Int. J. Human-Computer Studies 80 (2015) 1–13

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

# Designing concurrent full-body gestures for intense gameplay☆,☆☆

Chaklam Silpasuwanchai, Xiangshi Ren*

*The Center for Human-Computer Interaction, Kochi University of Technology, 185 Miyanokuchi, Tosayamada-Cho, Kami-Shi, Kochi 782-8502, Japan*

ARTICLE INFO

*Article history:*
Received 8 April 2014
Received in revised form
20 February 2015
Accepted 21 February 2015
Communicated by E. Motta
Available online 3 March 2015

*Keywords:*
Full-body games
Interactivity
Kinect
Concurrent full-body gestures
User-elicitation approach
Choice-based elicitation approach

ABSTRACT

Full body gestures provide alternative input to video games that are more natural and intuitive. However, full-body game gestures designed by developers may not always be the most suitable gestures available. A key challenge in full-body game gestural interfaces lies in how to design gestures such that they accommodate the intensive, dynamic nature of video games, e.g., several gestures may need to be executed simultaneously using different body parts. This paper investigates suitable simultaneous full-body game gestures, with the aim of accommodating high interactivity during intense gameplay. Three user studies were conducted: first, to determine user preferences, a user-elicitation study was conducted where participants were asked to define gestures for common game actions/commands; second, to identify suitable and alternative body parts, participants were asked to rate the suitability of each body part (one and two hands, one and two legs, head, eyes, and torso) for common game actions/commands; third, to explore the consensus of suitable simultaneous gestures, we proposed a novel choice-based elicitation approach where participants were asked to mix and match gestures from a predefined list to produce their preferred simultaneous gestures. Our key findings include (i) user preferences of game gestures, (ii) a set of suitable and alternative body parts for common game actions/commands, (iii) a consensus set of simultaneous full-body game gestures that assist interaction in different interactive game situations, and (iv) generalized design guidelines for future full-body game interfaces. These results can assist designers and practitioners to develop more effective full-body game gestural interfaces or other highly interactive full-body gestural interfaces.



## 1. Introduction

Full-body based interaction (e.g., Kinect) has enabled more natural and intuitive input for video games. However, game gestures developed by designers may not always be the most suitable gestures for players. Indeed, players have reported difficulties in playing some full-body based games, particularly in interaction-intensive games (e.g., First Person Shooters/Action/Adventure) where several actions/commands may have to be executed at or nearly at the same time (e.g. Gamespot, 2011). Thus one key challenge in designing effective game gestural interfaces lies in defining suitable, efficient gestures that enable players to effectively perform multiple game actions/commands simultaneously and with ease.

Several studies in relation to full-body game interaction have been conducted (e.g., Hoysniemi, 2006; Norton et al., 2010), but few studies have considered the intense-dynamic nature of game environments in general. When a player's hand is occupied with “Shooting Zombies”, which other body parts and gestures might the player prefer to perform simultaneous actions such as “Reload” or “Use First Aid Kit” with. Since a literal “Jump” or “Climb” action can be tiring, is it likely that users will prefer a less tiring, more efficient gesture? What gestures would veteran gamers and non-gamers devise or envisage to enhance their interaction experiences?

To investigate these potentials, three user studies were conducted. In the first study, to explore general user preferences of game gestures, we used a user-elicitation approach asking participants to define their preferred gestures for different game actions/commands. We found a high consensus (agreement score) between participants' gestures as most participants defined physical gesture (mimicking real-world actions) with 1-hand as the most preferable input body modality. We also found a difference in preferences between gamers and non-gamers.

In the second study, to also consider simultaneous use of gestures where physical gestures may not always be possible, we asked participants to rate the suitability of different body parts (one and two hands, one and two legs, head, eyes, torso) for each game action/command. This second study was intended to help

☆An earlier half of this work was presented at the ACM CHI 2014 Conference.
☆☆This paper has been recommended for acceptance by E. Motta.

* Corresponding author. Tel.: +81 887 57 2209.
*E-mail address:* xsren@acm.org (X. Ren).

http://dx.doi.org/10.1016/j.ijhcs.2015.02.010
1071-5819/

designers consider a set of suitable and alternative body parts, since an alternative body part may be needed to execute other simultaneous gestures while a certain body part is already occupied. Through the study, we identified a set of suitable and alternative body parts and gestures for different game actions/commands.

In the third study, to develop a simultaneous gesture set, we initially asked three participants to define their preferred gesture set using the user-elicitation approach. However, we found that there was little consensus among participants. In addition, participants mentioned that it is difficult to imagine possible combinations of gestures. To assist participants, we adapted the original user-elicitation approach and introduced a novel choice-based elicitation approach. We found that this approach has a positive effect in assisting participants to discover and create suitable gestures, which resulted in a consensus set of simultaneous game gestures. Based on the three studies' findings, we highlight potentially useful design guidelines.

## 2. Related work

Our work built upon three research areas: (i) game gesture design, (ii) gesture design using a user-elicitation approach and (iii) full-body interaction. Our reviews revealed that the study in the *simultaneous use* of game gestures remains underexplored.

### 2.1. Game gesture design

There are several challenges regarding full-body game interaction (Gerling et al., 2012; Hoysniemi, 2006; Norton et al., 2010). The challenges include the fact that (i) video game actions do not always have real-world counterparts such as casting magic; (ii) the naturalness of full-body games is limited by space limitations, e.g., people cannot really walk; (iii) people's physical capabilities vary (child vs. young adults vs. older adults); and (iv) full-body games can be tiring and attract injury.

To explore full-body game interaction, the common traditional approach is the Wizard of Oz approach, where participants perform their preferable gestures with "wizard-generated" interaction. Höysniemi et al. (2004) used the Wizard of Oz approach to investigate full-body game interaction for children with basic game tasks such as running and jumping. They raised issues with the approach such as possible interaction delays and the possible effect of the avatar's movement on children's movements. Norton et al. (2010) also used the Wizard of Oz approach to explore full-body game interaction with basic game tasks including running, jumping, climbing, and crouching. By analyzing video recordings and interviews, they found that users can adapt to physical limitations of full-body interaction effectively. For example, users may perform natural locomotion first (e.g., physically walking forward and backward) but will switch to compensating locomotion (e.g., walking in place) when physical space is limited. In addition, they made an initial observation that although hand and arm are considered the most used body modalities they may not always be prominent as arms may be busy with other commands. This is coherent with our initial argument.

We build upon these studies to include more various game actions/commands, and to particularly consider the simultaneous use of game gestures, which remain underexplored.

### 2.2. Gesture design using user-elicitation approach

The Wizard of Oz approach might suffer from possible interaction delays and can be time-consuming in preparing the setup (Höysniemi et al., 2004). One recent, inexpensive approach to design suitable gestures is the user-elicitation methodology, originally based on the guessability technique (Nielsen et al., 2004; Wobbrock et al., 2005), where participants were asked to define gestures for commands/actions. The methodology was proposed under the assumption that game gestures created by designers do not necessarily match users' mental models and may suffer from a large gulf of execution. Comparing with the Wizard of Oz approach, instead of providing "wizard-generated" interaction and feedback, the user-elicitation approach first provides visual cues (i.e., video clips) demonstrating the effect of each command/action, then asks the participants to perform their preferred gestures to trigger the effect, without any further feedback.

Regarding effectiveness, Morris et al. (2010) and Wobbrock (2009) suggested that a user-elicitation approach can produce more preferable gestures, compared to gestures designed by HCI experts who are likely to develop more "physically and conceptually complex gestures than end-users". Regarding learnability, Nacenta et al. (2013) compared memorability between three gesture sets: an author-defined set, a user-defined set, and a randomly-defined set and found that the user-defined gestures are easier to learn and remember.

Motivated by the usefulness of the approach, many works have been conducted for different user groups. For example, Kane et al. (2011) employed a user-elicitation approach to study how blind people use touch screen gestures. They found differences in gesture preferences between blind and sighted people, e.g., blind people prefer edge-based gestures. Mauney et al. (2010) also investigated touch-screen gestures but across nine different cultures and found that there is generally a high level of agreement in gesture preferences. Connell et al. (2013) explored child-defined gestures for full-body interaction and found some specific characteristics of children, e.g., children rely more on egocentric (body-centered) gestures but gradually change to allocentric (world-centered) gestures over time. These results suggested that different user groups may have differences in gesture preferences (e.g., non-gamers vs. veteran gamers) and it is essential to understand those differences to design suitable gestures.

The user-elicitation approach also has been used for designing gestures in different interactive devices and contexts, e.g., tabletops (Wobbrock, 2009), mobile phones (Ruiz et al., 2011), humanoid robots (Obaid et al., 2012), TV (Vatavu, 2012), 3D remote object manipulation (Liang et al., 2012b), and phone-to-phone/tabletop/large-display interaction (Kray et al., 2010). Most works resulted with design practices and a consensus set of suitable gestures which were determined by the highest agreed gestures among participants.

In summary, our review shows that the user-elicitation approach can help understand user preferences and develop suitable gesture sets and practices. In addition we also found that little study has been done on user elicitation for full-body based video games, and that considering the dynamic, interaction-intensive nature of video games, thus we were motivated to fill this gap.

### 2.3. Full-body interaction

Full-body interaction includes the use of body movements and gestures for interacting with computers, which may be categorized into four kinds: (i) full-body only (e.g., Kinect), (ii) full-body plus external devices (e.g., Kinect + artificial gun), (iii) external device only (e.g., Wii Remote) and (iv) body-centric interaction (i.e., using the body as interaction space). In full-body only interaction, users interact with computers (e.g., TV, games) using full-body movements and gestures through motion-camera sensing devices. To enhance realism, full-body only interaction may be augmented with external devices such as artificial gun (Williamson et al., 2011) for interaction. In external device only interaction, body movement is recognized through controllers (e.g., Wii Remote) or mobile sensors (Cohn et al., 2012), rather than from a motion-camera sensing device. In body-centric interaction, designers explored extending the interaction space to the body. For example, Chen (2012) investigated "body-centric interaction" for mobile phones which extends the mobile device interaction space from only the 2D screen to body space, e.g., using the arm as menus. Shoemaker et al. (2010) also conducted similar work but on a large display, e.g,. touching the hip to open a toolbar on a map application. Harrison et al. (2012) investigated how the body can become the input space, e.g., using the palm as a phone input screen. In our work, we primarily focused on full-body only interaction as it is a typical interaction seen in many full-body games (i.e., Microsoft Kinect).

A large body of work has been conducted regarding full-body interaction. Bianchi-Berthouze (2013) investigated how body movement affects game engagement, and found that body movement enhances affective and social experiences in gameplay. In relation to movement, Isbister et al. (2011) suggested that a greater amount of movement appears to lead to a greater amount of enjoyment. Nijhar et al. (2012) found that increasing movement recognition precision leads to player's higher level of immersion. Pasch et al. (2009) identified movement-related factors that influence player engagement: natural control, mimicry of movements, proprioceptive feedback, and physical challenge. Tholander and Johansson (2010) studied what makes movement enjoyable by observing non-digital artefacts (e.g., real-world sports) and proposed eight design guidelines, e.g., allowing users to significantly improve their skills by allowing only small changes in movement could increase their sense of pride and mastery. Gerling et al. (2012) explored full-body interaction for the elderly. Some suggested guidelines include fatigue management, accessibility support, simple setup and easy gesture recall. Isbister and DiMauro (2011) analyzed movement-based design patterns such as kinesthetic mimicry and piecemeal movement. Overall, few studies have considered the simultaneous use of game gestures.

Regarding the application area of full-body interaction, it has greatly expanded in recent years, and includes using full-body interaction for interaction with large public displays (Freeman et al., 2013; Müller et al., 2012), TV controls (Vatavu, 2012), storytelling (Samanci et al., 2007) and collaborative virtual environments (Schaeffer et al., 2003). A recent research effect focuses on using full-body interaction for promoting physical well-being (i.e., exergames). Some research findings suggest that exergames motivate users in physical activity (Berkovsky et al., 2010; Peer et al., 2011; Yim and Graham, 2007) and enhance physical health (Hoysniemi, 2006; Macvean and Robertson, 2013). Exergames have also been used to address serious health problems such as cerebral palsy (Hernandez et al., 2013), stroke (Alankus and Kelleher, 2012), and chronic back pain (Schönauer et al., 2011).

In summary, we believe that designing suitable game gestures is crucial to reach high level of player enjoyment and immersion and will be beneficial for related application areas including the area of health and exergames and for the digital entertainment industry per se.

## 3. Experiment I: analyze user preferences

To first explore player's preferences in game gestures, a user-elicitation study was conducted, i.e., participants were asked to define single gestures for each common game action/command. Agreement scores, frequency of use of body parts, gesture types, and subjective assessment were analyzed.

### 3.1. Participants

Twelve university students (all males, $M=22.1$ years) were recruited. Regarding game experience, five of the participants regularly played games in both PC (mouse+keyboard) and consoles (game controller) with more than 15 game hours per week (veteran gamers). Regarding experience with Kinect or other motion game gestures, only seven of the participants (including all five veteran gamers) reported having some prior experience (one to three times). All participants were right-handed and each was paid $10.

### 3.2. Events[1]

We derived events from various genres of Kinect games including *Gunstringer* (First Person Shooters), *Blackwater* (First Person Shooters), *Forza 4* (Racing), *Kinect Joy Ride* (Racing), *Kinect Adventure* (Adventure), *Kinect Rush A Disney Pixar Adventure* (Adventure and Role-playing), *Rise of Nightmares* (Adventure and Role-playing), *Kinect Sports* (Fighting and Sport), *Virtual Tennis 4* (Sport), *London 2012 Olympics* (Sport). This list of games is in no way exhaustive but these games do cover various genres of motion gaming, thus implying that these games can serve as a good representative starting point. We also included a few actions not included in current Kinect games but we considered them to be typical in video games such as "Stealth Walking" and "Steal". We left out a few repetitive actions such as "Use First Aid Kit" and "Use Power Up" and generalized them into one common action, e.g., "Use Item". There are some actions such as "Catch Ball" or "Jump" where the resulting gestures could be obvious, but including them in our study enables us to observe any specific behaviors (e.g., 1-leg vs. 2-leg).

A total of thirty-two actions and commands were derived. The complete set of events includes first person shooter's actions (Shoot, Reload Gun, Viewpoint Change), racing actions (Drive, Accelerate, Shift gear), adventure/role-playing actions (Climb, Walk, Stop Walking, Run, Jump, Stealth Walking, Steal, Slash, Open Chest, Open Door, Pick Item, Use Item, Push Box, Shake Item), sport/fighting actions (Hitting Baseball, Catch Ball, Throw Ball, Row Boat, Roll Bowling, Racket Hitting, Kick, Guard), and system commands (Zoom-In, Zoom-Out, Open Menu, Select Menu Item).

### 3.3. Procedure

Our study design used a user-elicitation approach similar to that of Wobbrock (2009) and Ruiz et al. (2011). At the start of the experiment, participants were asked to define game gestures for 32 game events in randomized order. To identify the most preferred gesture and reduce any ordering effects, participants were allowed to choose the same gesture for multiple events. Each event was displayed along with the name of the event and a target scene on a large display. Target scenes were created by using Visual Studio 3D Toolkit to represent an effect (e.g., a treasure chest being opened, an opponent being slashed), and participants were asked to perform gestures to cause (trigger) the effect. Target scenes were not taken from animated screenshots of any existing video games, as we tried to keep the target scenes independent of any particular game, which might otherwise influence the resulting gesture. Some target scenes were not animated, e.g., "Stealth walking" or "Drive", as these scenes require the participants to see the actors in a third-person perspective to show the effect clearly, which may influence the participant's defined gestures; so we

[1] For simplicity, the term *"events"* will be used in place of *"actions and commands"* throughout this paper.

simply communicated the effect using instructions along with static images containing the interaction medium/context of the events (see Fig. 1).

During gesture definition, participants were instructed to think-aloud while performing their gestures, confirm the start and end of their performed gesture and describe the corresponding rationale. Participants stood approximately 1.8 meters away from the display while performing gestures. The experiment was audio and video recorded for later analysis. As with common elicitation studies, we did not want participants to take recognizer issues into consideration, i.e., to remove the *gulf of execution* (Hutchins et al., 1985) between the user and the device, and to observe the users' best possible gesture without users being affected by the recognition ability of the system – similar to the rationale described in Wobbrock (2009) and Ruiz et al. (2011).

A similar evaluation method to that of Wobbrock (2009) was used: “The gesture I performed is a good match for its purpose”; “The gesture I performed is easy to perform”; “The gesture I performed is tiring”. The questionnaire followed a 7-point Likert scale style with 7 as “strongly agree”. Participants evaluated the gestures immediately after each performed gesture to assist recall efforts. To improve consistency of the evaluation, after all gestures were performed and evaluated, participants were allowed to double-check their evaluation scores for each gesture, and when needed, to look at the video or readjust their scores if needed. In our study, only two participants revised their scores but without any dramatic changes (i.e., both participants changed their scores for only two gestures by the scale of one, e.g., 6 to 7) and did not affect our results. The experimental session took around 1-hour.

### 3.4. Results

Agreement scores, frequency of use of body parts, gesture types, and subjective assessment were analyzed. A total of 384 gestures were collected.

**Fig. 1.** In Experiment I, participants were asked to perform gesture for different common game events.

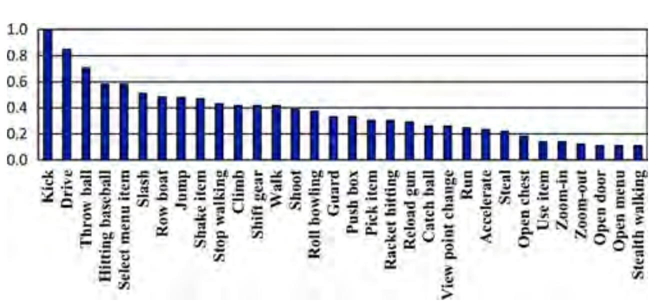



**Fig. 2.** Experiment I – overall agreement scores for events sorted in descending order.

#### 3.4.1. Agreement score

Wobbrock (2009)'s method was used to investigate the degree of consensus for each game event. The calculation of agreement score is as follows:

$$A_r = \sum_{Pi} \left( \left| \frac{P_i}{P_r} \right| \right)^2$$

*Pr* represents all gestures performed for event *r* and *Pi* is a subset of identical gestures from *Pr*. *Ar* ranges 0 to 1. Gestures were considered identical if they have similar trajectories and poses. For example, for the agreement score of the event “Walk”, a total of 12 participants and their corresponding 12 gestures were considered. This event was divided into 4 groups of identical gestures. The size of each group was 7, 3, 1, and 1. Therefore, the agreement score for “Walk” is

$$A_{walk} = \left(\frac{7}{12}\right)^2 + \left(\frac{3}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 = 0.42$$

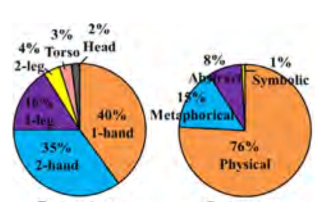
Fig. 2 shows the agreement score for each game event. The overall agreement score was 0.37, which is slightly higher than Wobbrock (2009) (0.32 and 0.28). Regardless of the high overall agreement score, system commands including “Open Menu”, “Zoom-In”, “Zoom-Out” achieved relatively low overall agreement scores (0.126).

#### 3.4.2. Use of body parts and gesture types

Fig. 3 shows the use of body parts. It is not surprising that one-handed gestures were the most preferred (40%), followed by two handed (35%), one leg (16%), two legs (4%), torso movement (3%), and head (2%). However, this pattern was not the same across all actions. For navigational events such as “Run”, “Walk’, “Stop Walking”, most participants preferred leg gestures. For the “View-point Change” event, participants preferred head input with a few using the torso. For the commands “Select Menu Item” and “Open Menu”, many participants preferred hands, however head input was used by some.

Regarding gesture types, it is important to highlight the difference in gesture preferences between veteran gamers and normal/non-gamers. We classified the defined gestures into four broad types (Wobbrock, 2009): physical (direct manipulation), symbolic (making an “OK” gesture), metaphorical (using a foot to “double click” items), and abstract (arbitrary mapping). The classification was conducted by the authors. To improve reliability, an independent rater performed the same classification with high interrater reliability (Kappa $=0.932$, $p<0.001$).

As shown in Fig. 3, most defined gestures resembled real-world gestures (physical gestures). On the other hand, non-physical gestures were mostly performed by veteran gamers, e.g., for the action “Jump”, non-gamers tend to prefer using an actual “Jump” action while veteran gamers preferred a less tiring gesture (e.g., raising a leg slightly above the ground). In any case, the found differences between veteran and novice gamers should be treated suggestively due to the small sample size of veteran gamers ($n=5$).




| | Physical | Metaphor | Abstract | Symbolic |
|---|---|---|---|---|
| Normal / non-gamers | 66% | 32% | 40% | 0% |
| Veteran gamers | 34% | 67% | 60% | 100% |

**Fig. 3.** The proportion of body parts used and gesture types (left) and veteran gamers vs. non/normal gamers (right).

*3.4.3. Subjective assessment*

Participants evaluated their defined gestures based on the Likert 7-scale rating (7 as strongly agree). The mean value is "The gesture I performed is a good match for its purpose?" (5.84); "The gesture I performed is easy to perform" (5.90); and "The gesture I performed is tiring" (5.51). The correlation between participants' good match ratings and agreement scores was found to be significant ($r = 0.746$, $p < 0.01$), suggesting that a good match rating is a good indicator of the consensus between users of the gestures performed.

### 3.5. Discussion

*Individual difference*: Our study achieved a relatively high agreement score compared to past works. The high agreement score was due to participants often mimicking their real-world gestures (e.g., kick) for performing game events. However, there were also some actions that exhibit slight variations which reduce its agreement score such as "Racket Hitting" – four different gestures were produced – two-arm swing, one-arm swing, wrist-flicking, or palm-gesture for "Racket Hitting". The variations in gestures reflect individual differences – some participants prioritize efforts required while some attempts to maximize the correspondence between game events and real-world gestures. One individual characteristic we found to affect gesture preference is gaming expertise. More specifically, gamers tend to define a more symbolic, efficient unique gesture library, while less-experienced gamers tend to define a more straightforward, physical gesture set. Possible explanations include the observation that veteran gamers tend to play games for a longer period of time, thus they prefer more efficient, less tiresome gestures to prolong their duration of play. Another possible explanation is that veteran gamers may tend to be more motivated to overcome challenges, thus they prefer more efficient gestures to competitively engage in those challenges (Nijhar et al., 2012; Pasch et al., 2009). Last, it may simply be due to the overall greater game experiences of veteran gamers, thus leading them to define more efficient game gestures.

To accommodate this difference, the idea of accommodating multiple gestures per game event has support (Norton et al., 2010). For example, it might be useful for designers to incorporate two possible gestures for one event – a physical gesture + a more efficient, symbolic gesture. Such accommodation (e.g., shortcut commands) has already been implemented in a desktop-based system facilitating novice and expert users' needs and thus implying that it could be useful for full-body interaction. Aside from gaming expertise, designers might also need to consider varying levels of player motivation (Mizobata et al., 2014) and also consider that the same gesture may not be valid for all users, places (Freeman et al., 2013; Müller et al., 2012) or situations (e.g., gamepad might be preferred over a long session of gameplay).

*2D vs. 3D interaction*: Regarding system commands such as "Zoom-in/out" and "Open menu", these commands achieved relatively low agreement scores. Although gestures for these commands were well-defined in 2D interfaces, they do not appear to be transferable to 3D interfaces. This indicates that designers might need to consider developing new 3D gestures for similar commands. Perhaps due to the larger degree of freedom when compared to 2D interaction paradigms (e.g., mobile interactions), the elicitation resulted with diverse gesture definitions. It may be beneficial for designers to investigate the right degree of freedom for users on similar system commands.

*Limitations*: Experiment I has several limitations – first, alternative body parts for each game event were not adequately explored in cases where the preferred body part was occupied with the task at hand. For example, it may not be possible to use a hand for both "Drive" and "Shift Gear". Alternative body parts should be identified to assist designers during the assignment of gestures; second, few possible simultaneous game gestures were discovered, i.e., participants mostly defined gestures based on their real-world actions but such actions may not always be applicable in highly interactive game situations and gestures more applicable to digital gaming might thus be overlooked. These issues were addressed in Experiment II (alternative body parts) and III (simultaneous gestures) respectively.

## 4. Experiment II: Identifying suitable and alternative body parts

In Experiment II, we seek to identify suitable and alternative body parts. We asked participants to rate the suitability of each of the body parts for each game event (Fig. 4).

### 4.1. Participants

Another twelve university students (1 female, $M = 22.4$ years) were recruited. Regarding game experience, six were veteran gamers who regularly played games on both PC and consoles with more than 15 hours per week. Regarding experience with Kinect or other motion game gestures, only eight of the participants (including all veteran gamers) had prior experience with full-body games (two to three times). All were right-handed. They were paid $10. The only female participant had no experience in video games. In our study, we did not observe any gender-specific differences between the female participant and other male non-gamer participants.

### 4.2. Events

For better coverage, 8 more game events were added for analysis including adventure/role-playing actions (Crawl, Roll Body), sport/fighting actions (Dodge, Headbutt, Punch), racing actions (Break, Boost), first person shooters actions (Hide).

### 4.3. Procedure

Participants were asked to define possible game gestures for 40 game events using different body parts in randomized order. These body parts were derived from our pre-study where we asked eight participants (6 males, $M = 21.3$ years) to nominate the possible

**Fig. 4.** In Experiment II, we asked participants to perform gestures using different body parts. In the figure above, we asked participants to "Shoot" using head (the participant moved head forward to "Shoot") and other body parts as well.

body parts for performing mid-air gestures. We grouped similar body parts (e.g., wrist/arm/finger → hand) and identified seven main body parts (one and two hands, one and two legs, head, eyes, torso) to be used in this study. The group classification was achieved with high interrater reliability (Kappa=0.985, $p<0.001$).

As opposed to performing *only one* gesture using any body part for each game event (Experiment I), Experiment II asked participants to perform *one gesture per one body part* (thus seven gestures per game event) and rate the suitability of each body part – "The body part I used is suitable". The questionnaire followed a 7-point Likert scale style with 7 as "strongly suitable".

Participants were instructed to skip when they felt that no suitable gestures could be performed using a particular body part. For each body part, participants were allowed to use different subparts (e.g., wrist/thumbs of the hand, elbow of the arm, feet of the leg). Similarly, we instructed participants that eye gestures could be any interaction performed using eye movements or fixations. Participants were also instructed that they were allowed to touch another body part (e.g., touching the left shoulder with right hand) where the gesture would be counted as a hand movement if the hand is the initiator. The setting was similar to that used in Experiment 1. The experimental session took around 1.5-hours.

### 4.4. Results

The agreement score, suitable and alternative body parts, and gestures were analyzed and reported.

#### 4.4.1. Agreement score

We used a similar approach to Experiment I to calculate the agreement score. Fig. 5 shows the agreement score for each game event. Gestures were considered identical if they have similar trajectories and poses. The mean agreement score was 0.19, which was much lower than Experiment I (0.37). This was expected due to the design of Experiment II which asked participants to perform each gesture using different body parts.

Many events (e.g., Kick, Drive, Shoot) that achieved high agreement scores in Experiment I have lower agreement score in Experiment II. For example, there were a total of 11 different gestures performed by participants to "Shoot", e.g., using one hand/two hand as a gun, using head movement, using the whole arm as a gun, using a kick-gesture.

On the other hand, events such as "Headbutt", "Racket Hitting", and "Catch Ball" achieved high agreement scores, given that most participants agreed that the most natural way to perform such actions is to imitate real-world actions.

Comparing Experiments I and II, the largest group of defined gestures was consistent, i.e., 31 out of 32 events in both Experiment I and II have the same largest group of defined gestures. For example, in both Experiment I and II, the most preferred gesture for "Shoot" is a one hand gun-gesture.

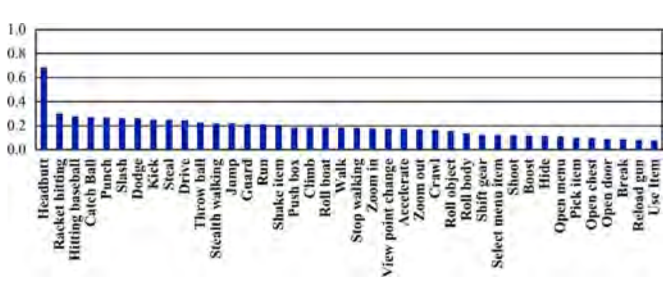



**Fig. 5.** Experiment II – overall agreement scores for events sorted in descending order.

| No. | Events | Suitability of body part: 1-hand | 2-hand | 1-leg | 2-leg | Head | Eye | Torso |
|---|---|---|---|---|---|---|---|---|
| 1 | Accelerate | 2.75 | 2.75 | 4.75 | 1 | 1 | 1 | 1 |
| 2 | Boost | 6.75 | 3.5 | 1.75 | 1 | 1 | 1 | 1 |
| 3 | Break | 2.75 | 1.75 | 6.25 | 1 | 1 | 1 | 1 |
| 4 | Catch ball | 4 | 6.5 | 1 | 1 | 1 | 1 | 1.25 |
| 5 | Climb | 1 | 6.5 | 1 | 2.5 | 1 | 1 | 1 |
| 6 | Crawl | 3.5 | 2.5 | 1 | 2 | 1 | 1 | 1 |
| 7 | Dodge | 1 | 1 | 1 | 1 | 1 | 1 | 6.5 |
| 8 | Drive | 6.25 | 6.75 | 1 | 1 | 1 | 1 | 2.25 |
| 9 | Guard | 5.5 | 6.75 | 4 | 1 | 1 | 1 | 1 |
| 10 | Headbutt | 1 | 1 | 1 | 1 | 7 | 1 | 1 |
| 11 | Hide | 2 | 2.5 | 1.25 | 2.5 | 1.75 | 1 | 6.25 |
| 12 | Hitting baseball | 4.5 | 4 | 1.75 | 1 | 1 | 1 | 1 |
| 13 | Jump | 1.75 | 1 | 3 | 5.5 | 1 | 1 | 1 |
| 14 | Kick | 1.75 | 2.25 | 7 | 2.75 | 1 | 1 | 1 |
| 15 | Open chest | 5.25 | 4.25 | 3.25 | 2.5 | 1 | 2 | 1 |
| 16 | Open door | 6.25 | 3.25 | 4.75 | 1 | 1 | 2 | 1 |
| 17 | Open menu | 3.75 | 6.25 | 2 | 1.75 | 1.75 | 1 | 1 |
| 18 | Pick item | 6.5 | 5.25 | 1.75 | 1 | 1 | 2.25 | 1 |
| 19 | Punch | 7 | 6.5 | 1.75 | 2.25 | 1 | 1 | 1 |
| 20 | Push box | 6.25 | 6.25 | 3.25 | 1 | 2 | 1 | 1 |
| 21 | Racket hitting | 5.75 | 4 | 1.75 | 1 | 1 | 1 | 1 |
| 22 | Reload gun | 4.75 | 5.75 | 2.25 | 1 | 1 | 1 | 1 |
| 23 | Roll body | 2.5 | 2 | 1 | 2.5 | 1 | 1 | 1 |
| 24 | Roll object | 6 | 6.25 | 3.25 | 1 | 2 | 1 | 1 |
| 25 | Row boat | 5.25 | 6.25 | 1.75 | 1 | 1 | 1 | 1 |
| 26 | Run | 1 | 3.5 | 2 | 5.5 | 1 | 1 | 1 |
| 27 | Select menu item | 5.5 | 4 | 3 | 2 | 1 | 2.25 | 1 |
| 28 | Shake item | 6.5 | 6 | 1 | 1 | 1 | 1 | 1 |
| 29 | Shift gear | 6.5 | 2.25 | 3.25 | 1 | 1 | 1 | 1 |
| 30 | Shoot | 5.75 | 5 | 1.5 | 1 | 1.25 | 1 | 1 |
| 31 | Slash | 6.25 | 5 | 2.5 | 1 | 1.75 | 1 | 1 |
| 32 | Steal | 6.5 | 3.25 | 1 | 1 | 1 | 1 | 1 |
| 33 | Stealth walking | 1 | 1.25 | 1 | 5.5 | 1 | 1 | 1 |
| 34 | Stop walking | 6 | 2.75 | 3.75 | 1.5 | 1 | 1 | 1 |
| 35 | Throw ball | 6.25 | 2.75 | 1 | 1 | 1 | 1 | 1 |
| 36 | Use item | 5.75 | 5.5 | 2 | 1.75 | 2 | 1 | 1 |
| 37 | View point change | 3.5 | 1 | 2.25 | 1 | 6.5 | 5 | 1.75 |
| 38 | Walk | 1.25 | 3.5 | 2 | 5.5 | 1 | 1 | 1 |
| 39 | Zoom-in | 4.5 | 5.75 | 2.5 | 1 | 2.25 | 1 | 1 |
| 40 | Zoom-out | 4.25 | 5.25 | 2.5 | 1 | 2 | 1 | 1 |

Most suitable body part
Most suitable alternative body part

**Fig. 6.** A set of suitable and alternative body parts (7 as most suitable).

| | | | Composite Score |
|---|---|---|---|
| 1-hand Drive (6.25) | 1-hand Shift Gear (6.5) | 1-leg Boost (1.75) | 14.5 |
| | 1-leg Shift Gear (3.25) | 1-hand Boost (6.75) | 16.25 |
| | Other body parts (1) | Other body parts (1) | 8.25 |
| 2-hand Drive (6.75) | 1-leg Shift Gear (3.25) | 1-leg Boost (1.75) | 11.75 |
| | Other body parts (1) | Other body parts (1) | 8.75 |
| Torso Drive (2.25) | 1-hand Shift Gear (6.5) | 1-hand Boost (6.75) | 15.5 |
| | 1-hand Shift Gear (6.5) | 1-leg Boost (1.75) | 10.5 |
| | 1-leg Shift Gear (3.25) | 1-hand Boost (6.75) | 12.25 |
| | 1-leg Shift Gear (3.25) | 1-leg Boost (1.75) | 7.25 |
| | Other body parts (1) | Other body parts (1) | 4.25 |

**Fig. 7.** Example of using the suitable and alternative body part from Fig. 6 to determine possible combinations of "Drive"+"Shift Gear"+"Boost". Designers may choose among "close" score solutions (e.g., highlighted blue). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)

#### 4.4.2. Suitable and alternative body parts and gestures

We analyzed the most suitable body parts and alternatives based on the participants' subjective assessments of "The body part I used is suitable". The most suitable body part for each event was selected based on the highest mean score. We also analyzed alternative body parts in case the most suitable body part was already occupied with the task at hand. Alternative body parts were selected on the grounds that they did not conflict with the most suitable body part and they presented with the highest mean score among all the alternatives. In cases where two body parts achieved equivalent highest mean scores and posed no conflict in body part usage, both were highlighted as the most suitable body part and no alternative was highlighted (e.g., Roll body).

Fig. 6 shows the suitable and alternative body parts for each game event. To apply Fig. 6, one can consider the score to determine the suitable combinations. For example, if designers were to design combined gestures of "Drive" + "Shift Gear" + "Boost", designers may need to consider multiple "close" solutions and their preference score (see calculation example in Fig. 7). Note that left and right hand/leg may be used simultaneously.[2]

For actions such as "Shoot", "Drive", "Open Door", "Open Chest", hands were most preferred. Legs were most preferred for navigating the game avatar. For leg gestures, the ability to maintain good body balance was the participants' top priority. For example, participants commented that while performing certain 1-leg gestures (e.g., raising one leg up) they can easily lose body balance and thus may not be effective during fast gameplay. Meanwhile, for 2-leg gestures, fatigue was the participants' primary consideration, thus designers may try to avoid tiresome 2-leg gestures. In cases when legs are occupied, "Walk", "Run", and "Stealth Walking" actions could be optionally performed using two-hands on the sides of the body swinging back and forth, with the swinging velocity determining the different actions. For viewpoint changes, there was a strong consensus for using head movement as the primary input, but in cases when the head input was already assigned,"View point change" could also be optionally performed using eye movement or 1-hand movement as a cursor indicating the point of view. Eye movement was also promising for target acquisition tasks (e.g., "Pick Item") but some participants mentioned that it could be difficult to control their eye movements. Torso gestures were most (and the only) preferred body part for torso actions (e.g., "Dodge", "Hide").

For system commands, symbolic gestures were preferred, e.g., "Zoom-in/out" action which could be performed using either a hand or a leg; a clenched fist can indicate "Zoom-in" and an open palm can indicate "Zoom-out" or, alternatively, raise a foot to 45°to indicate "Zoom-out" and lower the foot for "Zoom-in". "Open menu" could be performed by swiping an index-finger down in the air, or making a camera gesture (imitating a photo-taking gesture by having the thumb and index finger of each hand resembling an L-shape sign, combining both hands to form a rectangle) or by dragging one foot from 12 o'clock to 6 o'clock or by performing a rotating movement with the head. "Select menu item" could be performed using a hand or the eyes to move the cursor, or by moving the legs to a different position (up, down, left, right) of the clock (three o'clock, six o'clock, etc). The strategies of participants for system commands was that these commands should accommodate fast access during the middle of a intense gameplay.

### 4.5. Discussion

*Transferability between hand and leg*: Our resulting set of suitable and alternative body parts posed a potentially useful, interesting design implication – the possible transferability between hand and leg gestures. When the hand(s) is occupied, the leg(s) tends to be the participants' preferred alternative, and vice versa. For example, gun-gesture using hand may be the most preferable gesture for "Shoot", but when the hand is occupied with other tasks, participants suggested using kick-gesture for "Shoot". In a similar way, "Jump" can be performed alternatively by facing a palm up and raising the hand when the leg is occupied. Players could also simultaneously perform the "Use item" action by "double clicking" (tapping) one foot on the floor, instead of using hand(s). Together with "Drive" using driving-wheel gesture, "Break" and "Accelerate" can be simultaneously performed by stamping the foot to 12 o'clock or 6'clock. These transferability implications could prove handy for designers when designing simultaneous gestures. Furthermore, some participants suggested the possible gesture transferability between left and right hands in cases where one hand is already assigned. For example, one participant commented that although one hand is used for "Shoot", another hand may be used for "Use Item" or "Reload" gesture. Similar comments were reported for left and right leg.

*Limitation*: In Experiment II, we were able to identify various game gestures using different body parts (e.g., "Walk" using 2-hands swinging across the body, or 2-legs moving like actual walking, 1-foot on the 12-o'clock position). Nevertheless, there is a need to further investigate the preferred combinations of gestures. Simply combining these gestures based on the designers' intuitions may result in combined gestures that may not be anatomically comfortable or feasible. We addressed this issue in Experiment III.

## 5. Experiment III: simultaneous game gestures

In this experiment, we seek to develop a consensus set for simultaneous game gestures. Prior to this experiment, we conducted a pilot experiment where three participants were asked to elicit simultaneous gestures for a set of combined game events (e.g., shoot + reload). Nevertheless, although all our participants were veteran gamers, we found little consensus between their gestures. They also mentioned that the user-elicitation process was substantially difficult, as they commented that it was difficult to imagine possible combinations of gestures. Furthermore, they commented that the lack of any existing simultaneous gestures for interaction adds to the difficulty as there were few reference points.

Based on this challenge, we adapted the original user-elicitation method (Morris et al., 2010; Wobbrock, 2009) and introduced a novel choice-based approach for investigating suitable simultaneous gestures.

### 5.1. Choice-based elicitation approach

On the basis of human-centered design (Kling and Star, 1998) and the original user-elicitation method (Morris et al., 2010; Wobbrock, 2009), we proposed a choice-based elicitation approach which is intended for cases when users do not have clear ideas or familiarity with the output space, possibly due to the novelty or the complexity of the requested output and thus may not effectively produce suitable gestures as claimed by Morris et al. (2010). In our case, since the notion of simultaneous gestures is relatively uncommon for users, with few reference points, we used a more suggestive, hinted approach (choice-based elicitation approach) to guide and assist our users to discover better, more suitable simultaneous gestures.

The main difference between a choice-based elicitation approach and a user-elicitation approach is a predefined-list of possible gestures. The predefined gesture list registers two data columns, one for possible gestures, and another column for game events, with the relationship as many-to-one, respectively. The predefined gesture list is intended to assist users to discover new or better possibilities when asked to perform certain gestures. In any case, users are also encouraged to define their own gestures and do not necessarily need to pick gestures from the predefined list when they preferred other options. In our case, the predefined gesture list was populated by results from our first two experiments.

The choice-based elicitation approach makes a basic assumption that in an unfamiliar scenario, users may lack knowledge of the design space, thus they may not always define effective gestures (as seen in Experiment I). The second assumption is that "recognition" may be better than "recall" in an unfamiliar elicitation scenario, i.e., given a list of choices (gestures), users will be

[2] Participants reported that left and right body parts (hand/leg) may be used simultaneously.

able to better define/discover more suitable gestures (i.e., recognition), as opposed to only imagination from scratch (i.e., recall or original generation) which participants in our pilot study found to be relatively more difficult when the scenario is unfamiliar. The third assumption the choice-based approach made is that all participants should be experienced gamers in both general video games as well as full-body games, and understand the requirements of intense gameplay, so as to increase the likelihood of a suitable gesture set.

Corresponding to these assumptions, we hypothesize the benefits of choice-based elicitation approach as three-fold: first, it may allow participants to discover better, more suitable gestures from the predefined gesture list than they have originally come up with. Secondly, it may also allow participants to better understand the design space through examples, and allow them to become more "creative" in creating their own gestures. Thirdly, it may increase the likelihood of achieving a high level of consensus among participants. These benefits were investigated in this experiment in which we found positive results. Further details are described in the discussion section. In any case, one should also consider the possible threat of the choice-based elicitation approach which may prevent the definition of more novel gestures, i.e., participants may become "lazy" in defining their own gesture, and rely mainly on the predefined-list. We considered this threat in our study design.

### 5.2. Participants

Twelve veteran gamers (all males, $M=22.4$ years) were recruited. We selected primarily veteran gamers due to their better understanding of the needs and requirements, and of games in general (i.e., intended users). All reported to have experience with Kinect and motion game gestures. Ten participants were right-handed. They were each paid $10.

### 5.3. Events

We derived five combinations of common simultaneous game events or closely simultaneous events (transactional events[3]) from typical interactions in video games and full-body games, similar to Experiment I. The elicited events included the various combinations of the following game events: Shooting events (Shoot + Reload + Walk/Run/Stealth walking + Use First Aid Kit + View point change), Racing events (Steer + Accelerate/Brake + Shift gear + Boost + View point change), Adventure/Role-playing events (Walk/Run/Stealth walking + Open door/Steal/Open chest/Pick Item/Push box/Slash + View point change), Fighting/Sport events (Punch + Kick + Guard + Dodge + View point change), System events (Open menu → Select menu item → Use item).

We used "/" (e.g., Walk/Run/Stealth Walking) to depict a category of events that can only occur one at a time, e.g., "Walk" cannot be executed at the same time as "Run". We assumed these events were interchangeable, so we asked participants to define simultaneous gestures that combine each of the following (e.g., Shoot + Walk, Shoot + Run, Shoot + Reload).

### 5.4. Procedure

At the start of the experiment, participants were asked to define gestures for the five combinations of game events (i.e., Shooting, Racing, Adventure, Fighting, System). The order of the five groups was asked in a randomized order across participants.

To reduce the cognitive burden on participants during the elicitation process, participants were instructed to first perform a set of two combined events (e.g., Shoot + Reload; Stealth walking + Steal). Then the number of combined events gradually increased by adding one more event (e.g., Shoot + Reload + Use First Aid Kit), until the combinations cover the full set of events (e.g., Shoot + Reload + Zoom scope in + Run + Use First Aid Kit + View point change). The gradually-increasing combined events were randomized. Although in reality, all these events may not (or very rarely) occur simultaneously (all at the same time), by asking participants to assume the most difficult case, we were able to detect and avoid any body part conflicts in the final combination of gestures elicited by participants. Our observations showed that participants consistently revised their gestures to be more simultaneous-friendly when the number of events increased.

Combined events were portrayed on a large display by displaying them separately, with a plus-sign "+" between target scenes along with the name of the event, e.g., "Shoot target scene" (event name) + "Reload target scene" (event name). By not mixing the target scenes, this separation minimizes a possible deterministic influence in the selection of gesture combinations. Most target scenes were sourced from Experiments I&II.

During the elicitation, to ensure the choice-based approach would not minimize the level of creativity of the participants, participants were asked to first do their best to perform the indicated combined gesture without the presence of a predefined gesture list. After the first attempt where participants performed each combined gesture or could not come up with a combined gesture, participants were exposed to a predefined list of possible gestures (populated from the results of Experiment I&II), where users could opt to refine their gestures based on the predefined list to produce their preferred simultaneous gestures. After their redefinition, to make sure that users' mix-and-match simultaneous gestures were anatomically and kinetically feasible and comfortable, we asked participants to again perform the actual simultaneous gestures quickly.

Similarly as in Experiment I&II, a think-aloud protocol was used where participants indicated the start and end of their performed gestures, the body parts used, and described the corresponding rationale. Other experimental place settings were similar to that used in Experiment 1. At the end of the experiment, we also asked the participants to rate the usefulness of the choice-based approach and performed semi-structured interviews. The experimental session took around 1.5-hours.

### 5.5. Results

The agreement score, the resulting set of simultaneous game gestures and subjective assessment of the choice-based approach were analyzed and reported.

#### 5.5.1. Agreement score

We used a similar approach to Experiment I to calculate the agreement score. Two gesture combinations were considered identical if each gesture in both combinations had similar trajectories and poses, and if the two combinations contained exactly the same set of gestures. For example, for the agreement score of the scenario "Shooting", a total of 12 participants and their corresponding 12 sets of combined gestures were considered. This event was divided into 6 groups of identical combinations with sizes of 5, 2, 2, 1, 1, and 1. Therefore, the agreement score for the

[3] Transactional events refer to events that are executed quickly in sequence. Within transactional events, the prior event enables the execution of the next event, whereas simultaneous events do not pose such technical requirements. We used → to depict transactional events, while using + for simultaneous events.

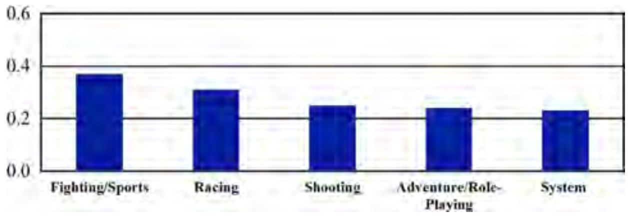



**Fig. 8.** Experiment III – overall agreement scores for the five combinations of game events sorted in descending order.


**Fig. 9.** User-defined simultaneous gesture set for shooting scenario. *Note*: the numbers depict the agreement score of each gesture, which are independent of the overall combination's agreement score.

Shooting scenario was

$$A_{shooting} = \left(\frac{5}{12}\right)^2 + \left(\frac{2}{12}\right)^2 + \left(\frac{2}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 + \left(\frac{1}{12}\right)^2 = 0.25$$

Fig. 8 shows the agreement score for each game event. The mean agreement score was 0.27, slightly lower than Wobbrock (2009) (0.32 and 0.28).

There was greatest consensus in the fighting and racing gesture sets, with shooting and adventure gestures scoring almost as low on agreement as the system gestures.

Agreement score of each gesture (which is independent of the combination score) was also calculated (see the numbers in Figs. 9–13), where low agreement score in the overall combination was shown to be influenced by the disagreement in some game events.

### 5.5.2. *User-defined simultaneous gesture set*

Of all the gestures performed, 95.3% of gestures originated from the predefined list - 61.1% of gestures were changed after seeing the proposed set and 34.2% were part of the predefined list although the gesture may have been invented by the participant without the knowledge of the set. Only 4.6% were newly created gestures elicited by participants. Although these gestures were newly created, participants reported that the predefined list provides a useful reference for imagining new gestures. All participants reported that without the gesture list, it would be difficult to imagine and design suitable gestures especially for simultaneous purposes with which they are not familiar. For example, participants said: "*The gesture list provides a good reference.*"; "*It's like a game, mixing and matching to find the best combinations.*"; "*The gesture list makes me much more creative*". This result clearly demonstrated the positive influence of this predefined list on the quality of the results.

Similar to past user-elicitation studies, we picked the largest group of combined gestures for our resulting set of simultaneous gestures (see Figs. 9–13). The priority of participants was simultaneous gestures

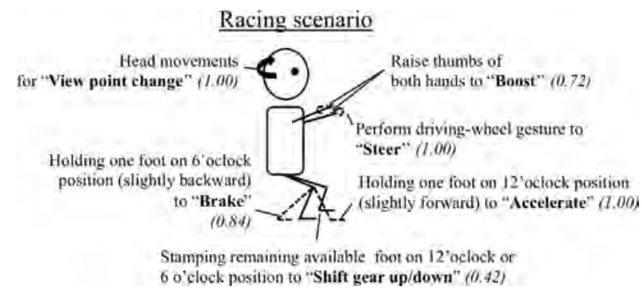



**Fig. 10.** User-defined simultaneous gesture set for racing scenario. The gesture set assumed a user in a sitting posture.


**Fig. 11.** User-defined simultaneous gesture set for adventure scenario. Participants reused navigation gestures (Run, Walk, Stealth Walking) from the "Shooting" scenario.

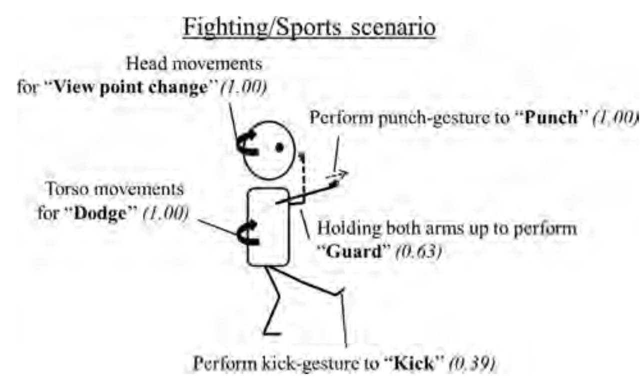



**Fig. 12.** User-defined simultaneous gesture set for fighting scenario. Gestures performed in this group were straightforward and mostly imitating real-world gestures due to the few body part conflicts.

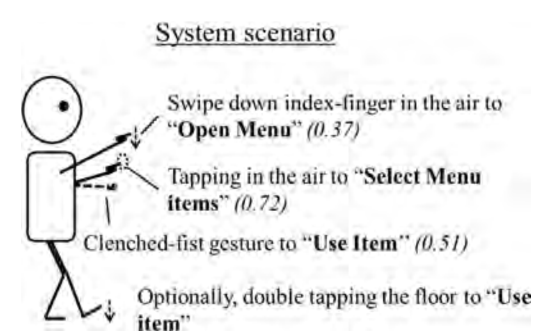



**Fig. 13.** User-defined simultaneous gesture set for system scenario. Unlike other simultaneous scenarios, these events were treated as transactional events (events in quick sequence) thus the same body parts were allowed to be used for different events.

that can be executed effortlessly and which are easy to learn/remember. Participants also preferred to reuse similar gestures across game events in different game groups.

*Shooting*–"Shoot" action was most preferably performed using one-hand imitating a gun-gesture and then flicking the gesture up (flicking the wrist up). Simultaneously, "Reload" action was most preferably performed by flicking that one-hand gun-gesture down. "Use First Aid Kit" was most preferably performed by performing a clenched-fist gesture using the remaining available hand. Head movement was most preferred for "View point change". For game avatar navigation, "Walk" was most preferably performed by holding one foot in the 12 o'clock position (i.e., pushing the foot slightly forward); "Stealth Walking" was most preferably performed by holding one foot in the 6 o'clock position (i.e., pushing the foot slightly backward); "Run" was most preferably performed by performing an actual run action, by jogging on the spot repeatedly.

Shooting scenario achieved a mean agreement score of 0.25. The differences came from minor disagreement of various gestures. For "Shoot", alternatively, some preferred a two-hand gun gesture (one hand as gun-gesture, while another hand holds the gun-gesture). For "Reload", some preferred using the other hand hitting the bottom of the gun-gesture to "Reload". For "Walk", some preferred to perform actual walk action. For "Stealth Walking", some preferred a slow-motion walking action.

*Racing*–"Steer" action was most preferably performed using two-hands to perform a driving-wheel gesture. "Accelerate" and "Brake" were most preferably performed by holding the foot to 12 o'clock and 6 o'clock respectively. While both hands are performing a driving-wheel gesture "Steer", most participants preferred to raise the thumbs of both hands to indicate "Boost". Head movement was most preferred for performing "View point change" action. For "Shift gear" action, the most preferable gesture was to stamp the remaining available foot to 12 o'clock and 6 o'clock position for shifting the gears up and down, respectively.

The Racing scenario achieved a mean agreement score of 0.31. The main differences came from the definition of "Boost", "Shift gear up/down" and "Brake" gestures. Alternatively, participants preferred pushing the driving-wheel gesture forward to "Boost". For "Shift gear up/down", some participants preferred stamping the foot on 3 o'clock and 9 o'clock position for shifting the gears up and down, respectively. For "Brake", some participants preferred to hold one foot on the center of the clock instead to the 6 o'clock.

*Adventure/role-playing*–For game avatar navigation, participants preferred similar gestures to navigation in Shooting events. Jump was most preferably performed by raising one leg slightly. Open door/Open chest/Push box/Steal/Pick Item/Slash events were most preferably performed using hand(s). For "Open Chest", "Open Door" and "Push box" actions, the most preferable gesture was to perform a palm gesture using 1-hand; for a "Slash" action, the most preferable gesture was to use one arm performing a slash-gesture; for "Steal" and "Pick Item" action, the most preferable gesture was to use one hand performing a grab-gesture. Head movement was most preferred to perform a "View point change" action.

The Adventure/Role-playing scenario achieved an agreement score of 0.24. The main differences came from whether participants preferred a hand or leg to perform certain actions. For example, to "Open Chest" or "Open Door", some participants performed a kick-gesture. Participants also showed different choices whether to use palm-gesture/punch-gesture/grab-gesture to "Open Chest", "Open Door", "Push Box", "Steal".

*Fighting/sports*–Events in this group were performed by fairly predictable and straightforward gestures, mostly imitating real-world actions due to the few body part conflicts, i.e., "Punch" was preferred by performing a punching gesture, with either hand. "Guard" action was preferred by holding both arms up on guard in front of the user's face. "Kick" action was preferred by performing a kick gesture. "Dodge" action was most preferably performed by moving the torso left and right accordingly. Head movement was most preferred for performing "View point change" action.

The Fighting/sports scenario achieved a relatively high agreement score of 0.37, due to obvious counterparts in the real-world. The main differences came from the definition of "Kick" gesture. Participants performed multiple variations of Kick such as straight-kick (most common), sideward-kick, knee-kick (using the knee to kick), feet-kick (using feet movement). For "Guard", some preferred one-handed guard.

*System*–"Open menu" command was most preferably performed by using the index-finger of either hand, and swiping down in the air to open a menu. "Select menu items" command was most performed by tapping in the air using the index finger and the "Use item" command was most preferably performed by performing a clenched-fist gesture. Nearly-equal in preference (6 participants), the "Use item" command can be optionally performed by "double clicking" (double tapping) one foot on the floor.

The System scenario achieved an relatively low agreement score of 0.23. "Open menu" could be alternatively performed by other symbolic gestures such as an opening-window gesture (virtually opening a window in the air), pushing-button gesture (pushing virtual button in the air using palm), and camera gesture. One could alternatively perform the "select menu item" using legs or eyes as the cursor. Given the relatively low agreement score of system events, it may be helpful to allow end-user customization on these events.

#### 5.5.3. Subjective assessment of choice-based elicitation approach

Participants were asked to rate the choice-based approach based on the Likert 7-scale rating (7 as strongly agree). Overall, participant responses to the choice-based approach were encouraging. The mean value is "Using the choice-based approach, the simultaneous gestures I performed were a good match for its intended purpose" (6.34); "The choice-based approach guides me to more easily discover better gestures" (6.23); "I would prefer using the choice-based approach in an unfamiliar elicitation scenario" (6.10); "The choice-based approach makes me more creative in creating new gestures" (5.32).

### 5.6. Discussion

We discuss the (i) effectiveness and challenges of the choice-based elicitation approach, (ii) informal comparison between our defined gesture set with existing Kinect games, and (iii) implications for full-body recognition technology.

#### 5.6.1. Effectiveness and challenges of the choice-based approach

Based on our subjective assessment and semi-structured interviews, we confirmed our hypothesis regarding the benefit of the choice-based elicitation approach. Participants commented on the usefulness of the choice-based approach in two primary ways: first, it allows participants to discover more efficient, suitable gestures (better than they originally imagined/elicited); second, it provides participants with a reference point from which they can further develop their own gestures. For example, participants commented that they felt that they become more "creative" after they saw the gesture list, and could subsequently come up with several creative gestures that did not exist in the original gesture list.

The choice-based elicitation approach also allows us to achieve a high consensus – agreement score of 0.27 – which is considered relatively reasonable when compared with past works (Wobbrock, 2009). We also compared the agreement score between the initial gesture set that participants came up with (before showing the predefined list) and with the final gesture set. The average agreement score for the final gesture set (0.27) was around two times higher than that of the initial gesture set (0.11). Overall, the choice-based elicitation approach showed a positive effect as we had hypothesized.

In any case, our study raised several issues for choice-based elicitation. *First*, participants commented that excessive choices would add cognitive burdens to participants, and suggested that approximately three to six choices would be the optimal number of choices. Too few or many choices could reduce the effectiveness of the choice-based approach, thus there is a need to choose which appropriate gestures are to be listed. We regard such a selection process to be at the core of any choice-based elicitation approach, and it requires solid, objective criteria if designers are to ensure that the intended quality of the predefined gesture list is as free of deterministic bias as possible. In our case, we selected the top-6 most-elicited (preferred) gestures from Experiment I&II to be listed. In other cases, the choices can depend on the criteria set by designers (e.g., top-X most comfortable). To prevent any possible ordering effect, it is also important to randomize the order of choices in the predefined-list. *Second*, there might be a possibility of a choice-based approach to reduce participants' level of creative imagination and confidence. It is thus important to not expose the predefined-list before participants elicit their first trial gesture or give up on elicitation due to lack of ideas. Participants should also be encouraged to define their original gesture when possible. *Third*, the approach requires a predefined gesture list whose creation can be time-consuming and it requires expert-participants who may not always be available. Based on these issues, there is a need to further investigate and compare the effectiveness of the choice-based elicitation approach with the traditional user-elicitation approach.

#### 5.6.2. Comparing our gesture set with existing games

We compared our user-defined gesture sets with some existing Kinect games. We found that many Kinect games do not accommodate the simultaneous use of gestures. For example, existing First Person Shooting Kinect games (e.g., Gunstringer, Ghost Recon, BlackWater) primarily used hand(s) as the input modality, such as raising the left hand up to indicate "Jump", using a punch-gesture for "Shoot", using a touching-elbow gesture for "Reload", and a hand to control "View point change". On the other hand, locomotion is automated, probably to reduce physical complexity. Although these hand gestures may be comfortable to execute, they do not facilitate the simultaneous use of game gestures. We proposed in our user-defined gesture set that "View point change" can be controlled by head movements, or alternatively by eye movements; "Running" can be achieved by jogging on the spot, "Reload" can be done by flicking the wrist down, so as to allow "Shoot" and "Reload" to be achieved simultaneously, instead of the touching-elbow gesture which would interrupt the "Shoot" action.

Existing racing Kinect games (e.g., Kinect Joy Ride, Forza 4) also mainly used a hand for interaction, such as performing the driving-wheel gesture to "Steer", and pulling back the wheel gesture to "Break". We proposed in our gesture set that leg modality can be leveraged for performing "Break", "Accelerate" and "Shifting gear" which can better accommodate the simultaneous use of gestures.

We found that several existing Kinect games limit the player's degree of control. For example, some games (e.g., Forza 4, Kinect Joy Ride, BlackWater) attempt to reduce physical complexity by removing and automating certain actions (e.g., Walking, Accelerate). However, although automating certain actions might remove physical complexity, all of our participants prefer the ability to control their movements (stop walking, walking, running). This is because the sense of control is important for achieving high immersion in the game. For example, one participant stated: "*Many existing Kinect racing games are boring because there is no realism. I cannot control my speed. I cannot feel immersed in racing; I am simply stretching my arm.*"

Some Kinect games used more "efficient" gestures (punch-gesture for "Shoot") over the natural gesture (gun-gesture for "Shoot") but this may not always be beneficial. For example, although the punch-gesture may perform more precise shooting, this might disrupt immersion. One participant stated: "*I would prefer gun-gesture over any other gesture for shooting, if not, I would never able to become immersed in the game*". The tradeoff between efficiency and naturalness should be carefully balanced.

For the System scenario, such as open menu or selecting menu items, existing Kinect games often employed a simple hand-pointing method (i.e., using hand as cursor). Our user-defined gesture set informs that users prefer a symbolic gesture for opening a menu (e.g., swiping down index-finger in the air), which can be less time-consuming compared to waving the hand for a long time to select menu items. Participants' priorities are to facilitate easy and fast access to system commands while in the middle of an intense gameplay.

For the Adventure and Sports scenario, we did not find much difference between existing games and our gesture set. The gestures are mostly pantomimic (one-to-one correspondence between the gesture and game events), e.g., punch-gesture for punching. This is probably because adventure and sport game events have obvious real-world counterparts in which there are few conflicts in body-part usage.

#### 5.6.3. Implications for full-body recognition technology

Many of the simultaneous gestures had strong implications for full-body recognition technology in highly interactive game scenarios. For example, with the need of simultaneous gestures in highly interactive scenarios, the recognizer should able to support recognizing (and distinguishing) simultaneous gestures at one time.

In addition, our resulting gesture set also indicates that the recognizer should be able to accommodate any slight variations of the similar gestures. For example, to perform "Kick", players may perform kick gesture at different heights and angles, with each gesture posing slight natural variations. Another example is "Shoot" where some participants may use one-hand gun gesture, while others may prefer holding the gun-gesture using two-hands.

With detailed gestures such as "gun-gesture", "index-finger swipe" gesture or "clenched-fist" gesture, the recognizer should be able to distinguish gestures to the details of the hand posture, number of fingers used, and which finger was used. For combinations of events with a low agreement score (e.g., system scenario), end-user customization should be enabled.

Considering the features of current technologies (e.g., Kinect), we believe that it is technologically feasible to meet these needs but the main challenge lies in detection and analysis algorithms (i.e., image processing). Some studies regarding natural variations (Fothergill et al., 2012) and finger detection (Liang et al., 2012a) have already been started and showed promising results, but less has been done on the detection of simultaneous gestures. This reflects the need for concurrent gestures detection and analysis algorithms, particularly for video games and other fast-paced interaction applications.

## 6. Design principles and guidelines

We summarized our findings into generalized design guidelines.

- *Guideline 1: Prioritize events*: One difficult decision designers may encounter is regarding the assignment of gestures. For example, it can be difficult to decide whether to use the most suitable body part or the alternative body part for one event.

Our studies show that participants define gestures according to priority. Participants commented that they would used the most suitable body part for crucial/frequently-used game events (e.g., hand for Shoot), while using symbolic/abstract gestures for less frequently-used game events (e.g., leg for open menu). The participants' general strategy is to minimize their effort to learn and remember combined gestures and also maintain the maximum degree of naturalness possible. Thus it is important for designers to prioritize which events are more significant than others.

- *Guideline 2: Prioritize immersion over playability*: Several existing Kinect games simplified the interaction by removing certain actions such as "Walk" or "Accelerate". However, this can have an adverse effect on player's immersion, as one of our participant stated, "*Game gestures should be the same as real-world actions wherever possible, if not, I can easily get distracted and I cannot feel immersed*". The lack of correspondence between the real-world and virtual-world can disrupt player's immersion and presence in the game. All our participants mentioned that immersion is among their top reasons in playing full-body game interactions instead of using traditional controllers (e.g., a keyboard).
- *Guideline 3: Use the hand moderately*: Although the hand is unsurprisingly the most preferred body input modality, it should be used moderately considering possible simultaneous use with other gestures. One should also consider the distribution of fatigue as one participant mentioned "*Using only hand is super tiring, I prefer my fatigue to be equally distributed across my body, which is also more fun.*"
- *Guideline 4: Exploit transferability between leg and hand and right and left body parts*: We found the possible transferability between hand and leg gestures, and between left and right body parts (e.g., left hand and right hand). This suggests that when certain body parts are occupied, designers may exploit this transferability when designing gestures. For example, when "Shift Gear" cannot be performed by a hand, the event then can be performed by a leg. In a similar manner, when the left hand is occupied with "Shoot", then one can use the right hand to perform other hand events (e.g., Use Item).
- *Guideline 5: Accommodate high tolerance for recognizing gestures*: We found that although all participants prefer a similar gesture for one event, they execute slightly differently in velocity and displacement based on their individual preferences. For example, for "Kick", participants kicked at different heights, angles and velocities. Thus recognition systems should tolerate such natural variations. Providing customization may help address this issue, e.g., allowing participants to customize different velocity/height for different kinds of "Kick".
- *Guideline 6: Gesture reuse*: Gesture reuse is important to assist the learnability of users. For example, our study found that reversible gestures were preferred by participants for reversible actions (Walk/Stealth Walking, Accelerate/Brake, Shift gear up/down). Similarly, participants preferred the same gesture for "Walk" in the Shooting scenario, and for "Accelerate" in the Racing scenario. Given this information, the idea of reusing gestures and designing reversible gestures has support.
- *Guideline 7: Design multiple gestures for one event, when needed*: We found individual differences between gamers and non-gamers. To accommodate differences in game expertise (novice vs. expert), one approach is to design more than one gesture for one event just as in desktop-based system. For example, designers can design two gestures for "Jump" - actual jump gesture for novice players and a more symbolic jump gesture (e.g., raising one leg slightly above the floor) for expert players.
- *Guideline 8: Reducing fatigue by a small amount can have great impact*: A small reduction in fatigue can have great impact in lengthy gameplay. For example, three participants mentioned that flicking the wrist up to "Shoot" is very different from flicking the lower arm in a long gameplay. Gestures should be designed to optimize the required efforts in lengthy gameplay.
- *Guideline 9: Design kinetically feasible combined gestures*: Designers should be careful to design kinetically-feasible combined gestures. Participants preferred simultaneous gestures that can be executed with ease and efficiency. Uncomfortable posture/gestures should be avoided (e.g., moving the head left while moving the torso right may cause injury). This suggests that full-body game designers should work with doctors who are specialized in motor control and coordination.

## 7. Limitations and future work

Our study resulted in a consensus set of simultaneous gestures that could be used in designing full-body game gestural interface. Nevertheless, our study posed two limitations: First, our simultaneous gestures assumed a typical set of combined game events, however, actual, more advanced commercial games might contain additional game events, thus the resulting simultaneous gestures might not be adequate. To address this limitation, we also provided a set of suitable and alternative body parts and gestures (Experiment II) and generalized guidelines (Experiment III) to guide designers. Second, due to the mechanistic nature of the user elicitation approach, we did not specifically examine the human capacity to perform simultaneous gestures (i.e., how many simultaneous gestures a user can perform at one time). To address this limitation, we conducted an informal follow-up study (without any feedback system) where we asked the twelve veteran gamers from Experiment III to perform simultaneous gestures according to the user-defined gesture sets of the five scenarios. Although they found that the gesture set to be kinetically comfortable, they commented that they can perform at best two to three simultaneous gestures and at most four. Further experiments will be required to verify this qualitative hypothesis.

An important next step is to comprehensively evaluate our user-defined gesture set for its technical feasibility and effectiveness. We will implement a full-body recognizer supporting the detection of our user-defined gesture set and we will measure the recognition accuracy and subjective assessment.

Another important next step is to further investigate the effectiveness of the choice-based elicitation approach. It will be beneficial to explicitly compare agreement scores and gesture sets resulting from the traditional user-elicitation approach and the choice-based elicitation approach in different scenarios. This will help evaluate and confirm the effectiveness and limitations of the choice-based elicitation approach.

## Acknowledgments

The authors would like to thank the members of The Center for Human Computer Interaction in Kochi University of Technology, particularly Okamoto Masashi, for their assistance. The authors would also like to thank the anonymous reviewers and the associate editor for their insightful comments.

## References

Alankus, G., Kelleher, C., 2012. Reducing compensatory motions in video games for stroke rehabilitation. In: Proceedings of CHI 2012. ACM, pp. 2049–2058.

Berkovsky, S., Coombe, M., Helmer, R., 2010. Activity interface for physical activity motivating games. In: Proceedings of IUI 2010. ACM, pp. 273–276.

Bianchi-Berthouze, N., 2013. Understanding the role of body movement in player engagement. Hum. Comput. Interact. 28 (1), 40–75.
Chen, X.A., 2012. Body-centric interaction with mobile devices. In: Proceedings of TEI 2012. ACM, pp. 385–386.
Cohn, G., Morris, D., Patel, S., Tan, D., 2012. Humantenna: using the body as an antenna for real-time whole-body interaction. In: Proceedings of CHI 2012. ACM, pp. 1901–1910.
Connell, S., Kuo, P.-Y., Liu, L., Piper, A.M., 2013. A Wizard-of-Oz elicitation study examining child-defined gestures with a whole-body interface. In: Proceedings of IDC 2013. ACM, pp. 277–280.
Fothergill, S., Mentis, H., Kohli, P., Nowozin, S., 2012. Instructing people for training gestural interactive systems. In: Proceedings of CHI 2012. ACM, pp. 1737–1746.
Freeman, D., LaPierre, N., Chevalier, F., Reilly, D., 2013. Tweetris: a study of whole-body interaction during a public art event. In: Proceedings of C&C 2013. ACM, pp. 224–233.
Gamespot, 2011. Rise of Nightmares Review. ⟨http://www.gamespot.com/rise-of-nightmares/reviews/⟩, (accessed 30-03-2013).
Gerling, K., Livingston, I., Nacke, L., Mandryk, R., 2012. Full-body motion-based game interaction for older adults. In: Proceedings of CHI 2012. ACM, pp. 1873–1882.
Harrison, C., Ramamurthy, S., Hudson, S.E., 2012. On-body interaction: armed and dangerous. In: Proceedings of TEI 2012. ACM, pp. 69–76.
Hernandez, H.A., Ye, Z., Graham, T.N., Fehlings, D., Switzer, L., 2013. Designing action-based exergames for children with cerebral palsy. In: Proceedings of CHI 2013. ACM, pp. 1261–1270.
Hoysniemi, J., 2006. International survey on the Dance Revolution game. Comput. Entertain. 4 (2).
Höysniemi, J., Hämäläinen, P., Turkki, L., 2004. Wizard of oz prototyping of computer vision based action games for children. In: Proceedings of IDC 2004. ACM, pp. 27–34.
Hutchins, E., Hollan, J., Norman, D., 1985. Direct manipulation interfaces. Hum. Comput. Interact. 1 (4), 311–338.
Isbister, K., DiMauro, C., 2011. Waggling the form baton: analyzing body-movement-based design patterns in Nintendo Wii games, toward innovation of new possibilities for social and emotional experience. In: Whole Body Interaction. Springer, pp. 63–73.
Isbister, K., Rao, R., Schwekendiek, U., Hayward, E., Lidasan, J., 2011. Is more movement better?: A controlled comparison of movement-based games. In: Proceedings of FDG 2011. ACM, pp. 331–333.
Kane, S.K., Wobbrock, J.O., Ladner, R.E., 2011. Usable gestures for blind people: understanding preference and performance. In: Proceedings of CHI 2011. ACM, pp. 413–422.
Kling, R., Star, S.L., 1998. Human centered systems in the perspective of organizational and social informatics. SIGCAS Comput. Soc. 28 (1), 22–29.
Kray, C., Nesbitt, D., Dawson, J., Rohs, M., 2010. User-defined gestures for connecting mobile phones, public displays, and tabletops. In: Proceedings of MobileHCI 2010. ACM, pp. 239–248.
Liang, H., Yuan, J., Thalmann, D., 2012a. 3D fingertip and palm tracking in depth image sequences. In: Proceedings of MM 2012. ACM, pp. 785–788.
Liang, H.-N., Williams, C., Semegen, M., Stuerzlinger, W., Irani, P., 2012b. User-defined surface+motion gestures for 3d manipulation of objects at a distance through a mobile device. In: Proceedings of APCHI 2012. ACM, pp. 299–308.
Macvean, A., Robertson, J., 2013. Understanding exergame users' physical activity, motivation and behavior over time. In: Proceedings of CHI 2013. ACM, pp. 1251–1260.
Mauney, D., Howarth, J., Wirtanen, A., Capra, M., 2010. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In: Proceedings of CHI EA 2010. ACM, pp. 4015–4020.
Mizobata, R., Silpasuwanchai, C., Ren, X., 2014. Only for casual players?: investigating player differences in full-body game interaction. In: Proceedings of Chinese CHI 2014. ACM, pp. 57–65.
Morris, M.R., Wobbrock, J.O., Wilson, A.D., 2010. Understanding users' preferences for surface gestures. In: Proceedings of GI 2010. CIPS, pp. 261–268.
Müller, J., Walter, R., Bailly, G., Nischt, M., Alt, F., 2012. Looking glass: a field study on noticing interactivity of a shop window. In: Proceedings of CHI 2012. ACM, pp. 297–306.
Nacenta, M.A., Kamber, Y., Qiang, Y., Kristensson, P.O., 2013. Memorability of pre-designed and user-defined gesture sets. In: Proceedings of CHI 2013. ACM, pp. 1099–1108.
Nielsen, M., Störing, M., Moeslund, T.B., Granum, E., 2004. A procedure for developing intuitive and ergonomic gesture interfaces for HCI. In: Gesture-Based Communication in Human–Computer Interaction, vol. 2915. Springer, pp. 409–420.
Nijhar, J., Bianchi-Berthouze, N., Boguslawski, G., 2012. Does movement recognition precision affect the player experience in exertion games? In: Proceedings of INTETAIN 2011. Springer, pp. 73–82.
Norton, J., Wingrave, C.A., LaViola, Jr., J.J., 2010. Exploring strategies and guidelines for developing full body video game interfaces. In: Proceedings of FDG 2010. ACM, pp. 155–162.
Obaid, M., Häring, M., Kistler, F., Bühling, R., André, E., 2012. User-defined body gestures for navigational control of a humanoid robot. In: Proceedings of ICSR 2012. Springer-Verlag, pp. 367–377.
Pasch, M., Bianchi-Berthouze, N., van Dijk, B., Nijholt, A., 2009. Movement-based sports video games: investigating motivation and gaming experience. Entertain. Comput. 1 (2), 49–61.
Peer, F., Friedlander, A., Mazalek, A., Mueller, F.F., 2011. Evaluating technology that makes physical games for children more engaging. In: Proceedings of IDC 2011. ACM, pp. 193–196.
Ruiz, J., Li, Y., Lank, E., 2011. User-defined motion gestures for mobile interaction. In: Proceedings of CHI 2011. ACM, pp. 197–206.
Samanci, O., Chen, Y., Mazalek, A., 2007. Tangible comics: a performance space with full-body interaction. In: Proceedings of ACE 2007. ACM, pp. 171–178.
Schaeffer, B., Flider, M., Kaczmarski, H., Vanier, L., Chong, L., Hasegawa-Johnson, Y., 2003. Tele-sports and tele-dance: full-body network interaction. In: Proceedings of VRST 2003. ACM, pp. 108–116.
Schönauer, C., Pintaric, T., Kaufmann, H., 2011. Full body interaction for serious games in motor rehabilitation. In: Proceedings of AH 2011. ACM, pp. 4:1–4:8.
Shoemaker, G., Tsukitani, T., Kitamura, Y., Booth, K.S., 2010. Body-centric interaction techniques for very large wall displays. In: Proceedings of NordiCHI 2010. ACM, pp. 463–472.
Tholander, J., Johansson, C., 2010. Design qualities for whole body interaction: learning from golf, skateboarding and bodybugging. In: Proceedings of NordiCHI 2010. ACM, pp. 493–502.
Vatavu, R.-D., 2012. User-defined gestures for free-hand TV control. In: Proceedings of EuroiTV 2012. ACM, pp. 45–48.
Williamson, B.M., Wingrave, C., Laviola, J.J., Roberts, T., Garrity, P., 2011. Natural full body interaction for navigation in dismounted soldier training. In: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2011), pp. 1–8.
Wobbrock, J.O., Aung, H.H., Rothrock, B., Myers, B.A., 2005. Maximizing the guessability of symbolic input. In: Proceedings of CHI EA 2005. ACM, pp. 1869–1872.
Wobbrock, J.O., Morris, M.R., Wilson, A.D., 2009. User-defined gestures for surface computing. In: Proceedings of CHI 2009. ACM, pp. 1083–1092.
Yim, J., Graham, T.C.N., 2007. Using games to increase exercise motivation. In: Proceedings of Future Play 2007. ACM, pp. 166–173.