|P|}{|P|-1}\sum_{P_i \subseteq P}\left(\left|\frac{P_i}{P}\right|\right)^2 - \frac{1}{|P|-1}$$

where P is the total number of proposed gestures for a command r, and P_{i} is the number of a subset i of identical gestures from P. The values of AR range from zero to one. Moreover, Vatavu and Wobbrock proposed qualitative interpretations for AR (see Table 6 in Vatavu & Wobbrock, 2015): AR ≤ 0.100 implies a low agreement, 0.100 < AR ≤ 0.300 implies a moderate agreement, 0.300 < AR ≤ 0.500 implies a higher agreement, and AR > 0.500 signifies very high agreement. It is worth noting that when AR is equal to 0.500, the agreement and disagreement rates are equal.

Finally, an overall command-gesture score was calculated by summing the values of the seven variables (cf. Pereira et al., 2015): frequency of a specific gesture (i.e., popularity), learnability, match, effort, subjective fatigue, REBA, and preference ranking. Prior to summation, each of these variables was normalized (with mean = 5.0 and standard deviation = 1).

2.3. Results

The results contained gesture taxonomy, the agreement rate, and the selected task gesture set according to an adjusted overall score.

Gesture taxonomy

We classified the gestures along four dimensions: form, nature, binding, and body parts. They were mainly based on the taxonomy used by Wobbrock et al. (2009) and were adapted to match full body gestures. The first three dimensions are from Wobbrock et al. (2009) and have been commonly used in user elicitation studies since the seminal work of Wobbrock et al. (2009) (e.g., Dim et al., 2016; Piumsomboon et al., 2013; Tung et al., 2015; Wobbrock et al., 2009; Wu, 2013; Wu & Wang, 2012); the body parts dimension was included to examine the degree of involvement of the different body parts (cf. Obaid, Häring, Kistler, Bühling, & André, 2012). Within each dimension, there were multiple subcategories (Table 2).

Table 2. Taxonomy of body gestures.

Dimension	Category	Description
Binding	*Object-centric*	Location is defined with respect to object features.
	World-dependent	Location is defined with respect to world features.
	World-independent	Location is defined without world features.
Form	*Static*	Gesture is essentially a static posture.
	Dynamic	Gesture contains body movements.
Body parts	*One hand*	Gesture is performed by one hand.
	Two hands	Gesture is performed by two hands.
	Full body	Gesture involves movements of full body.
Nature	*Physical*	Gesture acts physically on an object.
	Symbolic	Gesture depicts a symbol.
	Metaphorical	Gesture indicates a metaphor.
	Abstract	Gesture-command mapping is arbitrary.

The *binding* dimension (Dim et al., 2016; Piumsomboon et al., 2013; Tung et al., 2015; Wobbrock et al., 2009) describes the relative location where gestures are performed. Object-centric gestures act on specific objects in the interface—for instance, clicking an icon on the desktop. World-dependent gestures are performed at a specific location, such as tapping the top-right corner of the display or dragging an object off screen. World-independent gestures ignore world features, and can occur anywhere—for instance, crossing the arms to indicate rejection of a command.

The *form* dimension indicates whether a gesture only essentially contains a static posture (Dim et al., 2016; Obaid et al., 2012; Tung et al., 2015; Wobbrock et al., 2009; Wu, 2013; Wu & Wang, 2012). A gesture consists of preparation, stroke, hold, and retraction (McNeill, 1992). For static gestures, a user simply moves the body part into a specific posture and holds it for a certain amount of time before retraction; for dynamic gestures, there is a clear stroke phase that includes multiple, coherent movements of specific body parts (Obaid et al., 2012).

The *body parts* dimension counts how many body parts are involved in a gesture, by distinguishing between one-handed gestures, two-handed gestures, and full body gestures that involve at least one other body part (Obaid et al., 2012).

The *nature* dimension is divided into symbolic, physical, metaphorical, and abstract categories (Dim et al., 2016; Obaid et al., 2012; Piumsomboon et al., 2013; Tung et al., 2015; Wobbrock et al., 2009; Wu, 2013; Wu & Wang, 2012), to reflect the different levels of semantic knowledge involved in the gestures. Symbolic gestures refer to visual depictions—for example, drawing an "X" for rejection. Physical gestures act directly on the content of the interface; examples include panning, scaling, and rotating. Metaphorical gestures are performed when users treat the object or interface as something other than what it is. For example, to zoom out, the hand moves backward as if it were a telescope. The mapping of an abstract gesture to an interactive task is considered arbitrary when there is no symbolic, physical, or metaphorical connection between them.

Figure 2 shows the breakdown of the collected gestures using the aforementioned taxonomy. Half of the gestures were world-independent. The users preferred dynamic gestures (81.03%), and performing them with one hand (69.47%). Moreover, they made a few more metaphorical gestures than abstract, symbolic, or physical gestures.

We further investigated the 15 commands by exploring how the proposed gestures were distributed among the four dimensions.

Figure 3 shows the analysis in terms of binding. The gestures of *open*, *delete*, and *move* were performed on a specific object, which was confirmed by showing that most of their proposed gestures were object-centric (97.83%, 84.78%, and 97.73%, respectively). Further, it appeared that users preferred to make world-independent gestures for *reject*, *gesture on/off*, *close*, *zoom in/out*, *volume up/down*, *back*, *help*, and *pan*. However, because *accept* and *menu access* usually occur at a specific location in the interface, the majority of the proposed gestures were world-dependent.

Figure 4 shows the analysis in terms of form. The 15 commands exhibit a trend toward the use of dynamic

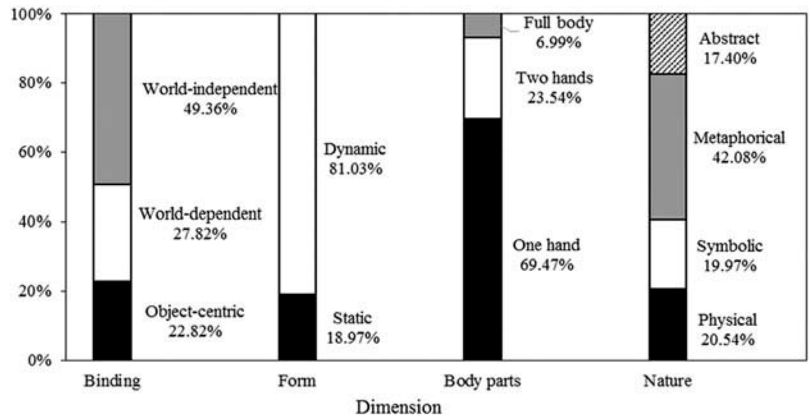

Figure 2. Percentage of gestures (N = 701) under each taxonomical category.

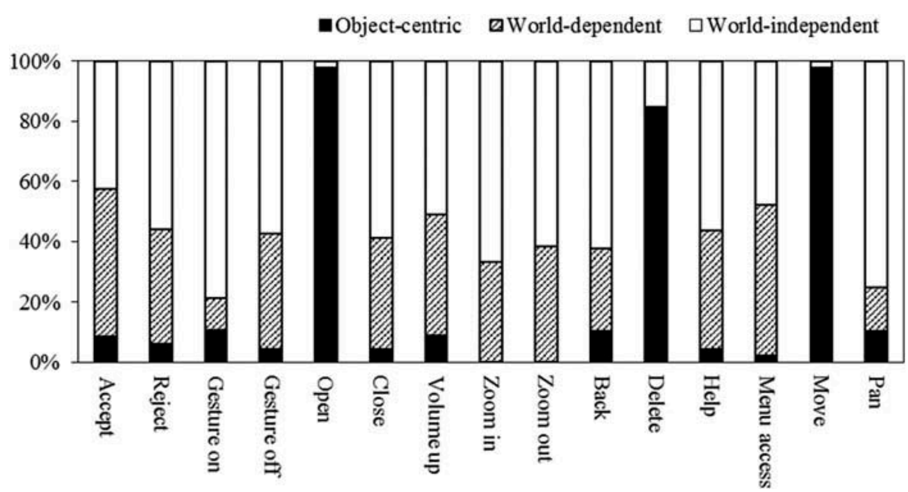

Figure 3. Distribution of proposed gestures (N = 701) in the binding dimension.

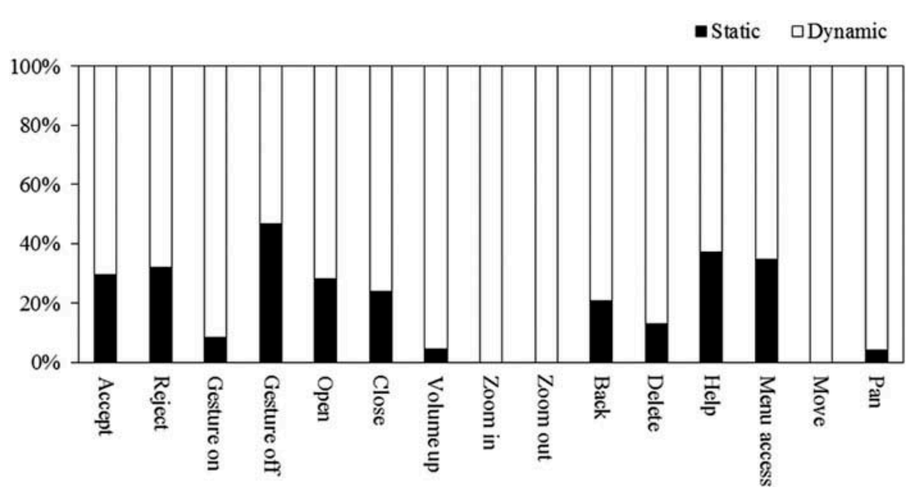

Figure 4. Distribution of proposed gestures (N = 701) in the form dimension.

gestures. It is worth noting that 46.81% of the gestures associated with the *gesture off* command are static, and most of these static gestures are symbolic (see Figure 6).

Figure 5 shows an analysis focused on the body parts dimension. For *zoom in/out* commands, users preferred using two hands. For the remaining 13 commands, users preferred using one hand.

Figure 6 shows the nature dimension distribution for each task. For *gesture on, zoom in/out, delete*, and *pan*, users preferred using metaphorical gestures.

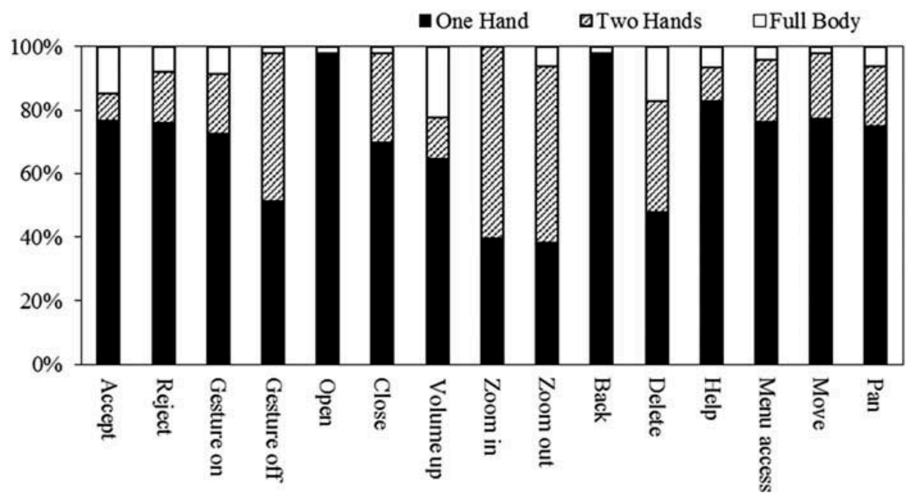

Figure 5. Distribution of proposed gestures (N = 701) in the body parts dimension.

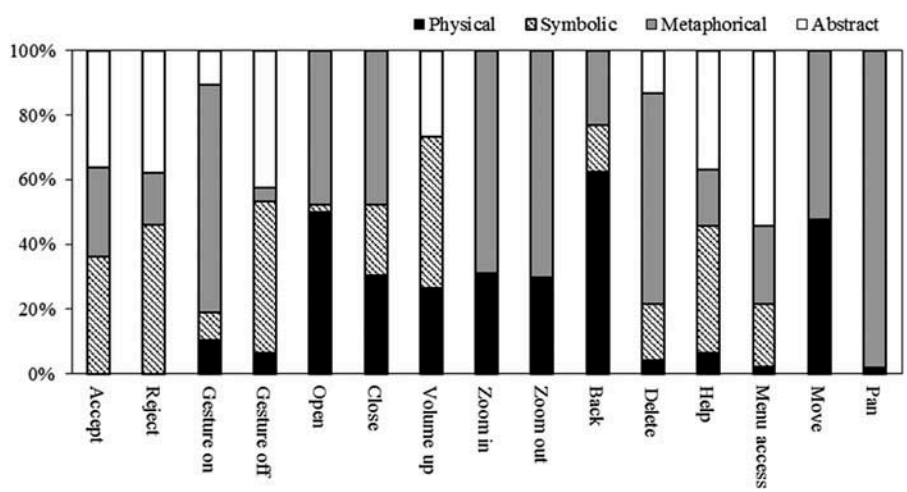

Figure 6. Distribution of proposed gestures (N = 701) in the nature dimension.

Agreement rate

Table 3 shows the AR for each command. Two commands have high agreement rates (*pan* and *open*), six commands have moderate agreement rates (*gesture on, volume up, zoom in, zoom out, back*, and *move*), and seven commands have low agreement rates (*accept, reject, gesture off, close, delete, help*, and *menu access*).

High-scoring gesture set

Table 4 lists the proposed gestures with the highest score for each command. In a few commands, the scores of the top two highest-scoring gestures were very close (a difference of less than 0.5); hence, we reported both of them.

Mental model observations

On the basis of the results from the aforementioned three parts, certain characteristics of users' mental models could be observed: (1) Participants were inclined to take advantage of their interactions with the outside world in their gesture elicitations. This could explain why the ARs for pan, open, gesture on, zoom in/out, back, move, and volume up were relatively high compared to the remaining seven commands (e.g., accept, reject, help, menu access). Most of these commands were fulfilled by performing gestures over concrete objects, and the interactive modes were analogous to those used in WIMP. (2) Participants did not prefer abstract gestures (only 17.40%); instead, they preferred gestures closely connected to real life, for example, gestures that have a clear

Table 3. Agreement Rate (AR) of each task.

Command	Number of Gestures	AR	Command	Number of Gestures	AR
Accept	46	0.073	Zoom out	47	0.202
Reject	50	0.063	Back	48	0.136
Gesture on	47	0.134	Delete	46	0.100
Gesture off	47	0.040	Help	46	0.041
Open	46	0.306	Menu access	46	0.054
Close	46	0.065	Move	44	0.196
Volume up	45	0.116	Pan	48	0.424
Zoom in	48	0.252			

Table 4. High-scoring gestures for each command in Experiment 1.

Command	AR	Gesture	Frequency (original)	Frequency	Learnability	Match	Ease	Fatigue	Preference	REBA	Overall Score
Accept	0.066	Thumb up	5	5.76	5.68	5.65	5.63	5.67	5.07	4.94	38.41
		Nod head	5	5.76	5.68	5.65	5.24	5.52	5.07	5.56	38.47
Reject	0.063	Shake head	4	5.81	5.7	5.96	5.87	5.5	4.99	5.01	38.83
Gesture on	0.134	Wave a hand	17	8.98	5.42	5.32	5.17	5.17	5.27	4.88	40.21
Gesture off	0.037	Cross arms	7	8.67	4.95	5.27	4.82	5.34	4.84	5.33	39.21
Open	0.171	Tap in via a finger	14	7.09	5.26	5.28	5.13	5.42	5.43	4.69	38.29
Close	0.056	Tap the top right corner via a finger	6	7.11	4.98	5.1	5.11	5.03	5.36	5.19	37.88
Volume up	0.127	Put a hand near an ear and slide away from it	7	6.17	5.48	5.97	5.32	4.44	5.46	4.51	37.35
		Move a hand upwards	13	7.94	5.11	4.5	4.94	4.94	4.99	5.03	37.45
Zoom in	0.252	Two hands move outwards in opposite directions	21	7.04	5.28	5.37	5.3	5.31	5.4	4.9	38.6
Zoom out	0.202	Close grip to scale via one hand	14	6.55	5.13	5.13	5.44	5.36	5.28	4.77	37.67
		Two hands move inwards	17	7.08	5.23	5.33	5.08	4.77	5.05	5.19	37.74
Back	0.137	Horizontally slide to the right via one hand	17	8.85	4.93	5.32	5.05	5.24	5.49	4.76	39.64
Delete	0.100	Grab and throw it away via a hand	7	6.29	5.38	5.27	5.67	5.43	5.47	5.33	38.07
		Vertical sliding on the file via a hand	13	8.93	4.68	4.67	4.87	5.14	5.02	5.55	38.55
Help	0.041	Hover in the top right corner via a hand	5	7.28	5.57	4.59	5.84	5.39	4.98	4.52	38.17
Menu Access	0.051	Slide down from the top via a hand	8	8.79	5.19	5.54	5.12	5.2	5.33	4.66	39.83
Move	0.196	Hover and drag via one hand	20	7.8	5.33	5.15	5.42	5.23	5.32	5.03	39.27
Pan	0.424	Horizontally slide via one hand	32	7.94	5.13	5.05	5.23	5.25	5.3	4.81	38.71

social meaning. (3) Participants preferred dynamic gestures (81.3%) over static ones, and they showed a strong trend toward using hand movements (93.01%) to fulfill the commands. (4) In line with previous studies (e.g., Wobbrock et al., 2009; Wu, 2013; Wu & Wang, 2012), participants tended to adopt reversible gestures for dichotomous tasks, for instance, "two hands move outwards in opposite directions" for *zoom in*, "two hands move inwards" for *zoom out*.

2.4. Discussion

In Experiment 1, we required participants to make at least two gestures for each command without any physical limitation, in order to find the most suitable gestures. Unsurprisingly, the proposed gestures involved not only hands, but also other body parts (e.g., head, legs, etc.); for at least three commands, participants preferred gesturing with their heads or arms. However, it is worth noting that most high-scoring gestures were related to the hand(s), suggesting their importance in gestural interaction with computers.

Although the user-defined approach is effective in revealing users' mental models for gestural interaction, it has a few limitations. One criticism of this method is that participants cannot fully understand the possibilities within the design space; because of this, critics may assert that the proposed gestures were intuitional, but may not have been "optimal." One effective method to overcome this drawback is to show users as many gestures as possible (e.g., Dim et al., 2016; Silpasuwanchai & Ren, 2015), and re-check the most suitable ones. In Experiment 2, we used this method to examine whether the high-scoring gestures in Experiment 1 were indeed the most suitable gestures.

3. Experiment 2: Subjective rating of proposed gestures

3.1. Method

Participants

A new group of twenty-five undergraduates (11 male) from Zhejiang University participated in the experiment. The participants were 22.83 (*SD* = 1.43) years old on average. None of them had experience with gestural interaction. The other aspects were the same as those in Experiment 1.

Experimental equipment and materials

To show participants as many proposed gestures as possible while keeping the length of experiment within an acceptable range, those gestures proposed by more than two participants were chosen as the target set for assessment. This resulted in a total of 75 gestures. Each command had 3 ~ 7 gestures. An experimenter, who was seated behind and to the right of the participants, was responsible for controlling the pace of the experiment and showing the subjective rating scales used in Experiment 1 to participants. Because Experiment 2 focused on the subjective evaluation of the proposed gestures, we did not calculate the REBA score, and hence, the Kinect 1.0 was not used. The other aspects were the same as those in Experiment 1.

Procedure and data analysis

The experimenter first explained the tasks to the participants; a practice command was then used to familiarize the participants with the procedure. Then, as in Experiment 1, before-and-after images and a description of the specific command were shown on the screen, and the participants were given unlimited time to understand the intention behind the command. Once the participants understood the command, a written description of a specific gesture was presented to them. They were asked to perform the corresponding gesture; moreover, if the participants performed the gesture incorrectly, they would be informed and corrected by the experimenter. Finally, they were required to evaluate the performed gestures in terms of learnability, match, ease, and subjective fatigue. Once all gestures related to the corresponding command had been completed, the participants ranked them. The next command was then shown. The commands and their corresponding gestures were presented in random order. The entire experiment lasted approximately 30 minutes.

The method for analyzing subjective data was identical to that in Experiment 1.

3.2. Results and discussion

Table 5 shows the gestures that yielded the top two highest scores. Eight commands had the same highest-scoring gesture between Experiments 1 and 2; therefore, the corresponding command-gesture combinations were recommended. On the other hand, six commands (*gesture on, gesture off, zoom in, delete, move*, and *help*) had distinct recommendations from Experiment 1, and one command (*zoom out*) overlapped with one of the two recommended gestures in Experiment 1. These findings imply that the user-defined method only partially reflects the preferences/requirements of users (cf. Choi et al., 2012; Dim et al., 2016; Silpasuwanchai & Ren, 2015). Below, we analyze the command-gesture combinations that differed between the two experiments and propose a gesture for each command by considering recognition accuracy, natural language principles, the mental models of the users, and the taxonomy distribution of each command (cf. Experiment 1).

For *gesture on*, the highest-scoring gesture in Experiment 2 was “click using a finger,” a typical gesture used in GUI interaction. However, this gesture was very similar to the one for the *open* command. In contrast, “wave a hand” was revealed to be the highest-scoring gesture for *gesture on* in Experiment 1 and the second highest-scoring gesture for *gesture on* in Experiment 2 and was easily distinguishable from the open command gesture. Therefore, we recommended using the “wave a hand” gesture from Experiment 1.

For *gesture off*, “draw a cross via a hand” contained two strokes, and the drawing routine was not fixed among users, which might have led to more recognition errors; “cross arms” was quick and had a fixed posture, posing little constraint on interaction. Therefore, the “cross arms” gesture from Experiment 1 was recommended.

For *zoom in* and *zoom out*, the “open and close grip” gesture can reduce physical risk and subjective fatigue, but this gesture poses stringent technical requirements on motion tracking and recognition. Considering that Experiment 1 revealed that two-hand gestures accounted for a higher proportion of the proposed gestures for these two commands (see Figure 5), and that it had better to offer symmetrical movements for them, we recommended having the two hands move outwards/inwards. However, if the technology is sufficiently sensitive to track the tiny gestures of the hand (e.g., Leap Motion; Weichert, Bachmann, Rudak, & Fisseler, 2013), the “open and close grip” gesture is highly recommended, particularly considering that the “close grip” for *zoom out* was also a high-scoring gesture in Experiment 1.

Table 5. Top two gestures revealed in Experiment 2, as well as the final recommended gestures.

Command	Experiment 2	Score	Same to Experiment 1	Final selection
Accept	Nod head	26.90	Yes	Nod head
	Draw a “√” via a hand	25.24	No	
Reject	Shake head	26.41	Yes	Shake head
	Cross arms	24.44	No	
Gesture on	Click using a finger	27.23	No	Wave a hand
	Wave a hand	25.82	Yes	
Gesture off	Draw a “×” via a hand	26.39	No	Cross arms
	Wave a hand	26.00	No	
Open	Tap in via a finger	28.07	Yes	Tap in via a finger
	Double click	27.50	No	
Close	Tap the top right corner via a finger	28.19	Yes	Tap the top right corner via a finger
	Grab the whole screen	25.89	No	
Volume up	Move a hand upwards	26.76	Yes	Move a hand upwards
	Put a hand near an ear and slide away from it	26.08	Yes	
Zoom in	Open grip to scale via a hand	27.23	No	Two hands move outwards in opposite directions
	Two hands move outwards in opposite directions	24.97	Yes	
Zoom out	Close grip to scale via a hand	27.81	Yes	Two hands move inwards
	Two hands move inwards	25.11	Yes	
Back	Horizontally slide to right via one hand	25.84	Yes	Horizontally slide to the right via one hand
	Slide up from the bottom via a hand	25.67	No	
Delete	Drag to the bottom right corner via a hand	25.99	No	Grab and throw it away via a hand
	Grab and throw it away via a hand	25.94	Yes	
Help	Draw a question mark using a hand	25.54	No	Draw a question mark using a hand
	Double click	24.60	No	
Menu Access	Slide down from the top via a hand	27.14	Yes	Slide down from the top via a hand
	Slide horizontally via a hand	25.82	No	
Move	Click and drag via one hand	27.62	No	Hover and drag via one hand
	Hover and drag via one hand	26.25	Yes	
Pan	Horizontally slide via one hand	27.70	Yes	Horizontally slide via one hand
	Grab and drag to the left via a hand	25.30	No	

For *delete*, the "drag to the bottom-right corner via a hand" gesture from Experiment 2 could be confused with the *move* command; the "vertical sliding over the file" gesture from Experiment 1 might have tapped other files. As the second highest-scoring gesture in Experiment 2, "grab and throw it away" was also the highest-scoring gesture in Experiment 1. Moreover, it had clear and distinct movements and was in line with our mental model. We hence chose this one as the most suitable gesture for *delete*.

For *move*, clicking to select a file can be mistaken as an attempt to open the file. We thus chose "hover and drag," revealed in both Experiments 1 and 2 as the more appropriate gesture.

Finally, for *help*, the highest-scoring gesture in Experiment 1 was "hover in the top-right corner via a hand," which, to a large extent, was affected by our GUI interaction experience (the *help* command was located in the top-right corner). However, if there was any active area (e.g., a button) in the top-right corner, using this gesture might have been inconvenient. "Draw a question mark using a hand" had no constraint on the interactive spot and belonged to world-independent gestures. Moreover, it was also recommended by Vatavu (2012). Therefore, we recommended using the gesture from Experiment 2.

A detailed demonstration of each command-gesture combination is shown in Figure 7.

4. General discussion

This study attempted to offer a gesture set supporting the most frequently used gestural interaction commands, by considering both hand and non-hand gestures. We achieved this by (1) combining the user-defined method with subjective assessments of user-defined gestures and (2) requiring participants to propose gestures without physical limitations. We considered that the two methods are complementary in meeting user requirements. The gestures obtained in Experiment 1 were often "easy to think of," which does not necessarily mean that they are optimal gestures. Using a new group of participants to evaluate the proposed gestures in Experiment 1, we reduced the limitations of the user-defined method and enabled users to better understand the design space. Indeed, we found that 53.3% of the recommended high-scoring gestures were shared between Experiments 1 and 2. Therefore, we suggest that future user-defined studies include a subjective assessment of all high-frequency gestures as an additional, yet important, step in gesture generation (see also Dim et al., 2016; Silpasuwanchai & Ren, 2015). On the other hand, we must point out that the user-defined method is quite effective in revealing the mental models of users, as five out of six command-gesture combination conflicts (according to highest score) between Experiments 1 and 2 were finally resolved by choosing the gestures offered by the user-defined method. Therefore, we consider that if design time is limited, adopting only the user-defined method is acceptable.

Based on the two experiments, a set of 15 gestures was suggested. The proposed gestures contained movements of the hand, arm, and head, suggesting the importance of including body movements beyond the motion of the hands for natural interaction in future gestural design. These 15 gestures exhibited three characteristics, which shed important light on the design principles of gestural interaction. In particular, (1) the participants preferred hand movements (13 of 15 gestures were related to the hands) and preferred using one hand to fulfill the commands, which is in line with the findings of Dim et al. (2016) and Silpasuwanchai and Ren (2015) . Therefore, accurate recognition of hand gestures is critical to the success of gestural interaction. (2) Participants intended to make gestures with clear social meaning—for instance, "draw a question mark using a hand" for *help*, "shake head" for *reject*, "nod head" for *accept*, "wave a hand" for *gesture on*, and "cross arms" for *gesture off*. These gestures are fairly common in our daily interpersonal communication, and users may implicitly treat the computer as a social agent and interact with it accordingly. Therefore, to foster natural interaction between humans and computers, designers should pay attention to enhancing natural communication between them and

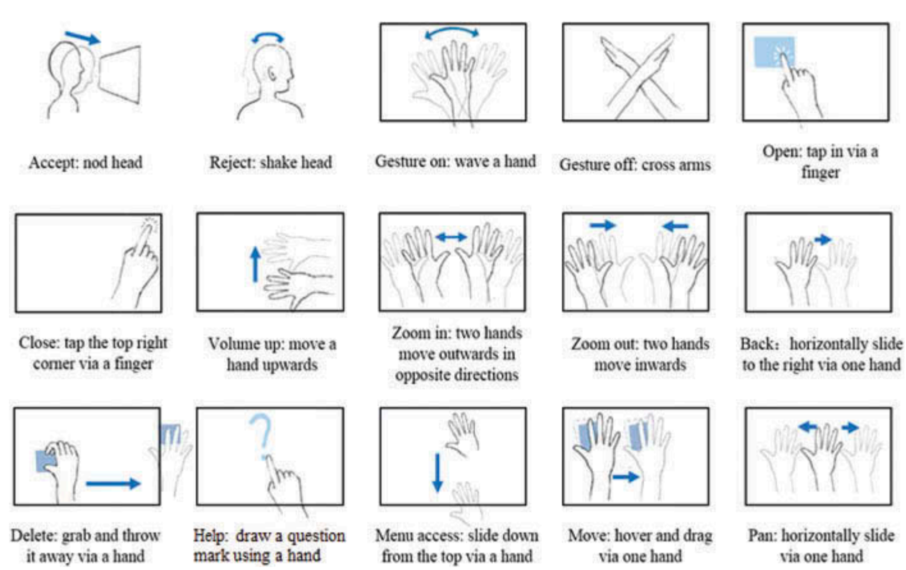

Figure 7. Recommended command-gesture combinations.

consider the adoption of social gestures from daily life. (3) Users were inclined to make dynamic gestures relative to static ones.

Of the 15 gestures, "wave a hand" for *gesture on* and "horizontally slide via one hand" for *pan* have been adopted in current gestural interaction devices, such as the Kinect-based Xbox. On the other hand, "shake head" for *reject*, "nod head" for *accept*, "grab and throw it away via a hand" for *delete* are not (to the best of our knowledge) used in current devices, nor have they been part of the gestures proposed in the few gestural interaction studies that utilized gestures of other body parts (Dim et al., 2016; Silpasuwanchai & Ren, 2015). However, it should be noted that these proposed gestures might be affected by culture. For example, for *accept* and *reject*, Western users tend to point their thumb up or down (Pereira et al., 2015; Piumsomboon et al., 2013; Vatavu, 2012), while Chinese users prefer nodding and shaking the head, respectively. Moreover, nodding and shaking the head seem to imply rejection and acceptance, respectively, for Bulgarians as well. Consequently, it is important to consider cross-cultural differences in gestural interaction. The remaining command-gesture combinations have been proposed in previous studies (Dim et al., 2016; Pereira et al., 2015; Piumsomboon et al., 2013; Vatavu, 2012; Vatavu & Zaiti, 2014; Wright et al., 2011; Wu & Wang, 2012).

The adoption of the proposed gestures also requires close cooperation with an appropriate interface design. For instance, the interface should offer feedback to help distinguish between the commands *open* and *move*, because both involve first selecting a file. One commonly used method is to float the file icon after hovering on a given file to inform the participant that it is ready to move. For "tap the top right" in the *close* command (it is noteworthy that there was no closing icon in the upper-right corner of the before-and-after image), the designer should place an exit sign in the upper-right corner of the interface to satisfy mental models formed in GUI interaction.

It should be noted that compared to other user-defined studies (Kühnel et al., 2011; Pereira et al., 2015; Piumsomboon et al., 2013; Vatavu, 2012; Vatavu & Zaiti, 2014; Wobbrock et al., 2009; Wu & Wang, 2012; Wu & Wang, 2012), the agreement rates in our study were relatively low, with only eight commands having moderate to high rates. This might have occurred for the following reasons: (1) Two-dimensional gestures usually involve direct contact with the interface, and previous 3D gesture studies focused on hand gestures. However, the 3D gestures in our study were conducted in the entire body space without constraint, leading to higher body variation. (2) Previous studies focused on touch-screen interaction, and users were more familiar with touch-screen interactions than full body-based gestural interaction. Most users had little experience in full-body interaction with computers. Indeed, during our survey to determine representative commands, it took us almost two months to find 31 users familiar with gestural interaction.

Finally, the current study had a few limitations. For instance, the selected 15 most commonly used commands may be constrained by the experience of the four expert users, because all of them were only experienced with Xbox. Moreover, there was no feedback during gestural enactment, which might have impeded the participants from being fully involved in the interaction and experiencing its real effect. However, this limitation is shared by most previous user-defined studies and must be addressed in future work. Finally, the proposed gestures might have been constrained by the interface in which they were explored. The match between gesture and interface is important for smooth human–computer interaction. Thus far, no interface has been directly designed according to the characteristics of gestural interaction. We used the Xbox interfaces to show the participants the before-and-after effects. A proper gesture-command match should also consider the properties of the user interface.

5. Key points

- To enhance the naturalness of gestural interaction, gestural designers should consider the movements of both hands and other body parts when developing body gestures.
- To attain greater compatibility between the generated gestures and gestural interaction commands, we combined the user-defined method (Experiment 1) and the subjective assessment of user-defined gestures (Experiment 2), which are complementary.
- Fifteen most commonly used gestural interaction commands across distinct application situations were selected based on literature review, expert evaluation, and a user survey.
- The proposed gestures exhibited three characteristics: participants preferred hand movements (particularly one-handed movements) and gestures with clear social meaning and were more inclined to make dynamic gestures than static ones.

Funding

This research was supported by the National Natural Science Foundation of China (31271089; 31571119), Project of Ministry of Science and Technology of the People's republic of China (2016YFE0130400).

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15* (2), 106–111.

Baudel, T., & Beaudouin-Lafon, M. (1993). Charade: Remote control of objects using free-hand gestures. *Communications of the ACM*, *36* (7), 28–35.

Bowman, D. A., & Wingrave, C. A. (2001, March). Design and evaluation of menu systems for immersive virtual environments. *Virtual Reality, 2001. Proceedings. IEEE*, 149–156, Yokohama, Japan. doi:10.1109/VR.2001.913781

Choi, E., Kwon, S., Lee, D., Lee, H., & Chung, M. K. (2012). Can user-derived gesture be considered as the best gesture for a command? Focusing on the commands for smart home system. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *56* (1), 1253–1257. doi:10.1177/1071181312561222

Choppin, S., Lane, B., & Wheat, J. (2014). The accuracy of the Microsoft Kinect in joint angle measurement. *Sports Technology*, *7* (1–2), 98–105.

Clark, R. A., Pua, Y. H., Bryant, A. L., & Hunt, M. A. (2013). Validity of the Microsoft kinect for providing lateral trunk lean feedback during gait retraining. *Gait & Posture*, *38* (4), 1064–1066.

De La Barré, R., Chojecki, P., Leiner, U., Mühlbach, L., & Ruschin, D. (2009, July). *Touchless interaction-novel chances and challenges*. International Conference on Human-Computer Interaction (pp. 161–169). Berlin, Germany: Springer.

Diego-Mas, J. A., & Alcaide-Marzal, J. (2014). Using Kinect (TM) sensor in observational methods for assessing postures at work. *Applied Ergonomics*, *45* (4), 976–985. doi:10.1016/j.apergo.2013.12.001

Dim, N. K., Silpasuwanchai, C., Sarcar, S., & Ren, X. (2016, June). *Designing mid-air TV gestures for blind people using user-and choice-based elicitation approaches*. Proceedings of the 2016 ACM Conference on Designing Interactive Systems (pp. 204–214). New York, NY: ACM.

Galna, B., Barry, G., Jackson, D., Mhiripiri, D., Olivier, P., & Rochester, L. (2014). Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease. *Gait & Posture*, *39* (4), 1062–1068.

Gao, Z., Li, J., Liang, J., Chen, H., Yin, J., & Shen, M. (2009). Storing fine detailed information in visual working memory–evidence from event-related potentials. *Journal of Vision*, *9* (7), 1–12.

Hignett, S., & McAtamney, L. (2000). Rapid entire body assessment (REBA). *Applied Ergonomics*, *31* (2), 201–205.

Igbaria, M., & Nachman, S. A. (1990). Correlates of user satisfaction with end user computing: An exploratory study. *Information & Management*, *19* (2), 73–82.

Kirakowski, J., Claridge, N., & Whitehand, R. (1998, June). *Human centered measures of success in web site design*. Proceedings of the Fourth Conference on Human Factors & the Web. Basking ridge, NJ, USA.

Kühnel, C., Westermann, T., Hemmert, F., Kratz, S., Müller, A., & Möller, S. (2011). I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, *69* (11), 693–704. doi:10.1016/j.ijhcs.2011.04.005

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, *7* (1), 57–78. doi:10.1080/10447319509526110

Lewis, J. R. (2002). Psychometric evaluation of the PSSUQ using data from five years of usability studies. *International Journal of Human-Computer Interaction*, *14* (3–4), 463–488. doi:10.1080/10447318.2002.9669130

Liang, S. F. M. (2013). *Control with hand gestures in home environment: A review*. Proceedings of the Institute of Industrial Engineers Asian Conference 2013 (pp. 837–843). Singapore: Springer.

Lund, A. M. (2001). *Measuring usability with the USE questionnaire*. Usability Interface: Usability SIG Newsletter, October. Accessed January 11, 2005. Retrieved fromhttp://www.stcsig.org/usability/newsletter/0110_measuring_with_use.html

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago, IL: University of Chicago Press.

Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., & Wobbrock, J. O. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21* (3), 40–45.

Nacenta, M. A., Kamber, Y., Qiang, Y., & Kristensson, P. O. (2013, April). *Memorability of pre-designed and user-defined gesture sets*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1099–1108). New York, NY: ACM.

Nielsen, M., Störring, M., Moeslund, T. B., & Granum, E. (2003, April). *A procedure for developing intuitive and ergonomic gesture interfaces for HCI*. International gesture workshop (pp. 409–420). Berlin, Germany: Springer.

Norman, D. A. (2010). Natural user interfaces are not natural. *Interactions*, *17* (3), 6–10.

Norman, D. A., & Nielsen, J. (2010). Gestural interfaces: A step backward in usability. *Interactions*, *17* (5), 46–49.

Obaid, M., Häring, M., Kistler, F., Bühling, R., & André, E. (2012, October). *User-defined body gestures for navigational control of a humanoid robot*. International Conference on Social Robotics (pp. 367–377). Berlin, Germany: Springer.

Pang, X., Guo, R., Yao, N., Yu, J., Wang, C., & Gao, Z. (2014). Human factor studies on gestural interaction: Past, present, and future. *Chinese Journal of Applied Psychology*, *20* (3), 243–251.

Pantic, M., Pentland, A., Nijholt, A., & Huang, T. S. (2007). *Human computing and machine understanding of human behavior: A survey*. Artifical Intelligence for Human Computing (pp. 47–71). Berlin, Germany: Springer.

Patrizi, A., Pennestri, E., & Valentini, P. P. (2016). Comparison between low-cost marker-less and high-end marker-based motion capture systems for the computer-aided assessment of working ergonomics. *Ergonomics*, *59* (1), 155–162.

Pereira, A., Wachs, J. P., Park, K., & Rempel, D. (2015). A user-developed 3-D hand gesture set for human–computer interaction. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *57* (4), 607–621.

Piumsomboon, T., Clark, A., Billinghurst, M., & Cockburn, A. (2013, September). *User-defined gestures for augmented reality*. IFIP Conference on Human-Computer Interaction (pp. 282–299). Berlin, Germany: Springer.

Schmitz, A., Ye, M., Shapiro, R., Yang, R. G., & Noehren, B. (2014). Accuracy and repeatability of joint angles measured using a single camera markerless motion capture system. *Journal of Biomechanics*, *47* (2), 587–591.

Sharp, H. (2003). *Interaction design*. New York, NY: Wiley.

Silpasuwanchai, C., & Ren, X. (2015). Designing concurrent full-body gestures for intense gameplay. *International Journal of Human-Computer Studies*, *80*, 1–13.

Tian, F., Lyu, F., Zhang, X., Ren, X., & Wang, H. (2016). An empirical study on the interaction capability of arm stretching. *International Journal of Human–Computer Interaction*. doi:10.1080/10447318.2016.1265782

Tung, Y. C., Hsu, C. Y., Wang, H. Y., Chyou, S., Lin, J. W., Wu, P. J., ... Chen, M. Y. (2015, April). *User-defined game input for smart glasses in public space*. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 3327–3336). New York, NY: ACM.

Van Diest, M., Stegenga, J., Wörtche, H. J., Postema, K., Verkerke, G. J., & Lamoth, C. J. (2014). Suitability of Kinect for measuring whole body movement patterns during exergaming. *Journal of Biomechanics*, *47* (12), 2925–2932.

Vatavu, R. D. (2012, July). *User-defined gestures for free-hand TV control*. Proceedings of the 10th European conference on Interactive tv and video (pp. 45–48). New York, NY: ACM.

Vatavu, R. D., & Wobbrock, J. O. (2015, April). *Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit*. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1325–1334). New York, NY: ACM.

Vatavu, R. D., & Zaiti, I. A. (2014, June). *Leap gestures for TV: Insights from an elicitation study*. Proceedings of the 2014 ACM international conference on Interactive experiences for TV and online video (pp. 131–138). New York, NY: ACM.

Vogel, D., & Balakrishnan, R. (2005, October). *Distant freehand pointing and clicking on very large, high resolution displays*. Proceedings of the 18th annual ACM symposium on User interface software and technology (pp. 33–42). New York, NY: ACM.

Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, *13* (5), 6380–6393.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance*. New York, NY: Psychology Press.

Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). *Maximizing the guessability of symbolic input*. CHI'05 extended abstracts on Human Factors in Computing Systems (pp. 1869–1872). New York, NY: ACM.

Wobbrock, J. O., Morris, M. R., & Wilson, A. D. (2009, April). *User-defined gestures for surface computing*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1083–1092). New York, NY: ACM.

Wright, M., Lin, C. J., O'Neill, E., Cosker, D., & Johnson, P. (2011, June). *3D gesture recognition: An evaluation of user and system performance*. International Conference on Pervasive Computing (pp. 294–313). Berlin, Germany: Springer.

Wu, H. (2013). Design of bare-hand gestures for object manipulation in virtual reality. *Journal of Information and Computational Science*, *10* (13), 4157–4166. doi:10.12733/jics20102747

Wu, H., & Wang, J. (2012, November). *User-defined body gestures for TV-based applications*. Digital Home (ICDH), 2012 Fourth International Conference on (pp. 415–420). Guangzhou, China: IEEE.

Zaiţi, I.-A., Pentiuc, Ş.-G., & Vatavu, R.-D. (2015). On free-hand TV control: Experimental results on user-elicited gestures with Leap Motion. *Personal and Ubiquitous Computing*, *19* (5–6), 821–838.

Zhao, Y., Du, N., Xu, X., Gu, Q., Wang, L., Gao, Z., & Wang, C. (2014). The influence of user-tracking feedback format on gestural interaction's user experience: A Kinect-based usability study. *Chinese Journal of Applied Psychology*, *20* (4), 367–374.

About the Authors

Zhen Chen is a graduate student at Zhejiang University, Hangzhou. She majors in applied psychology, obtained Bachelor of Science from Zhejiang University, Hangzhou, China, 2015.

Xiaochi Ma is an undergraduate student at Zhejiang University, Hangzhou, China. She majors in applied psychology.

Zeya Peng is an undergraduate student at Zhejiang University, Hangzhou, China. She majors in applied psychology.

Ying Zhou is an undergraduate student at Zhejiang University, Hangzhou, China. She majors in applied psychology.

Mengge Yao is an undergraduate student at Zhejiang University, Hangzhou, China. She majors in applied psychology.

Zheng Ma is an undergraduate student at Zhejiang University, Hangzhou, China. She majors in applied psychology.

Ci Wang is a lecturer at Zhejiang University, Hangzhou. He majors in engineering psychology, and obtained Master of Science from Zhejiang University, Hangzhou, China, 2011.

Zaifeng Gao is a professor at Zhejiang University, Hangzhou. He majors in cognitive psychology, and obtained Ph.D from Zhejiang University, Hangzhou, China, 2009. His research focuses on engineering psychology and cognitive psychology.

Mowei Shen is a professor at Zhejiang University, Hangzhou. He majors in engineering psychology, and obtained Ph.D from Hangzhou University, Hangzhou, China, 1992.