Understanding Gesture Articulations Variability

Orlando Erazo, Yosra Rekik, Laurent Grisoni, José Pino

Understanding Gesture Articulations Variability

Orlando Erazo[1], Yosra Rekik[2], Laurent Grisoni[3], and José A. Pino[4]

[1] Universidad Técnica Estatal de Quevedo, Quevedo, Ecuador
`oerazo@uteq.edu.ec`
[2] University of Lille Sci. & Tech, CNRS, INRIA, Lille, France
`yosra.rekik@inria.fr`
[3] University of Lille Sci. & Tech, CNRS, INRIA, Lille, France
`laurent.grisoni@univ-lille1.fr`
[4] Department of Computer Science, University of Chile, Santiago, Chile
`jpino@dcc.uchile.cl`

Abstract. Interfaces based on mid-air gestures often use a one-to-one mapping between gestures and commands, but most remain very basic. Actually, people exhibit inherent intrinsic variations for their gesture articulations because gestures carry dependency with both the person producing them and the specific context, social or cultural, in which they are being produced. We advocate that allowing applications to map many gestures to one command is a key step to give more flexibility, avoid penalizations, and lead to better user interaction experiences. Accordingly, this paper presents our results on mid-air gesture variability. We are mainly concerned with understanding variability in mid-air gesture articulations from a pure user-centric perspective. We describe a comprehensive investigation on how users vary the production of gestures under unconstrained articulation conditions. The conducted user study consisted in two tasks. The first one provides a model of user conception and production of gestures; from this study we also derive an embodied taxonomy of gestures. This taxonomy is used as a basis for the second experiment, in which we perform a fine grain quantitative analysis of gesture articulation variability. Based on these results, we discuss implications for gesture interface designs.

Keywords: Mid-air Gestures, Whole Body Gestures, Gesture Articulation, Gesture Variability, Gesture Taxonomy.

1   Introduction

Mid-air gestural interactions are increasingly popular, with applications for entertainment [20], operating rooms [11], museums [23] or public spaces [44]. Although the evolution of new hardware devices, such as MS Kinect or Intel RealSense, has contributed significantly to this proliferation, there are still some challenges that need to be addressed. One of these challenges is the association between gestures and meanings/commands. There are various proposals that could be used for this goal, which mainly focus on a one-to-one mapping (e.g., [45, 27]). However, users may perform a specific gesture in different manners expecting it has equal meaning. For

example, a user may display a menu by "drawing" in the air a letter "M" in the horizontal/vertical/sagittal plane, using one or two hands, performing one or more strokes, holding a hand pose, or also using other body part(s). The absence of this variability of user's gestures in the mapping between gestures and commands is a design limitation that may conduct to poor user experiences. This problem may be tackled by understanding gesture articulations; i.e., the ways in which users produce gestures in the air.

Regardless of the usefulness of gestures, little is known about how users articulate gestures. One of the focuses of previous work is how to choose gestures to define a gesture vocabulary. Several criteria may be used for this purpose. One of them may be the appropriateness of the candidate gestures to the intended meaning by considering features like user preferences [45, 27, 24, 14], social acceptability [35], teaching/learning [15], or memorability [25]. The time efficiency to perform the various gestures is another criterion to select gestures [10]. Other researchers have focused on notations or formal gesture specifications [6, 38], gesture recognition [43], and also studied specific gestures [26, 13]. Though some efforts have been made for analyzing user behavior on performing gestures [12, 2], variability of gesture articulations has only been studied in the case of multi-touch input [32, 33]. However, those findings should not be directly applied to mid-air gestures given the different behavior and body parts that can be involved. Moreover, most studies have focused on gestures produced with hands; the few ones related to gestures made with other body parts have not analyzed variability in depth (e.g., [28, 17]). Consequently, the answers to the following questions are still missing: How do users perform gestures moving their hands in the air? Do they use only hands or also use other body parts? Do they use one or two hands? Do they perform gestures using one or several movements?

Given these open questions, we advocate in this work for the need of more in-depth user studies to better understand how users articulate gestures without haptic contact under unconstrained conditions toward finding a relation between many gestures to one command. We are mainly interested in analyzing the variability of gesture articulations produced by a same user for a same gesture type, but we also summarize results for all participants and also identify differences among participants. To reach this goal, we conducted a two-phase user study in which participants produced various articulations for the same gesture type. The study was done from a broad perspective, not focusing just on hand gestures; i.e., participants were neither instructed nor restricted to use only the hands. The collected data allowed analyzing the variability of user's gestures, and the results show users' preferences to produce the same gesture type in various articulations.

Thus, this paper contributes as follows: (1) a qualitative model to understand user's gestures execution; (2) a taxonomy of whole body gestures, utilized to classify the performed gestures and make comparisons with related works; (3) a detailed qualitative and quantitative analysis of gesture variability; (4) implications for designing applications based on gestures. We hope our results will prove useful to designers and practitioners interested in maximizing the flexibility of gestures set designs. In the long run, the presented exploration and contributions are first steps toward designing many gestures to one command.

2   Related Work

The main design problems of gesture interfaces are shifting toward finding the optimum mapping from gesture to function. To tackle this issue, a set of toolkits have been designed to advise practitioners on how to obtain a high recognition rate or help them on how to organize gesture sets [4]. Spano et al. [38] proposed GestIT, a compositional, declarative meta-model for defining gestures based on Petri Nets. Choi et al. [6] developed a method to organize and notate user-defined gestures in a systematic approach. These works can help develop gesture interfaces and notate gestures respectively, but they are restricted to gesture specification and do not cover execution of those gestures according to users' mental models. Other similar work could be borrowed from touch interfaces but with the same limitation. Model-based evaluation can also be used to analyze gestures. Current existing models allow estimating performance scores to produce mid-air gestures [10, 9, 39] that can be used to compare gestures and/or to create gesture sets, but they do not consider the possibility of performing a same gesture in different manners.

Consequently, two options have been defined in order to find the best mappings: (1) designers can rely on their own expertise or (2) organize user studies. However, the first option often leads to arbitrary gesture sets [36] that do not take into account users' mental model or opinions, resulting in misdesigns and frustrating user experiences [8]. Involving users into the design process represents a viable alternative for collecting important data to inform design. Wobbrock et al. [45] introduced a methodology for eliciting gesture commands from users. Follow-up work verified this methodology for mid-air gestures [40, 31, 28, 17, 29, 37]. The methodology consists of presenting non-technical users the effects of gestures and eliciting the causes meant to invoke them. Later, Morris et al. [24] proposed that user elicitation studies could be improved by generating various gestures, priming users, and involving partners. Furthermore, Hoff et al. experimentally tested the first and second suggestions in a follow-up work [14]. Though this and other methodologies that can be found in the literature (e.g., [27]) are user-centered, they only considered the relation between one gesture and one command.

Other previous studies have reported on users' gestures articulation variations and preferences. Nancel et al. [26] demonstrated that bimanual gestures were faster than the one-handed ones for panning and zooming given that these actions are complementary. Next, trying to understand mid-air hand gestures, Aigner et al. [2] found that users would prefer both different types of gestures and number of hands depending on the meaning of the gesture. Actually, as Silpasuwanchai and Ren [37] explain and suggest, a same gesture may not be valid in all cases, and hence, more than one gesture should be used for one command when needed. Meanwhile, two works focusing on user defined gestures for augmented reality and controlling robots, showed that more than half of gestures proposed by participants for the requested tasks fell in the dynamic category (i.e., gestures are expressed using strokes) or unimanual category [31, 28]. Later, Henschke et al. [12] verified that user's gestures changed over repetitions. Despite this progress, articulation variability of gestures has not been analyzed in a comprehensive way.

Recently, Rekik et al. [32] presented a comprehensive investigation on how users vary multi-touch gestures under unconstrained articulation conditions. Based on a proposed taxonomy, they evaluated user gesture variability and concluded that users equally use one or two hands and gestures are achieved using parallel or sequential combination of movements. They also noted that it is important to know whether their findings "can be applied to other type of gesture detection devices which do not require a contact surface" [32]. However, no study yet exists for investigating the various manners in which users could articulate mid-air gestures. This work tries to fill this gap by collecting and analyzing gesture articulations variability. Hence, we followed the methodology of [32], making the needed modifications, to understand mid-air gestures.

3   User Study

Our study was composed of two tasks. Like Rekik et al. [32], the goal of the first task was to familiarize participants with the experimental setup and to analyze their interaction styles using an uncontrolled experimental procedure. The analysis of this data served as a basis to derive both a qualitative model on user conception and production of gestures, and a taxonomy of whole body gestures. This taxonomy and this model were used to analyze the second task. The goal of the second task, like [32], was to perform a quantitative analysis of how users articulate mid-air symbolic gestures by following some specific instructions and exploring several ways to do it. The remaining details of our study are described below.

3.1   Participants

Twenty people (mean age = 27.5 years, SD = 4.3, 6 female) took part in the study. They were invited by mailing lists and social networks. Eighteen participants were right-handed. Graduate students and researchers from Europe, South America, Africa and Asia agreed to volunteer for the study (UI designers were not allowed participating). Ten participants self-declared having some previous experience on mid-air interaction for gaming (e.g., using Kinect).

3.2   Apparatus

The hardware setup, mounted in our laboratory, consisted of a notebook, a Kinect sensor and a display. The Kinect and the display were connected to a notebook equipped with an Intel Core i7 processor, 8 GB of RAM. Participants stood about 2.5 m away from the display (with a size of 1.8×1.4 m and a resolution of 1024×768 pixels). The Kinect was placed below the display at a height of 1 m. Kinect RGB data was used to videotape the interaction and give participants some feedback while performing gestures. The developed application interface consisted of an augmented video blending UI controls and the real environment. Augmented video was used trying to avoid participant distractions while performing the tasks [10, 9]. It means

participants were able to see themselves in the projected display like looking in a mirror. Also, the application showed the instructions to participants (i.e., the name of each required gesture type), the progress, and whether the gesture was considered right or wrong (i.e., a green check or a red "X" respectively). Inputs were considered right if participants followed the instructions correctly. No additional visual feedback was provided to prevent some effect [10, 32].

In addition, we decided to make use of a Wizard of Oz design [7, 27, 12] taking into account that participants were instructed to perform gestures in the articulation they preferred rather than being limited to the capabilities of a recognition system. The main idea was participants believe they are interacting "normally" with the system that provides the results/information when responses are in fact given by an experimenter (the "wizard"). Hence, the experimenter pressed some keys with the aim the system responded according to the participant's input (i.e., start, end, and right/wrong).

3.3 Procedure and Tasks

Participants were given the exact procedure for each task in paper sheets to avoid instructing them in various manners. They executed the gestures guided by the software when they were ready to start. Participants were also requested and encouraged to think-aloud while performing each gesture in both tasks (i.e., perform each gesture describing it aloud) [32]. Moreover, they had a short rest at the end of each task. Finally, we asked participants to provide some information by filling out a questionnaire. The specific instructions and differences of each task are described below.

Task 1: Open-ended gestures. We asked participants to produce as many gestures as possible that came to their minds such as gestures that had a meaningful sense to them or gestures that they would use to interact with applications (i.e., gestures had to be realistic for practical scenarios, easy to produce, easy to remember, and different enough from one another that they could be used for different actions). In addition, participants were asked to describe the gestures they performed using the think-aloud protocol. Participants decided when to start and stop. Thus, this task finished when participants could not produce additional gestures.

Task 2: Goal-oriented gestures. Participants received more explicit instructions to carry out the second task. They had to create gestures for a set of 20 gesture types (see Figure 1). The gestures set includes geometric shapes, letters, numbers, and symbols similar to the ones used in previous works on touch and touchless interactions [32, 33, 34, 3, 42, 10]. The selection of these gestures aimed to be general enough trying participants could articulate them without visual representations and under unconstrained conditions [33]. Moreover, this task focused on symbolic mid-air gestures instead of gestures for traditional actions utilized in mid-air interactions (e.g., pan, zoom, rotate, etc.), given the versatility of symbols to be generalized for other applications [34].

The software randomly asked participants to perform a gesture by showing only its name. Paper sheets were presented on demand just for the case in which a participant was not familiar with the corresponding gesture type. Participants had to think of the gesture articulations in which they would produce the gesture type after reading its name. They were instructed to create as many different gesture articulations as possible, trying to increase variety and creativity [24]. Though participants were free to select the different gesture articulations, we gave them the requirement that executions should be realistic for practical scenarios, i.e., easy to produce and reproduce later. Furthermore, we provided no instructions on the body part(s) to be used to produce gestures. This decision enabled us to analyze whether users could use and/or prefer using body parts other than hands to articulate gestures.
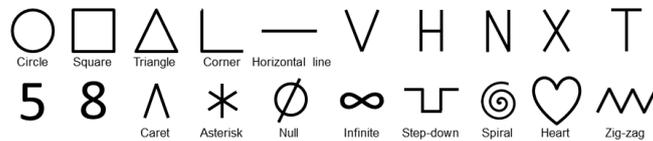
Fig. 1. The set of 20 gesture types used in the experiment.

4   Open-ended task results

The first task of the study consisted in producing gestures by following an uncontrolled experimental procedure. Participants had to appeal to their imagination to carry out the task. Thus, the gathered data allowed analyzing the various gesture articulations conceived and produced by participants.

4.1   General Observations

Regarding the collected quantitative data, participants performed a total of 117 gestures in this task. Each participant performed from 3 to 14 gestures (mean = 6, SD = 3, median = 5, mode = 3). We obtained the following features by analyzing all these gestures:

— All participants performed drawing or writing symbolic gestures (such as shapes, letters, or numbers). Furthermore, participants executed gestures for traditional interactions/actions. For instance, 12 participants produced gestures for scale, swipe, drag & drop, etc. 14 participants produced gestures for actions like typing in the air, wave, tap, clap, etc. One participant added gestures for selecting a group of objects while another participant added gestures for making a copy or paste actions.
— Participants produced both stroke and hold (keep a pose a short period of time) gestures. For instance, 17 participants performed at least one gesture stroke. Nine participants utilized poses at least once.
— Gestures were produced using one or more body parts. 15 participants executed gestures using both one and two upper limbs (i.e., arms, hands or fingers). From the remaining five, four utilized only one upper limb for all gestures, whereas one

used both upper limbs (but he also utilized the whole body in a few cases). Moreover, four out of the 20 also employed other body parts (different than upper limbs).

— 18 participants performed gestures using both single and multiple movements. The two remaining participants executed only single movement gestures, whereas nobody used only multiple movement gestures. Multiple movements were either parallel or sequential.

— 19 participants executed most of the gestures starting from a resting position, whereas the remaining one performed 75% of his gestures continuously, i.e., participants produced the gestures without adopting positions of resting or relaxation between strokes/movements only in a few cases. Resting positions consisted in having both hands/arms down and close to the hips or to the torso in most cases.

4.2   GCP: A Model for Gestures Conception and Production

Before analyzing in detail variability of gesture articulations it is necessary to understand how users conceive and produce gestures. Related works from psychology and neuroscience provide the basis to reach this goal. On the one hand, the framework proposed by Wong et al. [46] to define motor planning can be adapted and used to explain the user gesture conception. On the other hand, Kendon's and McNeill's proposals allow for analyzing gestures produced in midair [16, 22]. Based on both these related works and the aforementioned results, we derived GCP, a qualitative model for gestures conception and production (see Figure 2).
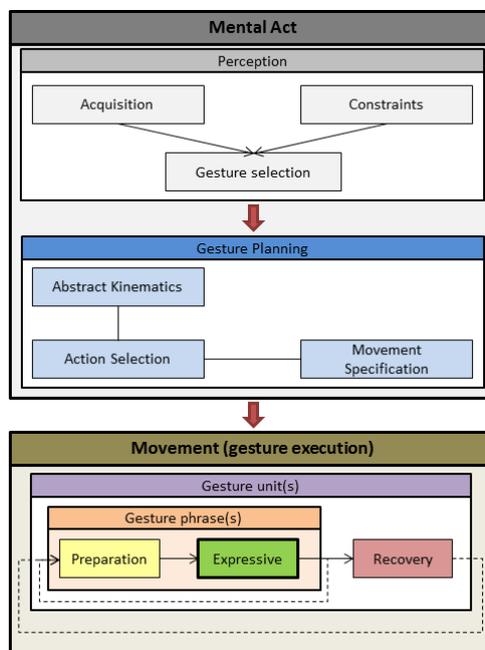
Fig. 2. GCP, a model on user conception and production of gestures (based on [46, 16, 22]).

According to GCP, a user initially needs to mentally prepare (think) before executing a gesture. During the mental act phase, which was adapted from [46], the user establishes/defines a motor goal (i.e., "what" processes), and then, s/he specifies the manner in which s/he will achieve that goal (i.e., "how" processes). In other words, the mental act consists of perception and gesture planning.

The user selects/defines/forms motor goals during perception. Perception consists of three processes: (1) acquisition or identification of proposed symbols/referents; (2) application of rules/constraints to perform gestures (e.g., the instructions given in our study); and (3) selection of the gesture to be performed.

Gesture planning in turn refers to how the required gesture will be produced; i.e., it defines the specific movement(s) to execute the gesture. It also involves several processes that may occur in sequence or in parallel. These processes are: (1) abstract kinematics of the gesture (i.e., how the gesture will look), which is optional for single gestures such as pointing; (2) selection of body end-effector(s) action (i.e., how the effectors/body parts will achieve the goal); (3) complete specification of motor commands needed to produce the gesture. The occurrence of these processes allows translating the motor goal into the movement that will correspond to the intended gesture.

Several phases can be observed when the user executes the gesture. Actually, the results obtained in the first task of our study are consistent with the temporal nature of gestures, which is described in terms of phases, phrases, and units [16, 22, 18]. The gesture execution starts with an optional physical preparation of the effectors selected during the mental act. During this preparation phase, the user moves the body part(s) from a resting or relaxation position to the position in which the meaning of the gesture is manifested. The peak of effort and shape are clearly expressed in the expressive phase, which is an obligatory phase. It must take the form of either stroke or hold [18]. These two phases, preparation and expression, are encapsulated in a gesture phrase (g-phrase). At the end of a g-phrase, the user may produce another one or continues to the next phase. Recovery [16] (or retraction [22]) is the last optional phase in which the effectors return to their initial or resting positions. Recoveries are not part of any g-phrase but together with one or more g-phrases are grouped into "kinematic units" labeled as gesture units (g-units). Thus, a g-unit is the "entire excursion from the moment the effectors begin to depart from a position of relaxation until the moment when they finally return to one" [16].

4.3   A Taxonomy of Mid-Air Gestures

Despite the GCP model allows capturing perception, planning, and execution of gestures proposed by our participants, it is insufficient to reach our goal. GCP only models the execution of gestures in a general way (i.e., looking at their temporal nature), and hence, it does not permit doing a fine grain analysis of gesture articulation. Therefore, we propose an embodied taxonomy of whole body gestures. In general, the different levels of this taxonomy cannot be seen serially or as partitionable attributes because these levels represent indivisible aspects of user gestures.

Table 1 depicts the proposed taxonomy. Overall it captures physicality, movement composition, and structure of gestures. Physicality captures the end-effectors or body

parts used to perform the gesture. Interestingly and contrarily to multitouch input [32], this level does not capture only hand gestures; it considers gestures performed with the whole body as well as with upper and lower limbs with the corresponding subdivisions. Movement level refers to the set of movements that compose a gesture. When a gesture is composed of more than one movement, these movements can be entered in parallel (i.e., multiple movements are articulated at the same time, e.g., using two hands to draw two sides of a "heart" shape at the same time) or in sequence (i.e., one movement after the other, such as in drawing the "plus" sign with one hand). However, not all gestures can be produced with parallel movements. In fact, only gestures containing a symmetry can be performed with parallel movements. Interestingly, wherever a gesture presented a symmetry, participants produced synchronous parallel movements to create that part of the gesture (i.e., some movements of the gesture were articulated with one movement at the same time and others were articulated in parallel, e.g., using two hands at the same time to draw the two diagonal symmetric lines of a "triangle" shape and then one hand to draw the horizontal line). The gesture may also involve a sequence of parallel movements (e.g., use both hands at the same time to articulate the two vertical lines of the "square" shape and then again use both hands to articulate the two horizontal lines). The last level refers to the structure of the gesture, which captures the state of the articulated movements. Considering this, a gesture may be a combination of single (static) or a series (dynamic) of poses that follow or not a path (like in [45]).

Table 1. A taxonomy of mid-air gestures.

|			
Physicality	Whole Body		
	Upper limbs	Arms	One-arm
			Two-arms
		Hands	One-handed
			Two-handed
		Fingers	One-finger
			Multi-fingers
	Lower limbs	Legs	One-leg
			Two-legs
		Feet	One-foot
			Two-feet
	Other, (e.g., head)		
Movement	Single		
	Multiple	Parallel	
		Sequential	Sequence of single movement
			Sequence of parallel movement
			Sequence of single and parallel movement
Structure	Static pose		
	Dynamic pose		
	Static pose with path		
	Dynamic pose with path		

5 Goal-Oriented Task Results

This section presents the results obtained from the second task in which participants produced various gesture articulations for specific symbolic gesture types.

5.1 Gesture Variations

Participants were instructed to propose as many articulation variations as possible for each gesture type. We collected 1,237 total samples for our set of 20 gesture types. On the average, our participants proposed 3.1 variations per gesture type (SD = 0.4, see Figure 3), a result that is in agreement with the findings of [30] for action gestures (mean 3.1, SD = 0.8). A Friedman test revealed a significant effect of gesture type on the number of variations ($\chi^2(19)$ = 96.053, $p < .001$). The "*" and "step-down" gestures presented the lowest number of variations (2.2 and 2.6 variations on the average respectively). The gesture with the maximum number of variations was "X" (3.9 on the average) for which our participants managed to easily decompose it into individual strokes that were afterward combined in many ways in time and space using different gesture physicality and structure (see Figure 4). For example, only 3 participants produced less than 4 gesture articulations for "X". These first results suggest that the specific geometry of the gesture enables users with different affordances of how to articulate that shape. Likely, the mental representation of a gesture variation implies a particular type of articulation which is tightly related to the gesture shape. We can also remark that for all gesture types the maximum number of variations was 4 or 5, except "triangle" and "V" that had 6 variations. The minimum number of variations was 1, except "square", "corner", "X", and "T" with 2 variations. Meanwhile, the means (averaged over all subjects) of the percentages of the gesture types for which each participant produced at least 4, 3, and 2 gestures were 49%, 70%, and 91% respectively. This result also suggests that, for some users and for some gesture types, the number of gesture articulation variations can be limited which can be explained by the previous practice but also by the geometrical shape of the gesture.
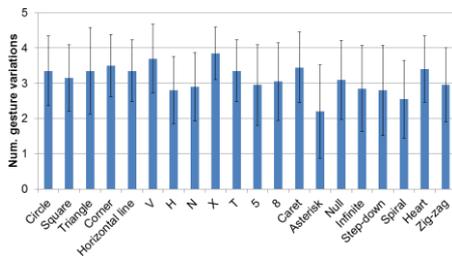
Fig. 3. Number of variations for each gesture type.
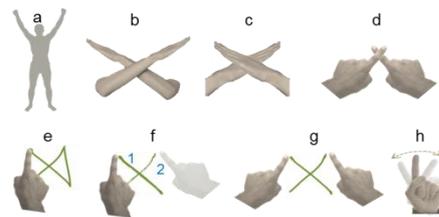
Fig. 4. Various articulation patterns for the "X" symbol produced with several poses (a-d); number of strokes (e-h), sequential (f), and parallel movements (g, b-d), using the whole body (a), arms (b), hands (c) and fingers (d-h). Numbers on strokes indicate stroke ordering.

5.2   Physicality Breakdown

Figure 5 shows the ratios (averaged over all users) of gestures for each gesture type and overall. We used only single levels for arms, fingers and lower limbs to simplify the analysis. A Friedman test revealed a significant difference in the ratios (averaged over all symbols) between the six physicality types ($\chi^2(5) =17.84$, $p < .001$). Post-hoc Wilcoxon signed-rank tests confirmed not significant differences only between arms and two-handed levels.

Friedman tests also revealed significant effect of gesture type on the ratio of the physicality types (Table 2). Referring to each physicality level, participants preferred one-handed gestures in all cases (53.4% on the average), especially for gesture types that may be considered as more difficult or strange to articulate. This is precisely the case of "spiral" that got the highest value and differed significantly from the other gesture types (except "infinite") according to the corresponding post hoc test. Similarly, the gesture types "asterisk", "infinite", and "zig-zag", that also obtained high values, were not significantly different from one another; and they showed differences between the other gesture types (except for "H", "N", and "5"). The next types are gestures made with fingers (one or multiple) and with two hands but with a short difference according to overall values (17.7% and 13.9% respectively). Notably, the highest values of finger type were for gesture types that represent numbers (i.e., "5" and "8", with no significant differences between them), which could be attributed to the fact that numbers can be easily represented using fingers. Actually, no significant differences were found between "5" and "8", as well as between them and other gesture types that can also be easily mapped into fingers (such as "circle", "square", "triangle", "V", "H"). Furthermore, finger gestures did not obtain high values for all gesture types (excluding one-handed type). For instance, some gestures were easier to map into two hands (e.g., "square", "T" and "step-down") and arms (e.g., "horizontal line", "V" and "X"). In addition, though whole-body type is represented in a relatively negligible ratio (5.8%), ratios between 14% and 20% were obtained for a few gestures (e.g., "heart" and "zig-zag"). Finally, gestures executed with lower limbs were observed only for six gestures with rates lower than 6%. Summing up, our participants produced their gestures mainly using one hand, and then, with fingers and two hands.

Although types of gestures produced with more than one upper limb did not get the best ratios, we performed an additional analysis of them given that about one third of gestures fell in these physicality types. 428 gestures performed with two-arms, two-hands, and fingers of both hands were analyzed according to the spatial relation of them. This relation can be as follows: folded (limbs act as a unit), act on each other (limbs act upon each other in a dynamic contact), symmetrical, complementary (e.g., one hand acts as a reference while the other one is moved), and independent. Effectors are in touch in the first two cases, whereas they are separated in the other cases. Figure 6 shows the global ratios of the five levels for each kind of effectors and overall. A Chi-square test revealed that the percentage of the spatial relation significantly differed by used effectors ($\chi^2(8) = 175.43$, $p < .001$). Actually, Figure 6 shows participants used most often two arms (67.4%) or several fingers (87.9%) as a unit, whereas

they preferred employing both hands symmetrically (55.6%). The other levels are negligible (less than 14% in all cases).

Table 2. Friedman tests for gesture type on level types.

Physicality		Movement		Structure	
Type	$\chi^2(19)$	Type	$\chi^2(19)$	Type	$\chi^2(19)$
Whole Body	91.40	Single	128.02	Static body pose	75.78
Arms	125.84	Parallel	144.58	Static arm pose	138.53
One-handed	122.12	Sequential	67.82	Static hand pose	91.21
Two-handed	46.92			Static hand pose and path	102.08
Fingers	85.70			One finger pose	121.77
Lower limbs	39.01			Other	62.92
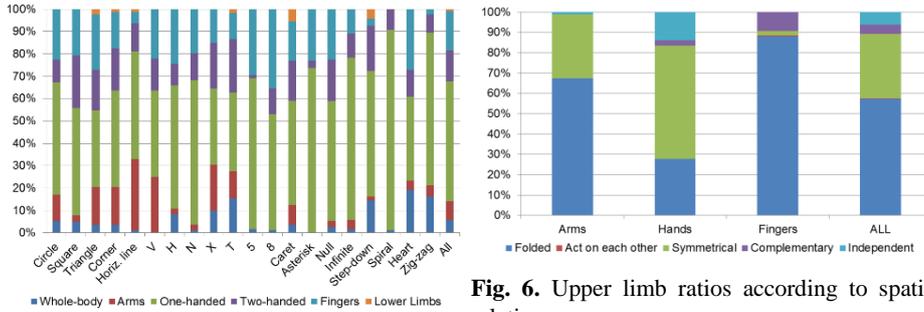
Fig. 5. Gesture physicality ratio.

Fig. 6. Upper limb ratios according to spatial relation.
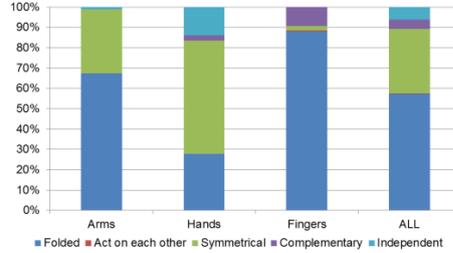
5.3 Movement Breakdown

Figure 7 shows gesture ratios for each gesture type according to movement synchronization type. The three types of sequential gestures were subsumed under a general sequential type due to the small number of occurrences especially in the case of sequence of parallel movements and sequence of parallel and single movements. A Friedman test revealed significant difference in the ratios (averaged over all gesture types) between the three movement synchronization types ($\chi^2(2) = 30.70$, $p < .001$). Post-hoc tests confirmed significant differences between all the movement synchronization types.

Similarly, gesture types showed a significant effect on the ratio of these three synchronization types as reported on Table 2. Overall participants performed more often single gestures; its average (60.5%) almost doubles the one of parallel gestures. However, parallel gestures were preferred to produce gesture types having a symmetry axis (e.g., "X", "triangle", and "T", etc.), a finding in agreement with previous work [32, 34], only symmetric gestures can be conveniently parallelized during articulation. Post-hoc tests showed significant differences between these three gesture types ("X", "triangle", and "T") and the remaining ones, except for the following pairs: ("trian-

gle", "rectangle"), ("T", "rectangle"), ("T", "V"), and ("T", "8"). All these gesture types can be produced using two end-effectors with ease, e.g., crossing the arms/hands/fingers to form an "X" (Figure 4 b-d). Other gesture types, such as "8" and "square", got in parallel type ratios lower than in single type, but they demonstrated a behavior similar to the previously described for "X", "triangle", and "T". Concerning sequential type, its overall ratio was very small (6.6%), but non-negligible ratios were observed for three gesture types: "H", "null", and "T". These gesture types showed no differences between them, as well as significant differences between them and the other gesture types (except between "*" and "null", and "*" and "T"). This finding suggests that participants may have produced gestures composed of a sequence of movements only when it was worth doing it; they still preferred the other two types.
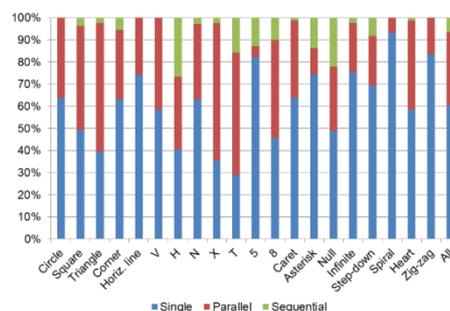

Fig. 7. Gesture movement composition ratio.

5.4 Structure Breakdown

This section provides details of gesture articulations based on both the dynamics (i.e., poses and paths) of gestures and the used end-effectors. The five structure types with highest scores were selected, while the remaining levels were grouped into another one labeled as "other". Figure 8 shows the corresponding ratios for each gesture type. Significant difference in the ratios between these six structure types was found ($\chi^2(5) = 74.11$, $p < .001$). Post-hoc tests confirmed not significant differences between static body pose and the levels one finger pose and "other". Likewise, Friedman tests revealed a significant effect of gesture type on the ratio of structure levels (see Table 2). In particular, participants preferred gestures made with one hand pose plus path (62.3%) to produce all gesture types, especially for those ones that may be considered more difficult or strange (e.g., "spiral", "infinite" and "zig-zag", which showed no difference from one another). Although static hand pose had a relatively non-negligible average ratio of 19.5%, it got the second place in most cases. It was outperformed by static arm pose to produce two gesture types, "horizontal line" and "X" (there was no difference between them), that could be mapped better into arms as explained above. Also, static poses held with the whole body or with one (index) finger of both hands were rarely used. Participants employed them especially for articulating the gesture types "X" and "T" (with significant differences between them

and the other gesture types, but no difference between both), which were in fact the gesture types more fairly distributed among the five structure types (excluding "other"). "Other" level was observable basically for "step-down" and "asterisk".

Figure 9 shows global ratios for each effector type according to structure levels. It reveals that participants performed few dynamic pose gestures (about 2%), and they mostly produced gestures that followed a path (63%) in comparison to only hold poses (37%); i.e., participants most often executed their gestures by holding a single (hand) pose while drawing the corresponding gesture type.
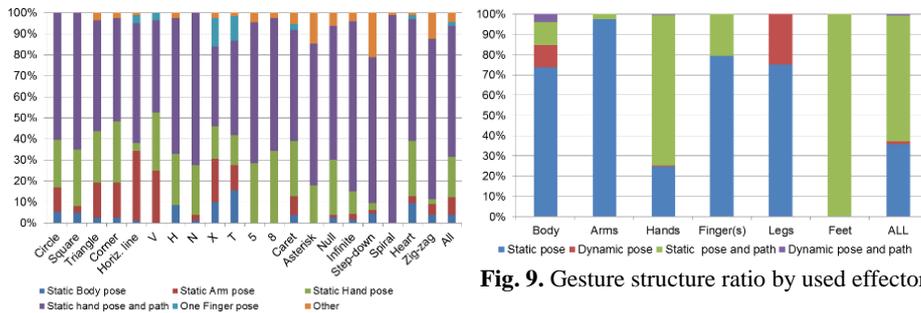

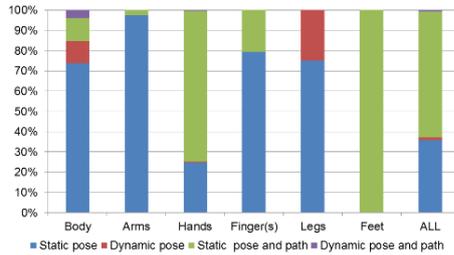**Fig. 8.** Gesture structure ratio.

Fig. 9. Gesture structure ratio by used effectors.

5.5 Mental Model Observations

At a high level, GCP facilitates the analysis and understanding of the second task. A user doing the task identifies the current symbol and defines a gesture for it (see Figure 2). Next, s/he executes that gesture departing from a resting position, and returning to it or going to another one. Overall we observed that participants executed their gestures in this manner, both gestures composed of single movement as well as those ones composed of multiple movements. Actually, only one subject tried to perform the gestures consecutively in most cases, i.e., without having a retraction. Contrary to a general/common participants' behavior, there are several particular observations that are worth mentioning:

– **Gesture shape complexity influences mid-air gesture input.** Overall participants were able to produce various articulations for the predefined gesture set, but they felt less creative for gesture types with complex/strange geometry (e.g., "spiral"). Two participants proposed exclusively drawing gestures in all cases. Nonetheless, they still used both hands and/or sequential movements to "draw" their symbolic gestures.
– **Preference for vertical plane.** While we did not constrain participants on the direction of the plane when articulating a gesture, all participants performed their gestures in the vertical plane. As an exception, one participant executed some gestures in the horizontal plane.
– **Gesture position, size and direction can be a source of variation.** Several participants in some cases changed the used hand, starting point, size and/or direction of paths to produce various articulations.

‒ **Gestures in ways few technologies could detect.** Though we instructed participants to articulate a gesture for each gesture type, some of them gestured in ways current technologies could not detect. One participant counted the number of fingers to define a number. Another participant touched his heart to define the "heart". One user walked by making three steps to define the "zig-zag". Furthermore, a participant proposed a few gestures by drawing a part of the shape, and next, performing another movement to deform it and get the desired figure. For example, he drew a horizontal line with a hand, and then, he put his hands apart on the line and moved them down to form the step-down symbol.

6   Discussion: Comparison with Elicitation Studies

The results described above provide evidence on how users articulate gestures, which should be compared with previous results. Table 3 shows a comparison between several previously proposed taxonomies and the one proposed in this work. Those taxonomies were proposed as part of studies on user-defined gestures for scenarios such as augmented reality [31], humanoid robot [28], storytelling [17], controlling a drone [29], and video games [37]. The values included in the table are given as the percentage of gestures that participants proposed for each gesture type in each work. Some levels were removed because there were not equivalent levels in one or other taxonomy (e.g., see physicality level).

Table 3. Comparison with other taxonomies.

		This work	[28]	[17]	[31]	[29]	[37]
Physicality	Whole Body	5.8	34.8	41.0		13.0	3.0
	One-handed	53.4	40.9	42.0		45.0	40.0
	Two-handed	13.9	24.2	17.0		42.0	35.0
	One-leg	0.8					16.0
	Two-legs						4.0
	Other						2.0
Movement	Single	60.5					
	Parallel	33.0			12.9		
	Sequential	6.6			4.4		
Structure	Static pose	35.7	15.7	45.0	10.1	11.0	
	Dynamic pose	1.3			10.8		
	Static pose with path	62.3	84.3	55.0	71.0	89.0	
	Dynamic pose with path	0.6			8.1		

Referring to physicality level, (one) hand was the most used end-effector in both our study and the other four studies that considered various body parts (i.e., [17, 28, 29, 37]). Though Piumsomboon et al. [31] only reported hand gestures, they also found a preference for one-handed gestures. Similarly, users may also prefer using one hand for gesturing in scenarios like product exhibition or public displays [21, 1, 5]. Actually, passers-by who stop to interact with a public display could hold a mobile phone or carry things on one hand [21, 1]. On the contrary, there is evidence in favor of two-handed gestures [26]. However, users may prefer bimanual gestures depending

on the nature of the tasks or the performed gestures, e.g., when one hand is employed as a reference or for zooming [2, 5]. Likewise, we found that participants employed more parallel movements than sequential movements. This finding is comparable with the results reported by Piumsomboon et al. [31] for the symmetric and asymmetric levels respectively. Concerning structure level, four of the works (i.e., [31, 28, 17, 29]) shown in Table 3 reported that participants articulated gestures using more strokes than holds, which is consistent with our results. Likewise, we found a high preference toward static hand pose with path gestures similar to [31].

In addition, although our study had other focus, we might also do another comparison with gesture elicitation studies, namely referring to production of gestures to enhance this type of studies. Morris et al. [24] suggested using five gestures, whereas Ho et al. [14] advised that requesting participants to produce more than three gestures would impact on practical utility of gesture elicitation. Though our participants were shown symbol names instead of desired effects of actions, they proposed at least three "natural" gestures for 70% (SD = 30%) of the utilized symbols on the average. Our study also included a first task in which the participants performed several gestures in a free manner, which could be comparable to priming. However, our participants were not able to reach the threshold of three gestures in all cases despite of this "priming". In conclusion, this finding suggests that proposing more than three gesture articulations can be not natural, which is consistent with [14] but different than [24].

Finally, our participants' characteristics are similar to those who took part in the studies used in Table 3 (i.e., [31, 28, 17, 29, 37]), but a difference between our work and the others must be noticed. While those works were mainly focused on finding gestures for specific scenarios (especially based on Wobbrock et al.'s methodology [45]), we are interested in analyzing the various manners in which users could produce gestures for a same command. In fact, participants were asked to propose a set of gestures for the needed actions in those works, whereas we asked participants to propose various gestures for a set of symbols (Task 2).

7   Design implications

Informed by our findings, we are able to outline a set of guidelines for designing mid-air gestures interfaces that address gesture ergonomics, design and recognizers with the aim of enabling several gestures to one command.

7.1   Mid-Air Gesture Ergonomics

Our findings indicate that strokes are preferred instead of poses to articulate gestures. These strokes were especially expressed by following a path, mostly with hands. Our findings also demonstrate that producing paths with hands matters to users more than the posture maintained while they do it. The participants generally kept the same pose while executing a gesture; i.e., they rarely used more than one posture between different strokes. Specific hand postures would be needed if users should discriminate between drawing paths and displacing the hand to the point where the path (or a part of

it) starts (i.e., the system is not capable of doing this distinction automatically), for example, to perform multi-stroke gestures. Likewise, and contrary to the findings for multi-touch gestures [32], paths performed with a variable number of fingers should not be a problem because users would adopt a single pose (e.g., putting together middle and index fingers, or touching the tips of thumb and index fingers).

Despite the high preference for drawing gestures reported here, we do not advocate that gestures based on poses should not be employed. They have proven to be useful in various scenarios (e.g., finger spelling [39]), but an additional issue is that postures should be learned to interact with applications [15]. Beyond this possible limitation, our results show that static poses maintained not only with hands/fingers may be suitable just for gesture types that could be easily produced through this structure (e.g., letter "X" or "T" and numbers). This finding indicates that the use of postures would depend on the facility to map gestures into the corresponding end-effectors. On the other hand, our results suggest that users would prefer static poses instead of dynamic poses, given that the second ones were very scarce in the second task.

Gestures expressed using either single strokes or single holds should be preferred according to the results of the second task. Although parallel movements may be used as a complement, sequential movements may not result "natural" to users. Unexpectedly, when participants produced the candidate multi-stroke gestures (i.e., gestures for symbols "H", "X", "T", "asterisk", and "null") using static hand poses with paths, they did it frequently using single strokes. Furthermore, participants did not worry or notice a need to discriminate between drawing paths and just moving hands.

Concerning used end-effectors, participants clearly preferred employing hands and fingers to produce symbolic gestures. Despite the high tendency to use one hand, the presence of two-handed, two-arms, and multi-finger gestures was also noticeable. Our findings suggest that users would use both hands symmetrically (principally to produce static hand pose with path gestures) more than folded hands. Conversely, two arms and several fingers would be mainly used in touch acting as a unit. Moreover, unlike previous work [2, 32], we observed no two-handed gestures in which a hand was used as a reference and the other one was used to express the gesture. Additionally, overall the use of arms or whole-body to execute gestures would be preferred depending on the ease to map gestures into them as mentioned above. Finally, more gestures executed with feet may have been expected, but they were hardly ever used by the participants. This may have happened due to any of these causes: the instructions were insufficient, the participants did not imagine these gestures, or simply, foot gestures were not good enough or "natural" to participants.

7.2   Mid-Air Gesture Design

Our findings show that inferring flexible input when articulating gestures would be more suitable when users are provided with little to no instructions or when symbols are unfamiliar or difficult to them. Otherwise, UI designers and researchers should observe how users articulate a gesture set before designing it. Familiar shapes should also be preferred to unfamiliar ones, and gesture articulations should be connected to users' previous gesture practice whenever possible. Additionally, gesture shapes with

complex geometries should be used with care, and learning and memorization should be integrated into the design of such gesture shapes. Moreover, the available methods to analyze the difficulty of symbolic gestures (e.g., [42, 10, 34]) should be taken into account during the design.

7.3 Mid-Air Gesture Recognizers

Many of the gestures we witnessed had strong implications for gesture recognition technology. Our results demonstrate that UI designers and researchers should design flexible recognizers that are invariant to users' preferred articulation patterns. Gesture recognizers should be trained with different articulation patterns in terms of physicality, synchronization, and structure. For example, for the same gesture type, our participants articulated it using different number of strokes that are combined sequentially or in parallel using arms/hands/fingers etc., and mixing path and pose structures, such as [19, 3, 41, 33].

8 Conclusion and next steps

We presented an investigation of users' gestures articulations variability. We outlined a model for gestures conception and production (GCP), and a taxonomy for mid-air gestures to carry out a qualitative and quantitative analysis. Our findings indicate that, additionally to hands, users would use other body parts to articulate gestures if the proposed gesture type can be mapped better into other body parts to hold postures. Similar to multi-touch input [32], gestures in mid-air could be articulated with single as well as multiple movements entered in sequence or in parallel. Our findings also suggest that users would prefer producing gestures in mid-air mainly using one hand to iconically describe single motion paths. This preference does not mean that users would not produce gestures in other articulations. These findings are important in the context of proposing new interaction techniques that make use of the variability of user gestures, and hence, this study is a first step toward enabling designers to use more than one gesture for a same command. These many-to-one mappings should lead to better user interaction experiences by giving more flexibility and avoiding penalization.

As a future work, we plan to study gesture variability in more interactive scenarios. The same or a similar methodology used here may be followed, but users would have to propose gestures for concrete applications. It may be similar to elicitation studies but adding variability analysis. Other aspects may be also considered in this future work, such as cultural differences (with a larger population) and different contexts.

References

1. Ackad, C., Clayphan, A., Tomitsch, M., Kay, J.: An in-the-wild study of learning mid-air gestures to browse hierarchical information at a large interactive public display. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 1227-1238. UbiComp'15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2750858.2807532

2. Aigner, R., Wigdor, D., Benko, H., Haller, M., Lindbauer, D., Ion, A., Zhao, S., Koh, J.T.K.V.: Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for hci. Tech. rep., https://www.microsoft.com/enus/research/publication/understanding-mid-air-hand-gestures-a-study-ofhuman-preferences-in-usage-of-gesture-types-for-hci/ (November 2012)

3. Anthony, L., Wobbrock, J.O.: A lightweight multistroke recognizer for user interface prototypes. In: Proceedings of Graphics Interface 2010. pp. 245-252. GI'10, Canadian Information Processing Society, Toronto, Ont., Canada, Canada (2010), http://dl.acm.org/citation.cfm?id=1839214.1839258

4. Barkhuus, L.: Television on the internet: New practices, new viewers. In: CHI'09 Extended Abstracts on Human Factors in Computing Systems. pp. 2479-2488. CHI EA'09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1520340.1520351

5. Chen, L.C., Chu, P.Y., Cheng, Y.M.: Exploring the Ergonomic Issues of User-Defined Mid-Air Gestures for Interactive Product Exhibition, pp. 180-190. Springer International Publishing, Cham (2016), http://dx.doi.org/10.1007/978-3-319-39862-4_17

6. Choi, E., Kim, H., Chung, M.K.: A taxonomy and notation method for three-dimensional hand gestures. International Journal of Industrial Ergonomics 44(1), 171-188 (2014)

7. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of oz studies: Why and how. In: Proceedings of the 1st International Conference on Intelligent User Interfaces. pp. 193-200. IUI'93, ACM, New York, NY, USA (1993), http://doi.acm.org/10.1145/169891.169968

8. Dezfuli, N., Khalilbeigi, M., Mühlhäuser, M., Geerts, D.: A study on interpersonal relationships for social interactive television. In: Procedings of the 9th International Interactive Conference on Interactive Television. pp. 21-24. EuroITV'11, ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/2000119.2000123

9. Erazo, O., Pino, J.A.: Predicting task execution time on natural user interfaces based on touchless hand gestures. In: Proceedings of the 20th International Conference on Intelligent User Interfaces. pp. 97-109. IUI '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2678025.2701394

10. Erazo, O., Pino, J.A., Antunes, P.: Estimating production time of touchless hand drawing gestures. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (eds.) Human-Computer Interaction – INTERACT 2015: 15th IFIP TC 13 International Conference, Bamberg, Germany, September 14-18, 2015, Proceedings, Part III. pp. 552-569. Springer International Publishing, Cham (2015), http://dx.doi.org/10.1007/978-3-319-22698-9_38

11. Gallo, L., Placitelli, A.P., Ciampi, M.: Controller-free exploration of medical image data: Experiencing the kinect. In: Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on. pp. 1-6 (June 2011)

12. Henschke, M., Gedeon, T., Jones, R.: Touchless gestural interaction with wizard-of-oz: Analysing user behaviour. In: Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction. pp. 207-211. OzCHI'15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2838739.2838792

13. Hespanhol, L., Tomitsch, M., Grace, K., Collins, A., Kay, J.: Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In: Proceedings of the 2012 International Symposium on Pervasive Displays. pp. 6:1-6:6. PerDis'12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2307798.2307804

14. Ho, L., Hornecker, E., Bertel, S.: Modifying gesture elicitation: Do kinaesthetic priming and increased production reduce legacy bias? In: Proceedings of the TEI '16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction. pp. 86-91. TEI'16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2839462.2839472

15. Ismair, S., Wagner, J., Selker, T., Butz, A.: Mime: Teaching mid-air pose-command mappings. In: Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services. pp. 199-206. MobileHCI'15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2785830.2785854

16. Kendon, A.: Gesture: Visible Action as Utterance. Cambridge University Press (2004)

17. Kistler, F., André, E.: User-defined body gestures for an interactive storytelling scenario. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) Human-Computer Interaction – INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II. pp. 264-281. Springer Berlin Heidelberg, Berlin, Heidelberg (2013), http://dx.doi.org/10.1007/978-3-642-40480-1_17

18. Kita, S., van Gijn, I., van der Hulst, H.: Movement phases in signs and co-speech gestures, and their transcription by human coders. In:Wachsmuth, I., Fr• ohlich, M. (eds.) Gesture and Sign Language in Human-Computer Interaction: International Gesture Workshop Bielefeld, Germany, September 17{19, 1997 Proceedings. pp. 23-35. Springer Berlin Heidelberg, Berlin, Heidelberg (1998), http://dx.doi.org/10.1007/BFb0052986

19. Kratz, S., Rohs, M.: Protractor3d: A closed-form solution to rotation-invariant 3d gestures. In: Proceedings of the 16th International Conference on Intelligent User Interfaces. pp. 371-374. IUI'11, ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/1943403.1943468

20. Leyvand, T., Meekhof, C.,Wei, Y.C., Sun, J., Guo, B.: Kinect identity: Technology and experience. Computer 44(4), 94-96 (2011)

21. Mäkelä, V., Korhonen, H., Ojala, J., Järvi, A., Väänänen, K., Raisamo, R., Turunen, M.: Investigating mid-air gestures and handhelds in motion tracked environments. In: Proceedings of the 5th ACM International Symposium on Pervasive Displays. pp. 45-51. PerDis'16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2914920.2915015

22. McNeill, D.: Hand and Mind: What Gestures Reveal about Thought. University of Chicago (1992)

23. Mehler, A., Lücking, A., Abrami, G.: Wikinect: image schemata as a basis of gestural writing for kinetic museum wikis. Universal Access in the Information Society pp. 333-349 (2015)

24. Morris, M.R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., Schraefel, M.C., Wobbrock, J.O.: Reducing legacy bias in gesture elicitation studies. Interactions 21(3), 40-45 (May 2014), http://doi.acm.org/10.1145/2591689

25. Nacenta, M.A., Kamber, Y., Qiang, Y., Kristensson, P.O.: Memorability of predesigned and user-defined gesture sets. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1099-1108. CHI'13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2470654.2466142

26. Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., Mackay, W.: Mid-air pan-and-zoom on wall-sized displays. In: Proceedings of the SIGCHI Conference on Human Factors in

Computing Systems. pp. 177-186. CHI'11, ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/1978942.1978969

27. Nielsen, M., Störring, M., Moeslund, T.B., Granum, E.: A procedure for developing intuitive and ergonomic gesture interfaces for hci. In: Camurri, A., Volpe, G. (eds.) Gesture-Based Communication in Human-Computer Interaction: 5th International Gesture Workshop, GW 2003, Genova, Italy. pp. 409-420. Springer Berlin Heidelberg, Berlin, Heidelberg (2004), http://dx.doi.org/10.1007/978-3-540-24598-8_38

28. Obaid, M., Häring, M., Kistler, F., Bühling, R., André, E.: User-defined body gestures for navigational control of a humanoid robot. In: Ge, S.S., Khatib, O., Cabibihan, J.J., Simmons, R., Williams, M.A. (eds.) Social Robotics: 4th International Conference, ICSR 2012, Chengdu, China, October 29-31, 2012, pp. 367-377. Springer Berlin Heidelberg, Berlin, Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-34103-8_37

29. Obaid, M., Kistler, F., Kasparavičiūtė, G., Yantaç, A.E., Fjeld, M.: How would you gesture navigate a drone?: A user-centered approach to control a drone. In: Proceedings of the 20th International Academic Mindtrek Conference. pp. 113-121. AcademicMindtrek'16, ACM, New York, NY, USA (2016), http://doi.acm.org/10.1145/2994310.2994348

30. Oh, U., Findlater, L.: The challenges and potential of end-user gesture customization. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1129-1138. CHI'13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2470654.2466145

31. Piumsomboon, T., Clark, A., Billinghurst, M., Cockburn, A.: User-defined gestures for augmented reality. In: Kotzé, P., Marsden, G., Lindgaard, G.,Wesson, J., Winckler, M. (eds.) Human-Computer Interaction – INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II. pp. 282{299. Springer Berlin Heidelberg, Berlin, Heidelberg (2013), http://dx.doi.org/10.1007/978-3-642-40480-1_18

32. Rekik, Y., Grisoni, L., Roussel, N.: Towards many gestures to one command: A user study for tabletops. In: Human-Computer Interaction – INTERACT 2013: 14th IFIP TC 13 International Conference, Cape Town, South Africa, September 2-6, 2013, Proceedings, Part II. pp. 246-263. Springer Berlin Heidelberg, Berlin, Heidelberg (2013), http://dx.doi.org/10.1007/978-3-642-40480-1_16

33. Rekik, Y., Vatavu, R.D., Grisoni, L.: Match-up & conquer: A two-step technique for recognizing unconstrained bimanual and multi-fnger touch input. In: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces. pp. 201-208. AVI'14, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2598153.2598167

34. Rekik, Y., Vatavu, R.D., Grisoni, L.: Understanding users' perceived difficulty of multi-touch gesture articulation. In: Proceedings of the 16th International Conference on Multimodal Interaction. pp. 232-239. ICMI'14, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2663204.2663273

35. Rico, J., Brewster, S.: Usable gestures for mobile interfaces: Evaluating social acceptability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 887-896. CHI'10, ACM, New York, NY, USA (2010), http://doi.acm.org/10.1145/1753326.1753458

36. Sheu, J.S., Huang, Y.L.: Implementation of an interactive tv interface via gesture and handwritten numeral recognition. Multimedia Tools and Applications pp. 1-22 (2015), http://dx.doi.org/10.1007/s11042-015-2739-6

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

38. Spano, L.D., Cisternino, A., Paterno, F.: A compositional model for gesture definition. In: Winckler, M., Forbrig, P., Bernhaupt, R. (eds.) Human-Centered Software Engineering: 4th International Conference, HCSE 2012, Toulouse, France, October 29-31, 2012. Proceedings. pp. 34-52. Springer Berlin Heidelberg, Berlin, Heidelberg (2012), http://dx.doi.org/10.1007/978-3-642-34347-6_3

39. Sridhar, S., Feit, A.M., Theobalt, C., Oulasvirta, A.: Investigating the dexterity of multi-finger input for mid-air text entry. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. pp. 3643-3652. CHI'15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2702123.2702136

40. Vatavu, R.D.: User-defined gestures for free-hand tv control. In: Proceedings of the 10th European Conference on Interactive Tv and Video. pp. 45-48. EuroiTV'12, ACM, New York, NY, USA (2012), http://doi.acm.org/10.1145/2325616.2325626

41. Vatavu, R.D., Casiez, G., Grisoni, L.: Small, medium, or large?: Estimating the user-perceived scale of stroke gestures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 277-280. CHI'13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2470654.2470692

42. Vatavu, R.D., Vogel, D., Casiez, G., Grisoni, L.: Estimating the perceived difficulty of pen gestures. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) Human-Computer Interaction – INTERACT 2011: 13th IFIP TC 13 International Conference, Lisbon, Portugal, September 5-9, 2011, Proceedings, Part II. pp. 89-106. Springer Berlin Heidelberg, Berlin, Heidelberg (2011), http://dx.doi.org/10.1007/978-3-642-23771-3_9

43. Vikram, S., Li, L., Russell, S.: Writing and sketching in the air, recognizing and controlling on the fly. In: CHI'13 Extended Abstracts on Human Factors in Computing Systems. pp. 1179-1184. CHI EA '13, ACM, New York, NY, USA (2013), http://doi.acm.org/10.1145/2468356.2468567

44. Walter, R., Bailly, G., Valkanova, N., M• uller, J.: Cuenesics: Using mid-air gestures to select items on interactive public displays. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices &; Services. pp. 299-308. MobileHCI'14, ACM, New York, NY, USA (2014), http://doi.acm.org/10.1145/2628363.2628368

45. Wobbrock, J.O., Morris, M.R., Wilson, A. D.: User-defined gestures for surface computing. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1083-1092. CHI '09, ACM, New York, NY, USA (2009), http://doi.acm.org/10.1145/1518701.1518866

46. Wong, A., Haith, A., Krakauer, J.: Motor planning. Neuroscientist 21(4), 385-398 (8 2015)