Pers Ubiquit Comput (2015) 19:821–838
DOI 10.1007/s00779-015-0863-y

ORIGINAL ARTICLE

On free-hand TV control: experimental results on user-elicited gestures with Leap Motion

Ionuţ-Alexandru Zaiţi[1] · Ştefan-Gheorghe Pentiuc[1] · Radu-Daniel Vatavu[1]

Received: 29 November 2014 / Accepted: 6 May 2015 / Published online: 24 June 2015

Abstract We present insights from a gesture elicitation study conducted for TV control, during which 18 participants contributed gesture commands and rated the execution difficulty and recall likeliness of free-hand gestures for 21 television control tasks. Our study complements previous work on gesture interaction design for the TV set with the first exploration of *fine-grained resolution 3-D finger movements and hand gestures*. We report lower agreement rates than previous gesture studies ($\mathcal{AR} = .158$) with 72.8 % recall rate and 15.8 % false positives, results that are explained by the complexity and variability of unconstrained finger and hand gestures. However, our observations also confirm previous findings, such as people preferring related gestures for dichotomous tasks and more disagreement occurring for abstract tasks, such as "open browser" or "show the list of channels" for our specific TV scenario. To reach a better understanding of our participants' preferences for articulating finger and hand gestures, we defined five measures for Leap Motion gestures, such as *gesture volume* and *finger-to-palm distance*, which we employed to evaluate gestures performed by our participants. We also contribute a set of guidelines for practitioners interested in designing free-hand gestures for interactive TV scenarios involving similar gesture acquisition technology. We release our dataset consisting in 378 Leap Motion gestures described by fingertips position, direction, and velocity coordinates to foster further studies in the community. This first exploration of viewers' preferences for fine-grained resolution free-hand gestures for TV control represents one more step toward designing *low-effort gesture interfaces for lean-back interaction* with the TV set.

Keywords Gesture interfaces · Interactive TV · Leap Motion · Gestures · Elicitation study · Recall likeliness · Gesture analysis · Gesture measures · Agreement rate · Agreement analysis · Home entertainment · User-defined gestures · Free-hand gestures

✉ Radu-Daniel Vatavu
vatavu@eed.usv.ro

Ionuţ-Alexandru Zaiţi
ionutzaiti@gmail.com

Ştefan-Gheorghe Pentiuc
pentiuc@eed.usv.ro

[1] Department of Computer Science, Electronics, and Automation and Integrated Center for Research, Development, and Innovation in Advanced Materials, Nanotechnologies, and Distributed Systems for Fabrication and Control (MANSiD), University Stefan cel Mare of Suceava, 720229 Suceava, Romania

1 Introduction

Television represents a valuable component in our lives, not only for delivering information and entertainment [14, 15], but also for creating the premises for enriched social interaction [2, 3, 12]. Television technology, content type, and content accessibility have evolved considerably over recent years. For example, we are currently witnessing inhabited, interactive, and Internet television systems [2, 4, 15, 36, 38] that are augmented by audio surround systems, ambient effects [39, 52], and second-screen devices [7]. However, the control device to interact with the TV set has remained virtually unchanged as we still control television today with the standard remote control. However, in the context in which researchers understand television and work with television concepts converging to interactivity [6], better designs of input devices and new interaction techniques are required for TV control.

We are interested in this work in understanding people's preferences for interacting with television via free-hand gestures. Although this topic has already received significant attention from both research and industry [5, 11, 16, 17, 21, 41, 47, 49, 50, 60], we address a different angle and technology for *low-effort gestures*. This work represents an extension of the previous investigation of Vatavu and Zaiţi [53] on free-hand gestures elicited with the Leap Motion controller [27]; see Fig. 1. In this work, we reach a better understanding of viewers' preferences for *short-range hand pose and 3-D finger movements for TV control* by analyzing agreement results with a new methodology [55] and by evaluating gesture performance with new measures. This way, we provide more insights on people's preferences and use of fine-grained finger gestures for television user interfaces that complements existing research on designing general gesture interaction for TV control [5, 11, 17, 21, 49, 60].

Our contributions are as follows: (1) we collect people's gesture preferences for interacting with the TV set using 3-D finger movements and hand poses that we acquire with the Leap Motion controller (we refer to such gestures as *leap gestures* in this work); (2) we report agreement results [55] for our participants' gesture proposals, and we select representative gestures for controlling 21 functions of the TV set, such as go to previous and next channel, but more abstract tasks as well, such as show the list of channels and open the Internet browser; (3) we introduce a set of five gesture measures to characterize the spatial and kinematic properties of leap gestures, such as gesture volume, gesture length, and finger-to-palm distance; (4) we contribute a set of design guidelines for gesture interfaces for the TV set; (5) we release our Leap Motion gesture dataset composed of 378 distinct samples collected from 18 participants together with companion software that reads the dataset and computes our measures. At the moment of writing this article, our dataset is the only publicly available dataset of Leap Motion gestures and one of the very few gesture datasets for research on interactive TV. Therefore, we hope that these resources will be useful to the community to foster new developments for free-hand gestures for TV control. In the long run, our exploration and contributions presented in this work are first steps toward designing *low-effort free-hand gestures for lean-back control* of the interactive TV set.

Fig. 1 Experimental setup for eliciting Leap Motion gestures for free-hand TV control [53]: participants perform free-hand gestures in the small space above the Leap Motion controller while watching content displayed on the TV screen

2 Related work

We position our research investigation in the large body of work on designing gestural interfaces for the interactive TV. Gesture interaction design for TV has been a very attractive topic in both academia [5, 11, 13, 19, 21, 25, 41, 49, 60] and industry [29, 40, 44] and has generated a large body of work to understand people's preferences for gesture interaction [22, 24, 28, 33–35, 43, 55, 58]. In this context, we want to understand the use of *fine-grained finger movements and hand poses* for executing functions on the TV set. Such gesture types have not been explored yet for the interactive TV context, probably because of the lack of accessible technology to capture such gestures. However, the dexterity and multi-functionality of the human hand has been thoroughly studied in psychology [20], which justifies a focused study of fine-grained gestures for the interactive TV. We believe that such gesture types are more suited than large whole-body movements [49, 51, 60] in the context of *lean-back* versus *lean-forward* paradigms of interacting with TV.

The experimental study reported in this work connects to previous gesture interface designs for the interactive TV. For example, Freeman and Weissman [17] proposed the first TV gesture interface that mapped the viewer's hand movements to a cursor displayed on the TV screen; Bobeth et al. [5] investigated the way older adults employ free-hand gestures to execute tasks on TV; Vatavu [50, 52] introduced augmented TV spaces for the control of which pointing gestures were employed; Dias et al. [13] were interested in designing gesture interfaces for specific applications running on TV; and Vatavu [56] examined audience kinesics in the form of body silhouettes captured by depth sensors and shared synchronously among remote TV viewers in order to enhance the user experience of social television. The PalmRC prototype of Dezfuli et al. [11] also approaches our rationale for exploring low-effort short-range gestures for TV. PalmRC leverages the palm of the user's hand as a supporting surface for finger touches to enable eyes-free control of the TV set.

In this work, we conduct a gesture elicitation study that implements the guessability methodology of Wobbrock et al. [59]. This methodology has been successfully applied for eliciting gesture commands for various application domains [24, 25, 34, 35, 43, 58] including gestures for the interactive TV [49, 51, 60]. For example, Vatavu [49] collected and analyzed free-hand and whole-body gestures captured with the Microsoft Kinect sensor in what constituted the first gesture elicitation study for the interactive TV. Vatavu [49] reported an average agreement rate of .415 for 12 standard functions on TV, which was computed from gestures collected from 12 participants. A follow-up exploration [51] extended the initial findings with more discussion and analysis of people's preferences for free-hand gestures versus TV remote controls. The study examined 22 functions for multi-screen entertainment systems, for which gesture commands were collected from 20 participants using the Microsoft Kinect sensor and an augmented remote control. Agreement rate analysis showed an average .430 agreement for remote gestures and .330 for whole-body gestures [51]. In another work, Vatavu [52] explored people's preferences for remote gestures to control television content spanning a physical-digital space in the implementation of the AROUND-TV prototype for augmented television. Wu and Wang [60] also collected user-defined hand and body gestures for TV control. In this work, we conduct a gesture elicitation study that addresses free-hand gestures captured using the Leap Motion controller, gestures that we analyze using the new agreement rate formula and methodology of Vatavu and Wobbrock [55].

Beyond this body of work on gesture elicitation and designing gesture interfaces for the TV, we believe there is much opportunity in leveraging the fine-grained movements of fingers for *low-effort lean-back interaction* with the TV set. To this end, we focus in this work on short-range finger movements and hand poses performed in a small volume of space. We capture such gestures with the Leap Motion controller [27], an acquisition device that provides a large amount of fingertip data, i.e., position and velocity coordinates at 200 frames per second for up to 10 fingers with precision of 0.01 mm. Such features attracted the attention of the research community toward this device due to the many opportunities for designing interactive applications; see [1, 8, 9, 18, 30, 32, 45]. The Leap Motion controller can also been used in combination with other gesture capture technologies, see [32, 45], to increase the acquisition range of users' gesture input. In this work, we employ the Leap Motion controller to collect short-range finger and hand gestures, a category of gestures that has been left unexplored so far for the interactive TV. We believe that such gesture types are likely to represent an optimal choice for lean-back TV control in the line of simple eyes-free alternatives to the TV remote control [11] and toward simple and efficient interaction with multi-screen TV systems [50, 52, 54, 57].

3 Experiment

We conducted a gesture elicitation experiment [10, 55, 58, 59] to collect people's preferences for Leap Motion gesture commands in the context of the interactive TV.

3.1 Participants

Eighteen (18) volunteers (four females) participated in our study. Mean age was 25.0 years (SD = 3.1). All participants were right handed. Ten participants had no previous experience with 3-D gesture interfaces and the other eight had used occasionally Nintendo Wii or Microsoft Kinect controllers for video games.

3.2 Apparatus

A 40-in. (102 cm) Sony TV and a Leap Motion controller were connected to a computer running Microsoft Windows 8.1 and our custom gesture acquisition software that implemented the experiment design. The Leap Motion controller [27] is a 3-D tracking device that detects and tracks targets with a precision up to 0.01 mm in a 3-D space of 0.23 m^3 with a field of view of 150 degrees. The Leap Motion controller reports data in the form of position, direction, and velocity coordinates for up to 10 fingers at a rate of 200 fps [27]. The controller was conveniently placed for our participants at comfortable arm reach; see Fig. 1 for the experiment setup.

3.3 Referents

We selected 21 functions (called from now on "referents" according to the terminology employed in [55, 58]) that represent common tasks to execute during television watching, e.g., change channels and adjust volume, but we also considered new functions recently made available on Smart TVs, e.g., open the web browser. Referents were divided into four categories: (1) nine basic TV commands (the BASIC category): open, close, go to next and previous channel, volume adjustments, and menu commands; (2) three generic commands (GENERIC): yes, no, and ask the system for help; (3) six channel commands (QUICK-CHANNEL): go to favorite and second favorite channels, access a random channel, go back to the last channel, and go to specific channels identified by their numbers; and (4) three feature-related commands (TV-FEATURE), such as show the TV guide, show the list of all channels, and open the web browser. Table 1 lists all the 21 referents employed in our

Table 1 Set of referents used for the gesture elicitation experiment

No	Referent	Description
BASIC referents (9)		
1	Open	Open the TV set
2	Close	Close the TV set
3	Next	Go to next channel
4	Previous	Go to previous channel
5	Volume up	Increase sound volume
6	Volume down	Decrease sound volume
7	Volume mute	Turn off volume
8	Open menu	Open a generic contextual menu
9	Hide menu	Hide/close the contextual menu
GENERIC referents (3)		
10	Help	Ask system for help (e.g., show the Help screen)
11	Yes	Enter affirmative answer to a system elicited Yes or No question
12	No	Enter negative answer to a system elicited Yes or No question
QUICK-CHANNEL referents (6)		
13	Go to favorite channel	Quick access to user's favorite channel
14	Go to 2nd favorite channel	Quick access to user's second favorite channel
15	Go to random channel	Have the TV choose a channel to watch, at random
16	Go to channel 7	Quick access to channel number 7
17	Go to channel 27	Quick access to channel number 27
20	Last channel	Quick access to the last channel that the user watched
TV-FEATURE referents (3)		
18	TV Guide	Open the TV guide
19	Show channels list	Show the list of available TV channels
21	Open browser	Open web browser

study. Our set of referents is similar to those used in previous gesture elicitation studies, e.g., Vatavu [49] employed 12 referents (our BASIC and GENERIC categories); Wu et al. [60] used 18 referents (out of which 9 are included in our BASIC category—Wu et al. [60] focus more on content play, such as the "fast forward" or "play song" functions); and Morris [35] used 15 referents in a study focused on the content displayed in a web browser. While we relied on these previous studies to inform our set of referents, we also considered new referents specific for television watching. For example, we included functions to gain quick access to important channels, i.e., the favorite channel, but also two referents to help us understand how participants prefer to refer to channel numbers using hand gestures, i.e., "Go to channel 7" and "Go to channel 27".

3.4 Task

Participants were seated in a comfortable chair at approximately 2 m from the TV set. The experimenter was present for the entire duration of the study with the role to introduce participants to the features of the Leap Motion controller during the training session and to supervise the data collection procedure during the actual elicitation experiment. Before running the study, participants were given some time to familiarize themselves with the equipment and to discover its active sensing area, i.e., the 3-D volume above the device in which the hand is detectable by the device. During this stage, the Leap Motion visualizer was active on the TV screen (see Fig. 1) so participants could observe not only the tracking capabilities, but also the limitations of the Leap Motion controller.

Each referent was presented to each participant using a text message displayed on the TV screen, and participants were asked to propose a suitable gesture command for that referent. Participants took as much time as they needed to propose gestures. Once they were confident about their gesture proposals, gestures were recorded by the Leap Motion controller and annotated by our software. The experimenter noted down the execution details of each gesture, e.g., "the participant hold the palm facing down and rotated the arm around its own axis to hold the open palm facing up".

Referents were presented in a random order with 21 trials, one trial per referent. At the end of the experiment, participants filled in a questionnaire in which they went through all the referents one more time and tried to recall their own gesture proposals. For each gesture, participants rated the degree to which their gestures fit the referents using scores ranging from 1—"no fit at all" to 5—"very well fit". During this process, participants were also asked to perform the gesture one more time so that the experimenter could also rate how easily they were able to remember their own gestures, which he did on a 5-point Likert scale with 1 denoting "immediate recall" and 5 "no recall". The experimenter's evaluation on whether a gesture was correctly replayed by the participant was informed by his notes taken just before. If participants were not able to recall their gestures for some referent, they were allowed to think of a new one. Participants also rated their likeliness to remember gestures with a score ranging from 1 denoting "very easy to remember" to 5 − "very difficult". They also rated on 5-point Likert scales whether they preferred the proposed gesture or a TV remote button (a TV remote was available for participants to consult at this stage). The experiment took on average 35 min per participant.

4 Results

4.1 Agreement between participants

We measured the degree of agreement between our participants' gesture proposals by calculating individual agreement rates [55, 58, 59] for each referent with the agreement rate formula of Vatavu and Wobbrock [55]:

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

where r is the referent for which agreement is computed and P_i represent subsets of identical or similar gesture proposals from the entire set of gestures collected from our $|P| = 18$ participants. Agreement rates vary between 0 (corresponding to the case in which every participant proposed a distinct gesture for the same referent r) and 1 (perfect consensus between participants, all suggesting the same gesture for a given referent r). We refer the reader to Vatavu and Wobbrock [55] for more details about the agreement rate formula as well as detailed run-through examples of agreement rate analysis. We also measure consensus between participants with Kendall's W coefficient of concordance [23], which is a normalization of the Friedman statistic used to assess the agreement between multiple raters with a number ranging between 0 (no agreement at all) and 1 (perfect agreement):

$$W = \frac{12\sum_{i=1}^{|R|}\left(R_i - \overline{R}\right)^2}{|P|^2(|R|^3 - |R|)} \quad (2)$$

where |P| is the number of participants and |R| the number of referents, i.e., $|P| = 18$ and $|R| = 21$ for this study, R_i represents the total rank of referent i and $\overline{R}$ the mean of all ranks R_i.

We found a mean agreement rate of .158 (SD = .063) across all referents, see Fig. 2a. This small value for agreement was confirmed with Kendall's W coefficient that was .254 ($\chi^2(20) = 91.439, p < .001$). As Kendall's coefficient is related to the average of Spearman rank correlation coefficients between pairs of rankings [23] (p. 276), we can interpret the magnitude of its effect as medium according to Cohen's suggested limits for appreciating effect size, i.e., the value of W is less than .300, but greater than .100. Agreement analysis with the V_{rd} test statistic of Vatavu and Wobbrock [55] showed a significant effect of referent type on agreement rates ($V_{rd(20,N=378)} = 560.973$, $p < .001$).

The highest agreement rate was obtained for the "Next" and "Previous channel" referents (.601 and .516, respectively), for which most participants proposed hand movements to the left and right. We detected a significant difference between the two rates ($V_{rd(1,N=36)} = 4.568, p < .05$) and computed a coagreement of .438 showing that 67 pairs out of all the $18 \times 17/2$ pairs of participants were in agreement for both these referents. The lowest agreement was obtained for abstract tasks, such as "Open browser", "Show channels list", and "Volume mute" (Fig. 2a) that scored agreement rates of .033, .026, and .020, respectively. "Volume mute" was the only referent for which the agreement rate was not significantly greater than zero ($V_{rd(1,N=36)} = 3.000, n.s.$). Figure 2c shows the average agreement rates computed for each of the four categories of referents. The highest agreement rate was .225 for the Basic category (Kendall's $W = .310, \chi^2(8) = 44.584, p < .001$), followed by .140 for Quick-Channel ($W = .161, \chi^2(5) = 14.504, p < .05$), .126 for Generic ($W = .082\ \chi^2(2) = 2.943, n.s.$), and .026 for TV-Feature ($W = .094, \chi^2(2) = 3.391, n.s.$). These results are explained by the fact that the Basic category includes referents with embedded scale range information (e.g., concepts such as up and down, next and previous, etc.) for which users are more likely to reach agreement, while the TV-Feature category includes abstract tasks. For reference, we list all participants' gesture proposals for the entire set of 21 referents under the "Appendix" section.

4.2 Experienced versus novice users

Eight (8) participants had previously used gestures for video games. To understand the effect of previous experience on consensus, we calculated agreement rates

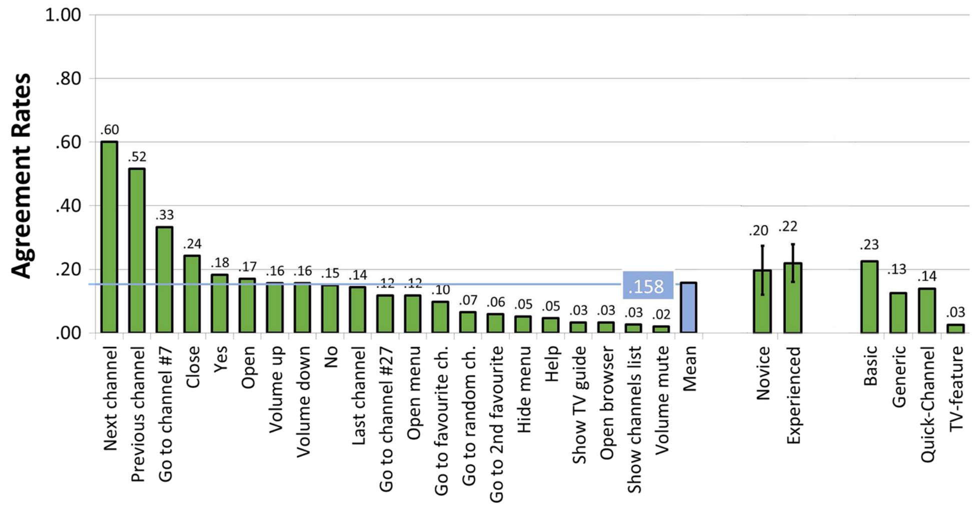


Fig. 2 Agreement rates computed using the formula of Vatavu and Wobbrock [55] for all the 21 referents in our set (**a**), comparison between novice and experienced users (**b**), and comparison between various categories of referents (**c**)

distinctly for the two groups. Results showed higher agreement for the experienced group, .220 versus .197, see Fig. 2b, but a Mann–Whitney test did not detect any significant difference ($U = 172.000, Z = -1.223, n.s$), showing that previous practice with free-hand and body gestures from other application domains, such as gaming, had no influence on participants' consensus for our specific scenario for TV set control.

4.3 Relationship between agreement rates and thinking time

Our participants spent in average 20.5 s (SD = 5.0 s) to figure out suitable gesture commands for each referent. We found a significant negative correlation between agreement rates and thinking time (Pearson's $r_{(N=21)} = -.550$, $p = .01$); see Fig. 3. This result is surprising, because it shows that the more time participants took to think about potential gesture commands, the less agreement resulted in the end. This result can be interpreted in two ways. First, we believe that the more time participants allocated to the task, the more creative they wanted to be and, consequently, they produced gesture commands less likely to be proposed by others. Second, participants' first gesture choice (i.e., the gesture choice after a minimum thinking time) was likely to be discovered by other participants as well, probably due to some internal mechanism of understanding referents, e.g., move hand to left and right for moving to next or previous items in a list.

4.4 Gesture goodness

Participants used a 5-point Likert scale to rate how fit their gesture proposals were for each referent (i.e., gesture goodness), with 1 denoting a gesture command "no fit at all" to its referent, 2—"less fit", 3—"moderate", 4—"good fit", and 5—"very good fit". Overall, the median rating was 4, which shows good confidence of our participants in the gestures they proposed. A Friedman test revealed a significant effect of referent type on self-reported goodness ($\chi^2(20) = 67.761, p < .001$). Four commands were rated "very well fit" (maximum on our scale): "Next channel", "Previous channel", "Volume up," and "Volume down", while the referents with the lowest ratings (3—"moderate fit") were "Go to favorite", "Go to 2nd favorite channel", "Volume mute", "Open browser", and "TV guide".

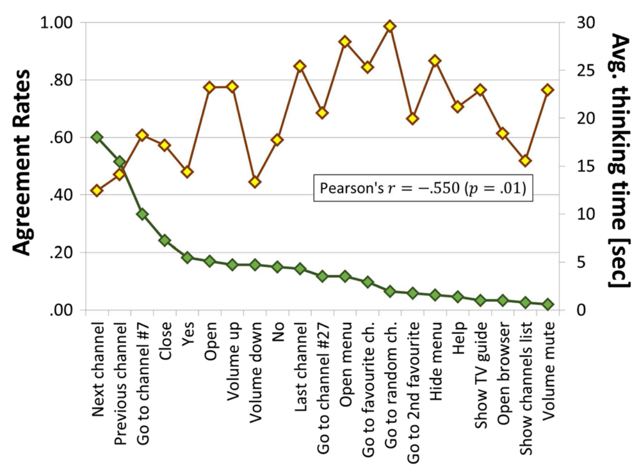


Fig. 3 Correlation between agreement rates and average thinking times, Pearson's $r_{(N=21)} = -.550$, significant at $p = .01$

4.5 Preference for gestures versus the TV remote

Participants rated their preferences about using gesture commands or remote controls using an 11-point Likert scale with values from 5 to 0 and then back to 5 with the

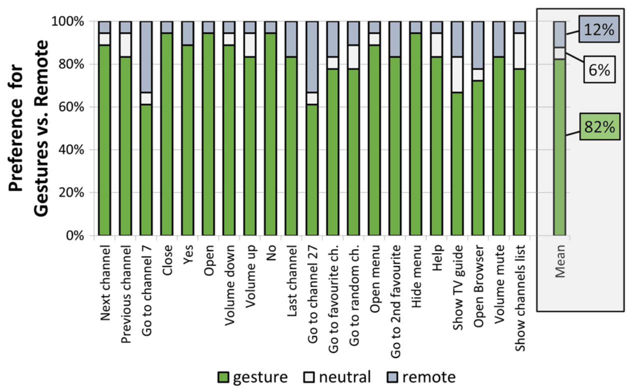


Fig. 4 Participants' self-reported preferences for using leap gestures versus the TV remote control, shown as percentages

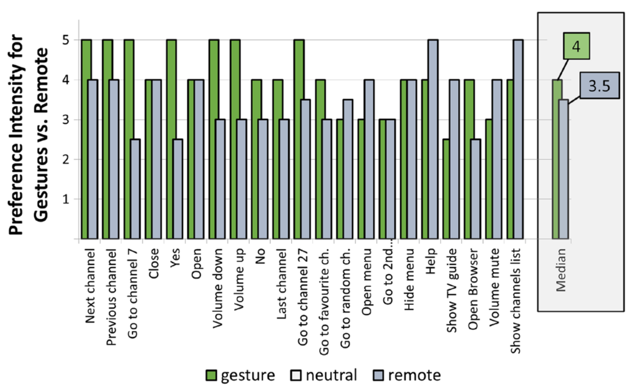


Fig. 5 Participants' self-reported preferences for using leap gestures versus the TV remote control, shown as intensity of the preference

left-most 5 levels encoding their preferences for gestures, 0 a neutral state, and the right-most 5 levels encoding preferences for the remote control. Results were in favor of gestures that were preferred in 82 % of all ratings versus 12 % for the remote, while 6 % were neutral responses; see Fig. 4. The intensity of the preferences measured on another 5-point scale showed a median score of 4 for gestures and 3.5 for the remote control; see Fig. 5.

4.6 Recall rate

Participants were asked how easy they found recalling gestures they had just proposed (i.e., recall likeliness), which they answered using a 5-point Likert scale ranging from 1—"very easy" to 5—"very difficult". The median rating across all participants and referents was 3 corresponding to "moderate difficulty". At the same time, participants had to perform each gesture once more and the experimenter observed their reaction time, which he rated on a 5-point Likert scale with 1 corresponding to "instantaneous recall" and 5 being "no recall at all". The

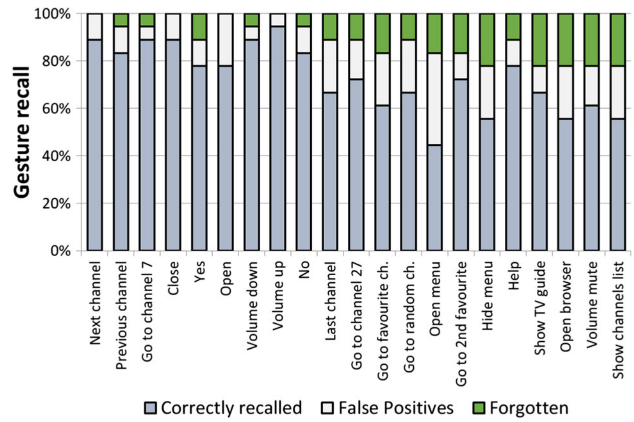


Fig. 6 Gesture recall results. NOTE: referents are listed in descending order of their agreement rate

experimenter's median rating was 1 as the majority of the participants recalled their gestures immediately. However, when analyzing this data further, we found that only 72.8 % of the participants' replays of gestures were correct (out of $18 \times 21 = 378$ gestures), while in 11.4 % of all cases, participants could not remember their gesture proposals and in 15.8 % cases they "recalled" the wrong gesture − a gesture that we refer to as a false positive; see Fig. 6.

A Friedman test showed a significant effect of referent type on the experimenter's rating ($\chi^2(20) = 53.391$, $p < .001$). The referents for which gestures were recalled with the highest accuracy were "Next channel" and "Volume up", while the lowest recall rates occurred for "Open menu" and "Open browser"; see Fig. 6. We also found a significant Pearson correlation between agreement rates and recall likeliness ($r_{(N=18)} = .618, p = .01$), showing that gestures with large consensus are also more likely to be recalled easier.

5 Gesture set

We collected 378 gestures (=18 participants × 21 referents) with corresponding fit-to-function ratings. Based on the agreement rate results (see Fig. 2), we assigned each referent with the gesture that received the highest agreement [58]. For references with low or no consensus at all, we selected for the gesture set one of the participants' gesture proposals that we believed best matched the referent based on our previous experience in gesture interface design. Results are listed in the "Appendix" at the end of this paper. Please note that this gesture set is by no means the definite set of gestures to use in all interactive TV applications. Its main goal is rather to inspire gesture interface designs for the TV rather than to act as a standard. For example, practitioners may opt for a combination of fine-grained finger movements and large arm gestures as in previous work [49, 51, 60], in which case they would use only some of our findings. All the gestures that we collected are available for download at the web address http://www.eed.usv.ro/~vatavu to enable further studies on gestures for the interactive TV.

6 Measures for analyzing Leap Motion gestures

In this section, we define measures to analyze finger and hand gestures captured with the Leap Motion controller. During the training stage of the experiment, participants were encouraged to explore the interactive space of the controller, which is about 0.23 m^3. We wanted to know how much of this volume was actually used by our participants when performing gestures as well as the amount of motion produced by fingers and hands during gesture articulation. To this end, we proposed a set of five features that we used to characterize the spatial and kinematic variations of leap gestures.

6.1 Gesture volume

We define the volume of a gesture as the volume of the bounding box of all fingertips and palm positions detected by the Leap Motion controller during the execution of that gesture, which we express in physical units, i.e., cm^3. Figure 7 illustrates average gesture volumes for all referents. GESTURE-VOLUME can be interpreted as the spatial effort to execute the gesture.

We found a significant effect of referent type on GESTURE-VOLUME ($\chi^2(20) = 74.219, p < .001$). The average volume of all gestures for all referents was 11,277 cm^3 (SD = 17,239). The top 10 % largest volumes were in the range of 13 and 48 % of the total volume tracked by the Leap Motion controller. These results indicate a high propensity of users toward low-volume gestures, reflective of low-effort movements of the hand and fingers. We also found gestures with volumes smaller than 1 cm^3, for which the hand and fingers were kept relatively stable in mid-air. These gestures were mostly symbolic, such as the peace sign or the "OK" sign.

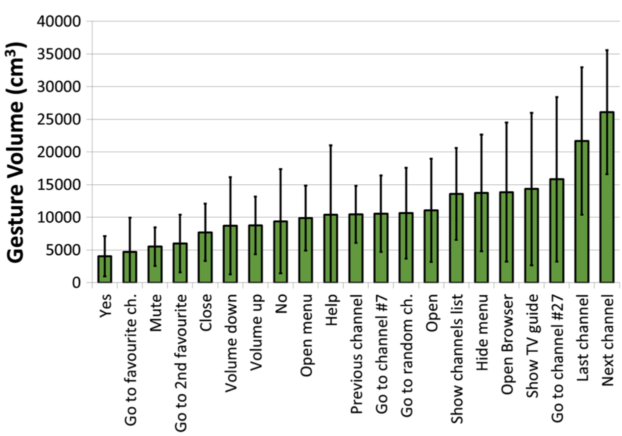


Fig. 7 Average GESTURE-VOLUME values for each referent. NOTE: referents are listed in ascending order of their average volumes; error bars show 95 % CIs

Dichotomous referents had similar GESTURE-VOLUME values, an exception being "Previous" and "Next", for which the average difference was 15,626 cm^3. Whereas the suggested gestures for the two referents are similar (i.e., the hand moves from left to right and from right to left), they are different in how participants articulated them, i.e., movements for "Next" were more ample.

During the experiment, we observed that participants had a tendency to work in a 2-D plane rather than to use the 3-D space enabled by the Leap Motion controller. To find out more, we computed for each gesture the degree by which it varies along each of the x, y, and z axis; see Fig. 8. The figure shows three columns (one for each axis) for each gesture proposed by each participant that encode the variation of that gesture on each axis. The axis with the largest variation is always shown in black color and full height, while the axis with the smallest variation is always white and 30 % of the cell's height. The color and height of the remaining axis are interpolated values between the minimum and maximum variations on the previous two axes. Note how the middle variation tends to stay close to one of the extremes. Actually, we found that the middle axis value was either in the bottom or the top third of the interval between the lowest and highest variations for 67 % of all the gestures, a finding that confirms quantitatively our observations that users prefer to articulate gestures in a 2-D plane.

6.2 Gesture length

We define the centroid of the hand as the average coordinate of data acquired from the Leap Motion controller. GESTURE-LENGTH is defined as the path length of the trajectory of the centroid of the hand during gesture articulation, which we compute as the sum of Euclidean distances between data retrieved at consecutive timestamps. We found a significant effect of referent type on GESTURE-LENGTH ($\chi^2(20) = 41.933, p < .005$). The average length was 75 cm (SD = 64.8); see Fig. 9. We also found a significant positive correlation between gesture volume and length (Pearson's $r_{(N=21)} = .653, p = .05$).

6.3 Finger count

We define FINGER-COUNT as the average number of fingers employed during gesture execution that were tracked by

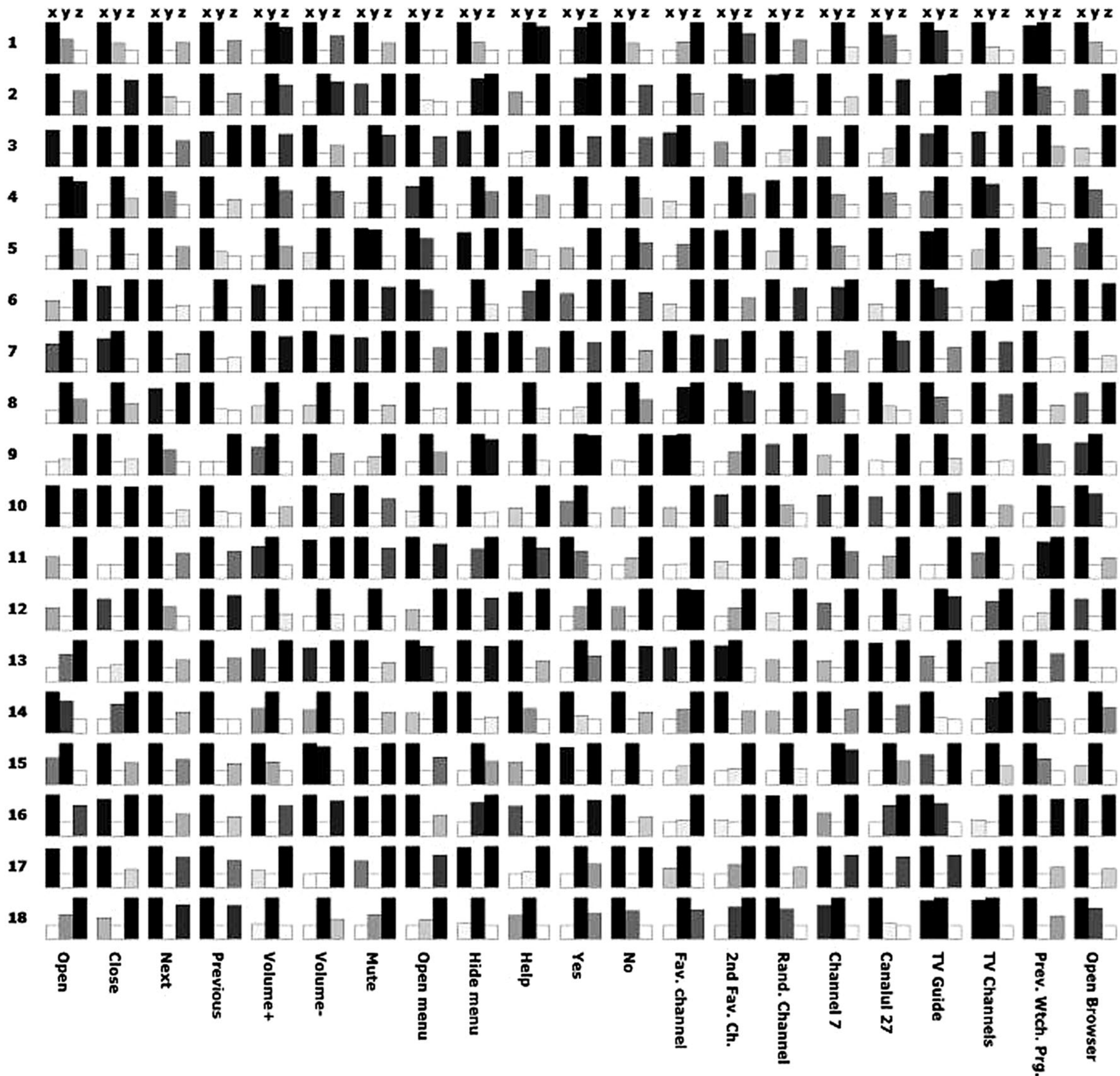


Fig. 8 Variations on the *x*, *y*, and *z* axes for each gesture performed by each participant. The *white column* shows the lowest variation and the black column shows the highest variation

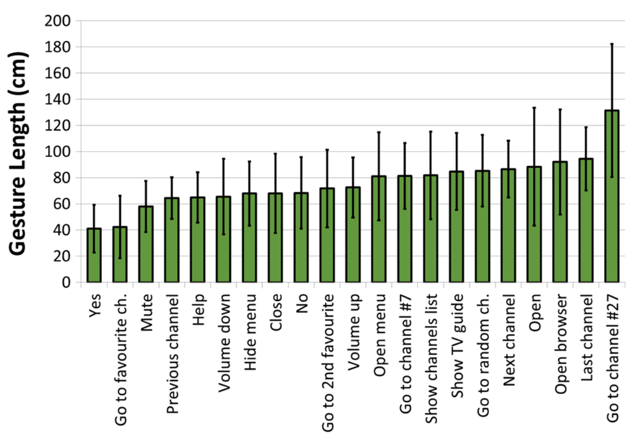


Fig. 9 Average GESTURE-LENGTH values for each referent. NOTE: referents are listed in ascending order of their average lengths; error bars show 95 % CIs

the Leap Motion controller. The average FINGER-COUNT was 1.68 (SD = 1.13); see Fig. 10. We did not detect any significant effect of referent type on FINGER-COUNT ($\chi^2(20) = 22.444, n.s.$ at $p = .01$). The low values of FINGER-COUNT are partly explained by participants' tendency to work in a 2-D plane following a model of touch screen interaction, for which they mostly employed the index finger, e.g., when drawing symbols or letters to nominate a referent. We also observed mid-air equivalents for touch gestures, such as directional swipes and pinch gestures used to increment or decrement values similarly to zooming in and out on touch screens. Participants also employed cultural gestures, such as the "peace" sign or the "OK" sign. Low FINGER-COUNT values are also explained by the limitations of the Leap Motion controller not able to handle

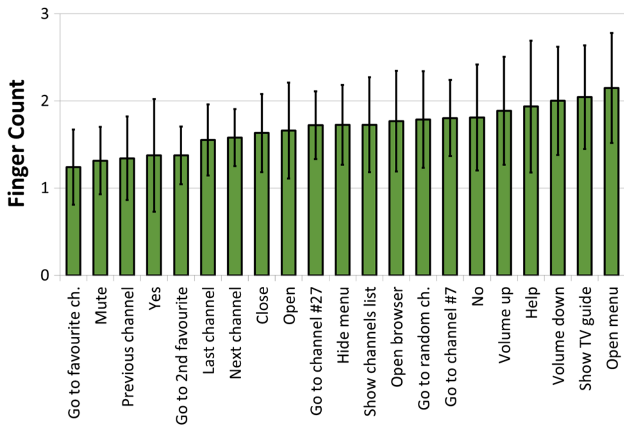


Fig. 10 Average Finger-Count values for each referent. Note: referents are listed in ascending order of their average finger count; error bars show 95 % CIs

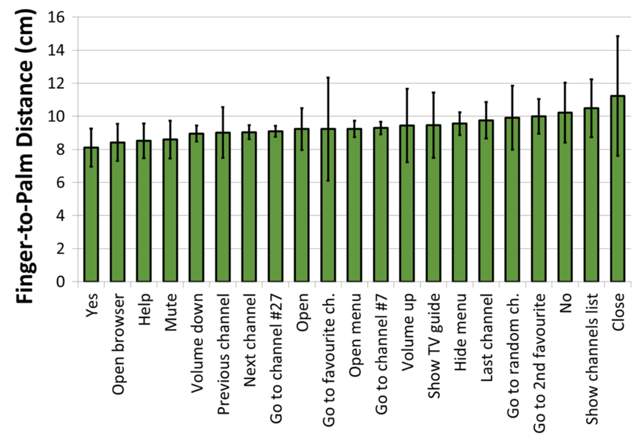


Fig. 11 Average Finger-to-Palm values for each referent. Note: referents are listed in ascending order of their average finger-to-palm distance; error bars show 95 % CIs

occlusions. For instance, the "OK" sign had the thumb finger hidden by the fist.

6.4 Finger-to-palm distance

The Leap Motion controller delivers the palm position for the hand and fingertip positions for all detected fingers. We define the Finger-to-Palm distance as the average variation of distances between fingertips and the palm during gesture articulation:

$$\delta(g) = \frac{1}{n}\left(\sum_{i=0}^{n-1}\frac{1}{NF_i}\sum_{j=0}^{NF_i-1}\|f_j - P\|\right) \quad (3)$$

where g denotes the gesture, n the number of frames that were collected from the Leap Motion controller during the articulation of g, NF_i the number of fingers detected in frame i, and $\|\cdot\|$ denotes the Euclidean distance between two points in 3-D. A similar measure was used in medicine as an outcome of clinical trials [46].

The average Finger-to-Palm distance was 9.4 cm (SD = 3.5); see Fig. 11. We did not detect any significant effect of referent type ($\chi^2(20) = 15.708, n.s.$ at $p = .01$). We believe this is because our participants employed similar motion patterns for fingers based on a model of touch screen interaction, which resulted in only minor variations in the values reported by the Finger-to-Palm measure.

6.5 Articulation speed

We compute the Articulation-Speed of a gesture as its length divided by its execution time, which we express in physical units, e.g., cm/s. Similar to Gesture-Volume, Articulation-Speed characterizes the participants' effort to execute gestures.

We found a significant effect of referent type on Articulation-Speed ($\chi^2(20) = 47.334, p < .005$). The average speed was 52.8 cm/s (SD = 37.3 cm/s); see Fig. 12. Using these results, we can classify gestures into low speed (two referents were assigned gestures with an average speed less than 40 cm/s), medium speed (15 referents had speeds between 40 and 60 cm/s), and high speed (four referents had average speeds faster than 60 cm/s). High-speed gestures involve motions of the hand rather than of fingers (e.g., move the hand from left to right and from right to left, move the hand or fist forward and backward, hand waving, etc.), whereas low-speed gestures employ fingers more (e.g., show the index finger, thumbs-up, perform a

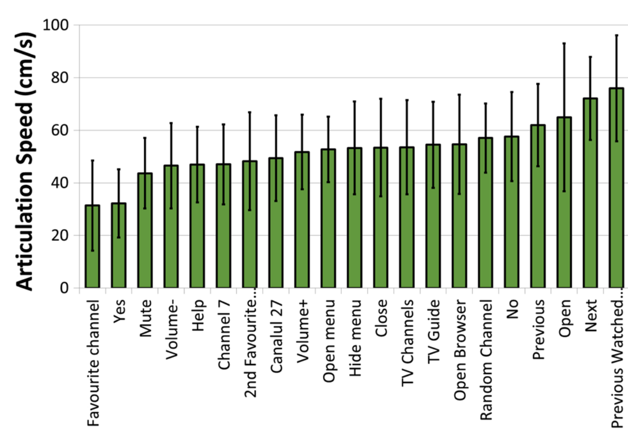


Fig. 12 Average Articulation-Speed values for each referent. Note: referents are listed in ascending order of their articulation speed; error bars show 95 % CIs

click in mid-air, open palm, etc.). However, most referents belong to the medium-speed class (15 out of 21 referents), for which we can identify a mixture of gestures involving both hand and finger motions.

7 Implications for the design of hand and finger gestures for TV control

Our quantitative and qualitative results give insights into the way people define, rate, and evaluate, and later recall fine-grained resolution 3-D finger and hand gestures. In this section, we use our findings to provide a number of nine guidelines for practitioners interested in designing TV interfaces employing such gesture types and similar gesture acquisition technology:

1. *Fine-grained finger and hand gestures are personalized, so user-dependent gesture training is recommended.* We found a low agreement rate between our participants' gesture proposals (average $\mathcal{AR} = .158$). We also found an overwhelming preference for gesture commands instead of the TV remote control: in 82 % of all responses, participants preferred gestures over the remote. The result is surprising given the low agreement rate that we found earlier. However, this finding shows that finger gestures tend to be highly personalized and that user-dependent training is needed in order to avoid poorly designed interfaces with unintuitive mappings between gestures and functions [31, 37].
2. *Users fall back on previously acquired gesture interaction models, such as touch screen interaction—therefore, such expertise should be exploited for mid-air gesture input as well.* During the experiment, we observed an interesting behavior when participants started to think about gesture commands. When having to execute a more difficult task, our participants proposed gestures using a strategy that appeared as iterative design, which they followed through until they reached a simple and familiar gesture for that referent. For example, participants sometimes noted the similarity of the gesture they proposed with some touch screen gesture, such as mid-air directional movements for "Next" and "Previous channel" and touch screen directional swipes.
3. *Users show preference for 2D gestures, so design gestures that can be performed in 2D.* We found that our participants mostly employed the 3-D gesture-sensing device to articulate 2-D gestures. Most of the gestures we collected can be executed in a 2-D plane without any major loss of detail. For example, directional movements of the hand and drawing letters and symbols occurred mostly in a vertical plane. For some gestures, participants imagined a 2-D plane above the Leap Motion controller that they used as a support for drawing.
4. *Users prefer either hand motion or hand poses.* To find out more about our participants' gesture preferences, we classified gestures into the four classes of the taxonomy of Vatavu and Pentiuc [48]: simple static (i.e., hand poses), simple dynamic (sequences of hand poses, but no motion), complex static (only motion is important), and complex dynamic (both motion and hand pose are important). For 40 % of our participants' gestures, only motion was relevant, followed by 38 % of hand gestures involving only postures, either static or combinations of postures. Of all gestures, 22 % involved combinations of hand pose and motion. Figure 13 shows the distribution of gestures according to the 4-class taxonomy of Vatavu and Pentiuc [48].
5. *Users associate gestures and referents in ways that help them maximize recall rate.* This behavior was revealed by the recall percentages (see Fig. 6) with similar values for dichotomous gestures. When encountering referents with opposite effects, e.g., "Next" and "Previous channel", "Volume up" and "Volume down", most participants considered that their corresponding gestures should also be similar.
6. *There is some preference for culture-specific gestures.* We observed several gestures with cultural meaning, e.g., thumbs-up, hand wave, fingers closing in the "shut-up" hand gesture, etc. However, cultural gestures that we report in this work are common for Western cultures and they may prove inappropriate for other cultures. Also, the right-to-left and left-to-right movements for "Previous" and "Next channel" are also probably connected with the left-to-right reading order in this culture.

Our findings also confirm results reported in a previous study addressing free-hand and body gestures [49]:

7. *Exploit hand pose to distinguish between different commands.* Hand pose is important to differentiate

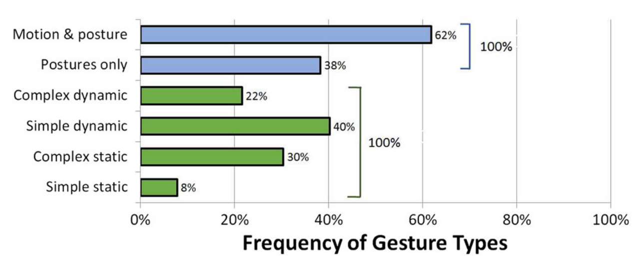


Fig. 13 Frequency distribution of participants' leap gestures according to the taxonomy of Vatavu and Pentiuc [48]

between gestures with similar hand motion. For instance, hand waving was executed by our participants with all fingers to denote "Help", but with only three fingers to mean "Go to the favorite channel". In a previous work, Vatavu [49] also reported the importance of hand pose to differentiate between free-hand gesture commands, an aspect that we were able to observe again in this study.

8. *Users draw letters in mid-air to execute referents that start with those letters.* In many cases, participants suggested letters to identify referents, especially abstract ones, such as letter "M" for "Open menu", "G" for "TV Guide", or "B" for "Open browser", etc. Sometimes, there were multiple letter suggestions for the same referent, such as "C" and "L" for "Show channels list". This finding recommends several gesture input techniques previously developed in the literature for other gesture acquisition technologies, such as augmented letters [42] and marking menus [26], for the interactive TV scenario as well. We also observed the use of symbols, such as drawing "@" to open the browser, or the universal quantifier symbol "∀" to access a random channel. Participants drew digits to specify channels by their numbers. We believe that these results are also explained by our participants' previous experience with touch screen devices on which they produce letters and symbols with stroke gestures.

We also witnessed cases in which participants performed gestures with the support of the non-dominant hand, which approach the idea of the PalmRC prototype of Dezfuli et al. [11]:

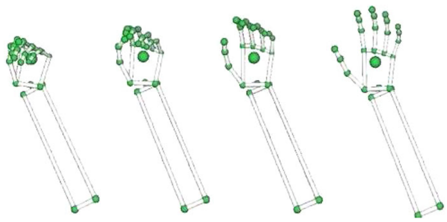
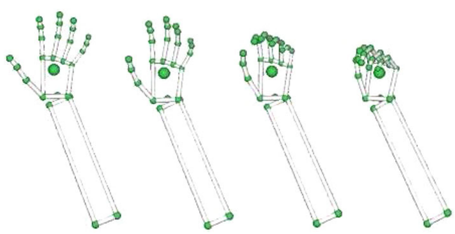
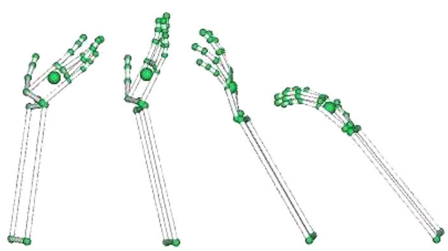
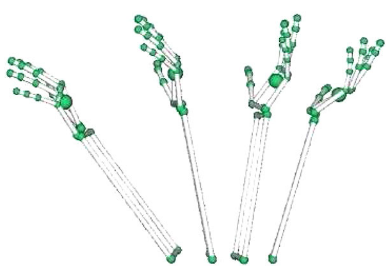
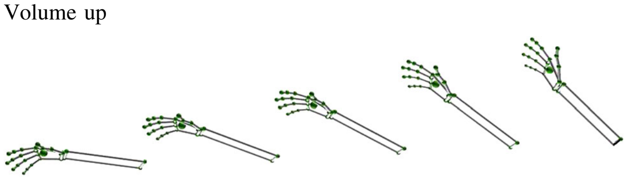
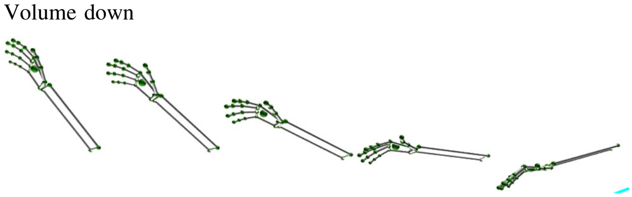
9. *Make use of concrete or imaginary support surfaces to assist users to articulate mid-air gestures.* In some situations, participants employed parts of their hands as active sensing areas. In other cases, participants performed gestures in a vertical plane. Yet in other situations, participants imagined a horizontal plane above the Leap Motion controller that they used as a reference for their gestures.

8 Conclusion

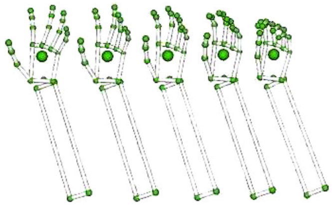
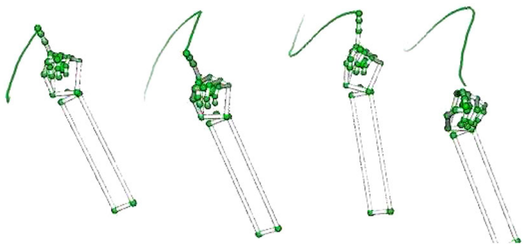
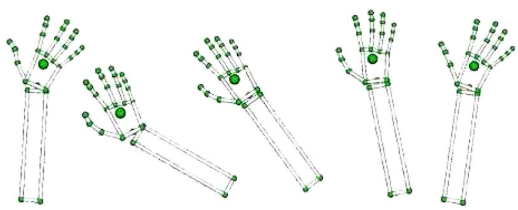
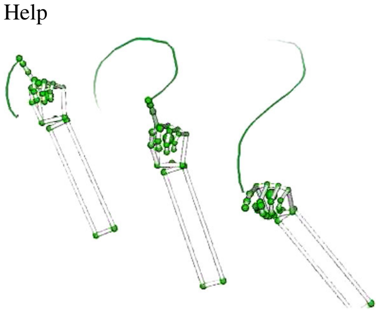
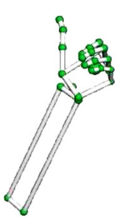
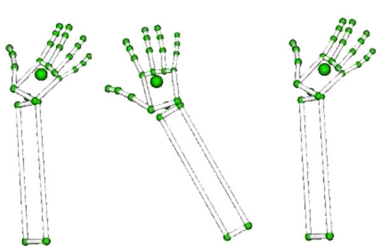
We presented in this work experimental results of the first study conducted on user-elicited fine-resolution finger and hand pose gestures for controlling the interactive TV. We delivered guidelines for working with such fine-resolution gesture types for iTV scenarios employing a Leap Motion controller or a similar gesture acquisition device. We compared our results with previous studies on free-hand gestures for iTV [49] and complemented their findings. To encourage further exploration of such gesture types in the interactive TV community, including development of recognition and interaction techniques, we make available our user-defined dataset composed of 378 gestures with recorded position, direction, and velocity coordinates for hands and fingers. We hope that this exploration on fine-resolution gestures will attract the community attention toward designing viable gesture alternatives for the remote control in the context of lean-back interaction with television.

Acknowledgments This research was supported by the liFe-StaGE project, "Multimodal Feedback for Supporting Gestural Interaction in Smart Environments", No. 740/2014, co-funded by UEFISCDI, Romania and OeAD, Austria. The infrastructure was provided by the University Stefan cel Mare of Suceava and was partially supported from the project "Integrated center for research, development and innovation in Advanced Materials, Nanotechnologies, and Distributed Systems for fabrication and control", No. 671/09.04.2015, Sectoral Operational Program for Increase of the Economic Competitiveness co-funded from the European Regional Development Fund.

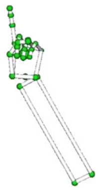
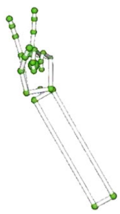
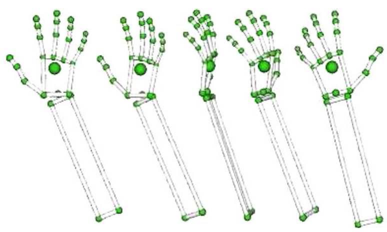
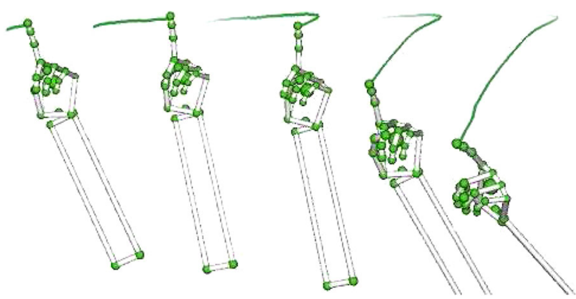
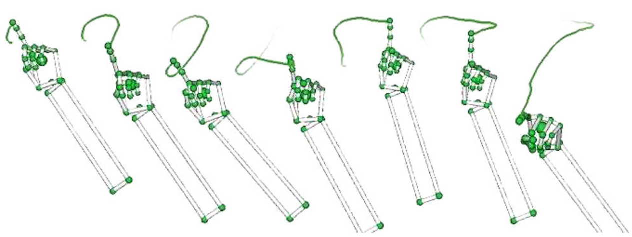
Appendix: Complete set of gestures collected during the elicitation experiment

No.	Referent	Participants' gesture proposals
1	Open 	*Open palm*, wave hand, move hand up, move fist up followed by wrist rotation, move hand down, move palm away from the body, thumbs-up
2	Close 	*Close palm*, wave hand, close hand into pinch, move hand down, fist moves up and down twice, move fist up followed by wrist rotation, hand performs the "go away" cultural gesture, draw "X", perform click in mid-air with the index finger
3	Next 	*Move hand from right to left* and *Move hand from left to right*, thumbs-up while move hand to the right, "OK" sign with the thumb, moving the index finger from left to right and then click in mid-air. Please note the two options in terms of movement direction (left-to-right and right-to-left) that correspond to two distinct metaphors: moving the viewing window (as it happens with scrolling actions and traditional GUI) and moving the items themselves [37]
4	Previous 	*Move hand from left to right* and *Move hand from right to left*, thumbs-up while move hand to the left, perform double click in mid-air with the index finger, draw circle clockwise. Please note the two options in terms of movement direction (left-to-right and right-to-left) that correspond to moving the viewing window (as it happens with the scrolling action and traditional GUI) or the items themselves [37]
5	Volume up 	*Move hand up*, hand in pinch pose expanding fingers, thumbs-up with hand moving to the right, thumbs-up with hand moving up twice, rotate imaginary button to the right, draw the "+" symbol in mid-air, open hand from thumb-index pinch, draw triangle, open palm, draw circle clockwise, hand performs the "go away" cultural gesture
6	Volume down 	*Move hand down*, hand changes pose from open palm to index-thumb pinch, thumbs-up with hand moving to the left, thumbs-up with hand moving down twice, rotate imaginary button to the left, move hand from left to right, draw triangle, hand closing fingers into pinch, draw circle counterclockwise, hand performs the "come here" cultural gesture

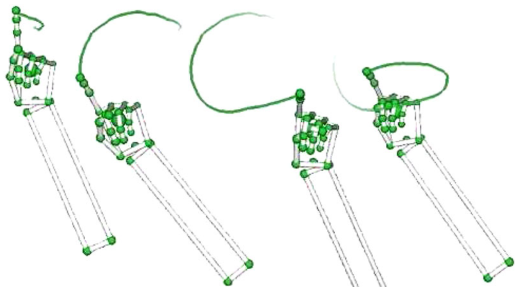
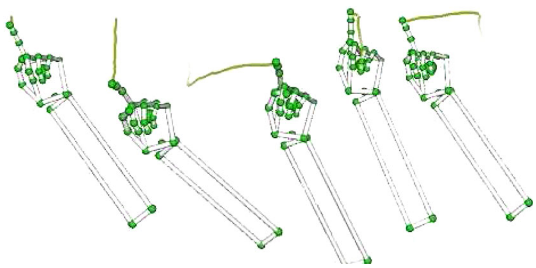
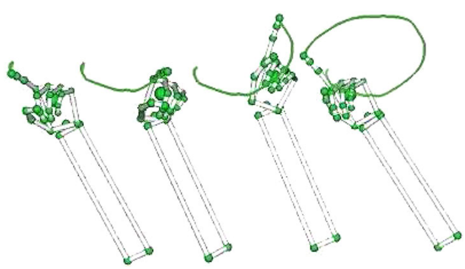
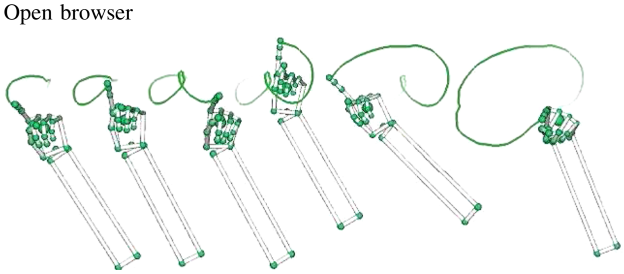
continued

No.	Referent	Participants' gesture proposals
7	Volume mute 	*Close fingers into pinch*, fist followed by extending the little finger, open palm, thumbs-down to thumbs-up, draw letter "X", close fist, open palm to index-thumb pinch, pinch with the thumb and little finger, open palm facing down while hand moves from left to right, draw crossed zero, draw circle counterclockwise, move hand down, move hand from left to right, move hand from right to left, move palm away from the body
8	Open menu 	*Draw letter "M"*, draw small "o", small "m", rotate wrist, move hand up, open fingers from the thumb-index pinch, perform click in mid-air with the index finger, hand in pinch pose opens fingers, wave hand, move hand down
9	Hide menu 	*Hand wave*, closed fist opens index and little finger, close fist, open hand rotates, hand changes from open palm to pinch with the thumb and index finger, move hand up, hand performs the "go away" cultural gesture, rotate palm from left to right, hand performs the "come here" cultural gesture
10	Help 	*Draw symbol "?" or letter "H" in mid-air*, hand performs the "come here" cultural gesture, hand performs the "peace" sign, wave fingers, finger snapping, move hand toward the body, wave hand, pinch and wave
11	Yes 	*Thumbs-up*, draw the "check" sign, three fingers down, draw letter "Y" in mid-air, thumbs-up rotated 90 degrees to the right, perform a click in mid-air with the index finger, open palm, pinch followed by thumbs-up
12	No 	*Hand wave*, close fist, three fingers down, move hand up, thumbs-up rotate from right to up, draw letter "X" in mid-air, move hand down

continued

No.	Referent	Participants' gesture proposals
13	Go to favorite channel 	*Show index finger*, thumbs-up, perform the "come here" cultural gesture, pinch between the thumb and little finger, move index finger down and up, draw the "check" sign, open palm, fingers snapping, draw star, thumbs-up rotated to right, move hand down and back then and forward and down
14	Go to 2nd favorite channel 	*Show two fingers (index and middle)*, thumbs-up rotated 90 degrees to the right, pinch with the thumb and the ring finger, two fingers up, "peace" sign left to right, fingers snapping twice, draw digit "2" in mid-air, two fingers up followed by thumbs-up, move hand forward two times
15	Go to random channel 	*Palm facing down rotates to palm up*, draw circle clockwise, draw circle anticlockwise, draw circle anticlockwise three times, close fingers, open palm rotates down to up, wave with fingers, perform a click in mid-air with the index finger, open palm, draw letter "R", wave hand, draw the universal quantifier symbol "∀"
16	Go to channel #7 	*Draw digit "7" in mid-air*, various finger configurations to indicate number "7"
17	Go to channel #27 	*Draw number "27" in mid-air*, various finger configurations to indicate number "27"

continued

No.	Referent	Participants' gesture proposals
18	TV Guide 	*Draw letter "G" in mid-air*, move hand up, thumbs-up, draw letter "M", move palm in front of the body, move hand down, move hand from left to right, show the index finger, pinch followed by wave, open palm
19	Show channels list 	*Draw square in mid-air*, four fingers move down, open palm, draw letters "C" and "L", move hand down, hand performs the "come here" cultural gesture, wave hand, move hand up, pinch followed by wave, perform double click in mid-air with the index finger, move hand toward the body, draw letter "L" in mid-ar, draw circle anticlockwise
20	Last channel 	*Draw circle anti-clockwise*, index and middle fingers move to the left, fingers snapping, two fingers up, thumbs-up moving back, hand rotates an imaginary button to the left, wave hand
21	Open browser 	*Draw symbol "@"*, move palm down twice, close fist, draw letter "W" three times as in "WWW", hand performs the "go away" cultural gesture, pinch twice, pinch followed by thumbs-up, move index finger from left to right, move hand from right to left, move hand down, move palm in front of the body, show three fingers

References

1. Aslan I, Uhl A, Meschtscherjakov A, Tscheligi M (2014) Mid-air authentication gestures: an exploration of authentication based on palm and finger motions. In: Proceedings of the 16th international conference on multimodal interaction (ICMI '14). ACM, New York, NY, USA, pp 311–318
2. Barkhuus L (2009) Television on the internet: new practices, new viewers. In: Proceedings of CHI '09 extended abstracts on human factors in computing systems (CHI EA '09). ACM, New York, NY, USA, pp 2479–2488
3. Bernhaupt R, Obrist M, Weiss A, Beck E, Tscheligi M (2008) Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. Comput Entertain 6(1), Article 5 (May 2008), 23 pages
4. Benford S, Greenhalgh C, Craven M (2000) Inhabited television: broadcasting interaction from within collaborative virtual environments. ACM Trans Comput Human Interact 7(4):510–547
5. Bobeth J, Schmehl S, Kruijff E, Deutsch S, Tscheligi M (2012) Evaluating performance and acceptance of older adults using freehand gestures for TV menu control. In: Proceedings of the 10th European conference on interactive TV and video (Euro-ITV'12). ACM, New York, NY, USA, pp 35–44
6. Cesar P, Chorianopoulos K (2009) The evolution of TV systems, content, and users toward interactivity. Found Trends Human Comput Interact 2(4):279–373

7. Cesar P, Bulterman, DC, Jansen AJ (2008) Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content. In: Tscheligi M, Obrist M, Lugmayr A (eds) Proceedings of the 6th European conference on changing television environments (EuroITV '08). Springer, Berlin, pp 168–177
8. Christian S, Alves J, Ferreira A, Jesus D, Freitas R, Vieira N (2014) Volcano salvation: interaction through gesture and head tracking. In: Proceedings of CHI '14 extended abstracts on human factors in computing systems (CHI EA '14). ACM, New York, NY, USA, pp 297–300
9. Codd-Downey R, Stuerzlinger W (2014) LeapLook: a free-hand gestural travel technique using the Leap Motion finger tracker. In: Proceedings of the 2nd ACM symposium on spatial user interaction (SUI '14). ACM, New York, NY, USA, pp 153–153
10. Cooke NJ (1994) Varieties of knowledge elicitation techniques. Int J Hum Comput Stud 41(6):801–849
11. Dezfuli N, Khalilbeigi M, Huber J, Mller F, Mhlhuser M, Palm RC (2012) Imaginary palm-based remote control for eyes-free television interaction. In: Proceedings of the 10th European conference on interactive TV and video (EuroITV '12). ACM, New York, NY, USA, pp 27–34
12. Dezfuli N, Khalilbeigi M, Mhlhuser M, Geerts D (2011) A study on interpersonal relationships for social interactive television. In: Proceddings of the 9th international conference on interactive television (EuroITV '11). ACM, New York, NY, USA, pp 21–24
13. Dias T, Variz M, Jorge P, Jesus R (2013) Gesture interaction system for social web applications on smart TVs. In: Proceedings of the 10th conference on open research areas in information retrieval (OAIR '13), pp 225–226
14. Elsweiler D, Mandl S, Lunn BK (2010) Understanding casual-leisure information needs: a diary study in the context of television viewing. In: Proceedings of the 3rd symposium on information interaction in context (IIiX '10). ACM, New York, NY, USA, pp 25–34
15. Fanciulli M (2008) Principles of entertainment in inhabited television. In: Proceedings of the working conference on advanced visual interfaces (AVI '08). ACM, New York, NY, USA, pp 5–12
16. Freeman D, Vennelakanti R, Madhvanath S. Freehand pose-based gestural interaction: studies and implications for interface design. In: Proceedings of the 4th international conference on intelligent human computer interaction (IHCI '12), pp 1–6
17. Freeman WT, Weissman CD (1995) Television control by hand gestures. In: Proceedings of the IEEE international workshop on automatic face and gesture recognition
18. Hantrakul L Kaczmarek K (2014) Implementations of the Leap Motion device in sound synthesis and interactive live performance. In: Proceedings of the 2014 international workshop on movement and computing (MOCO '14). ACM, New York, NY, USA, 4 pages
19. Jeong S, Song T, Kwon K, Jeon JW (2012) TV remote control using human hand motion based on optical flow system. In: Proceedings of the 12th international conference on computational science and its applications (ICCSA'12). Springer, Berlin, Heidelberg, pp 311–323
20. Jones LA, Lederman SJ (2006) Human hand function. Oxford University Press Inc, New York
21. Juhlin O, Önnevall E (2013) On the relation of ordinary gestures to TV screens: General lessons for the design of collaborative interactive techniques. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '13). ACM, New York, NY, USA, pp 919–930
22. Kane SK, Wobbrock JO, Ladner RE (2011) Usable gestures for blind people: Understanding preference and performance. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '11). ACM, New York, NY, USA, pp 413–422
23. Kendall MG, Babington Smith B (1939) The problem of m rankings. Ann Math Stat 10(3):275287
24. Kray C, Nesbitt D, Dawson J, Rohs M (2010) User-defined gestures for connecting mobile phones, public displays, and tabletops. In: Proceedings of the 12th international conference on human–computer interaction with mobile devices and services (MobileHCI '10). ACM, New York, NY, USA, pp 239–248
25. Kühnel C, Westermann T, Hemmert F, Kratz S, Müller A, Möller S (2011) I'm home: defining and evaluating a gesture set for smart-home control. Int J Hum Comput Stud 69(11):693–704
26. Kurtenbach G, Buxton W (1991) Issues in combining marking and direct manipulation techniques. In: Proceedings of the 4th annual ACM symposium on user interface software and technology (UIST '91). ACM, New York, NY, USA, pp 137–144
27. Leap Motion. https://www.leapmotion.com/
28. Lim JH, Jo C, Kim D-H (2012) Analysis on user variability in gesture interaction. Convergence and hybrid information technology. Springer LNCS 7425, pp 295–302
29. Magic Remote. http://www.lg.com/global/magicremote/
30. Macek T, Labský M, Vystrčil J, Luksch D, Kašparová T, Kunc L, Kleindienst J (2014) Interactive car owner's manual user study. In: Proceedings of the 6th international conference on automotive user interfaces and interactive vehicular applications (AutomotiveUI '14). ACM, New York, NY, USA, Article 29, 4 pages
31. Malizia A, Bellucci A (2012) The artificiality of natural user interfaces. Commun ACM 55(3):36–38
32. Matos N, Santos A, Vasconcelos A (2014) Kinteract: a multi-sensor physical rehabilitation solution based on interactive games. In: Proceedings of the 8th international conference on pervasive computing technologies for healthcare (PervasiveHealth '14). ICST, Brussels, Belgium, pp 350–353
33. 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 '10 extended abstracts on human factors in computing systems (CHI EA '10). ACM, New York, NY, USA, pp 4015–4020
34. Morris MR, Wobbrock JO, Wilson AD (2010) Understanding users' preferences for surface gestures. In: Proceedings of graphics interface 2010 (GI '10). Canadian Information Processing Society, Toronto, Ont., Canada, Canada, pp 261–268
35. Morris MR (2012) Web on the wall: insights from a multimodal interaction elicitation study. In: Proceedings of the 2012 ACM international conference on interactive tabletops and surfaces (ITS '12). ACM, New York, NY, USA, pp 95–104
36. Nadamoto A, Tanaka K (2005) Complementing your TV-viewing by web content automatically-transformed into TV-program-type content. In: Proceedings of the 13th annual ACM international conference on multimedia (MULTIMEDIA '05). ACM, New York, NY, USA, pp 41–50
37. Norman DA (2010) Natural user interfaces are not natural. Interactions 17(3):6–10
38. Olsen DR, Partridge B, Lynn S (2010) Time warp sports for Internet television. ACM transactions on computer–human interaction, vol 17, no 4, Article 16, (December 2010), 37 pages
39. Phillips Ambilight. http://www.philips.co.uk/c/televisions/33092/cat/
40. Philips uWand. http://www.uwand.com/
41. Ren G, O'Neill E (2013) Freehand gestural text entry for interactive TV. In: Proceedings of the 11th European conference on interactive TV and Video (EuroITV '13). ACM, New York, NY, USA, pp 121–130
42. Roy Q, Malacria S, Guiard Y, Lecolinet E, Eagan J (2013) Augmented letters: mnemonic gesture-based shortcuts. In: Proceedings of the SIGCHI conference on human factors in

computing systems (CHI '13). ACM, New York, NY, USA, pp 2325–2328
43. Ruiz J, Li Y, Lank E (2011) User-defined motion gestures for mobile interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '11). ACM, New York, NY, USA, pp 197–206
44. Samsung Smart TV: TV Gesture Book. http://www.samsung.com/global/microsite/tv/common/guide_book_5p_vi/waving.html
45. Sen TK, Sinko MW, Wilson AT, Hoque ME (2014) M.I.D.A.S. touch: magnetic interactive device for alternative sight through touch. In: Proceedings of the 16th international ACM SIGACCESS conference on computers and accessibility (ASSETS '14). ACM, New York, NY, USA, pp 307–308
46. Torok KS, Baker NA, Lucas M, Domsic RT, Boudreau R, Medsger TA Jr. (2010) Reliability and validity of the delta finger-to-palm (FTP), a new measure of finger range of motion in systemic sclerosis. Clin Exp Rheumatol 28(2 Suppl. 58):S28–S36
47. Vatavu RD, Pentiuc SG (2008) Interactive coffee tables: interfacing TV within an intuitive, fun and shared experience. In: Proceedings of the 6th European interactive TV conference (EuroITV '08). LNCS 5066, Springer, Berlin Heidelberg, pp 183–187
48. Vatavu RD, Pentiuc SG (2008) Multi-level representation of gesture as command for human–computer interaction. Comput Inform 27(6):837–851
49. Vatavu RD (2012) User-defined gestures for free-hand TV control. In: Proceedings of the 10th European conference on interactive TV and Video (EuroITV '12). ACM, New York, NY, USA, pp 45–48
50. Vatavu RD (2012) Point & click mediated interactions for large home entertainment displays. Multimed Tools Appl 59(1):113–128
51. Vatavu RD (2013) A comparative study of user-defined handheld vs. freehand gestures for home entertainment environments. J Ambient Intell Smart Environ 5(2):187–211
52. Vatavu RD (2013) There's a world outside your TV: Exploring interactions beyond the physical TV screen. In: Proceedings of the 11th European conference on interactive TV and video (EuroITV '13). ACM, New York, NY, USA, pp 143–152
53. Vatavu RD, Zaiti IA (2014) Leap gestures for TV: insights from an elicitation study. In: Proceedings of the 2014 ACM international conference on interactive experiences for TV and online video (TVX '14). ACM, New York, NY, USA, pp 131–138
54. Vatavu RD, Mancas M (2014) Visual attention measures for multi-screen TV. In: Proceedings of the 2014 ACM international conference on interactive experiences for TV and online video (TVX '14). ACM, New York, NY, USA, pp 111–118
55. Vatavu RD, Wobbrock JO (2015) Formalizing agreement analysis for elicitation studies: new measures, significance test, and toolkit. In: Proceedings of CHI '15, the 33rd ACM SIGCHI conference on human factors in computing systems. ACM, New York, NY, USA
56. Vatavu RD (2015) Audience silhouettes: peripheral awareness of synchronous audience kinesics for social television. In: Proceedings of the 2015 ACM international conference on interactive experiences for TV and online video (TVX '15). ACM, New York
57. Vatavu RD, Mancas M (2015) Evaluating visual attention for multi-screen TV: measures, toolkit, and experimental findings. Personal and Ubiquitous Computing. Springer, New York
58. Wobbrock JO, Morris MR, Wilson AD (2009) User-defined gestures for surface computing. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '09). ACM, New York, NY, USA, pp 1083–1092
59. Wobbrock JO, Aung HH, Rothrock B, Myers BA (2005) Maximizing the guessability of symbolic input. In: Proceedings of CHI '05 extended abstracts on human factors in computing systems (CHI EA '05). ACM, New York, NY, USA, pp 1869–1872
60. Wu H, Wang J (2012) User-defined body gestures for TV-based applications. In: Proceedings of the 2012 4th international conference on digital home (ICDH '12). IEEE Computer Society, Washington, DC, USA, pp 415–420