Leap Gestures for TV: Insights from an Elicitation Study

Radu-Daniel Vatavu
University Stefan cel Mare of Suceava
Suceava 720229, Romania
vatavu@eed.usv.ro

Ionuț-Alexandru Zaiți
University Stefan cel Mare of Suceava
Suceava 720229, Romania
ionutzaiti@yahoo.com

ABSTRACT
We present insights from a gesture elicitation study in the context of interacting with TV, during which 18 participants contributed and rated the execution difficulty and recall likeliness of free-hand gestures for 21 distinct TV 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 pose gestures. We report lower agreement rates (.20) than previous gesture studies and 72.8% recall rate and 15.8% false positive recall, results that are explained by the complexity and variability of unconstrained finger gestures. Nevertheless, we report a large 82% preference for gesture commands versus TV remote controls. We also confirm previous findings, such as people's preferences for related gestures for dichotomous tasks, and we report low agreement rates for abstract tasks, such as "open browser" or "show channels list" in our specific TV scenario. In the end, we contribute a set of design guidelines for practitioners interested in free-hand finger and hand pose gestures for interactive TV scenarios, and we release a dataset of 378 Leap Motion gesture records consisting in finger position, direction, and velocity coordinates for further studies in the community. We see this exploration as a first step toward designing low-effort high-resolution finger gestures and hand poses for lean-back interaction with the TV set.

Author Keywords
Gesture interfaces; interactive TV; Leap Motion; motion gestures; elicitation study; hand pose; recall likeliness.

ACM Classification Keywords
H.5.2 User Interfaces: Evaluation / methodology; Input devices and strategies (e.g., mouse, touchscreen).

INTRODUCTION
Television represents a valuable component in our lives, not only for delivering information and entertainment [11,12], but also for creating premises for enriched social interaction [1,2,9]. Over the years, content type, content accessibility, and underlying TV technologies have evolved considerably.

TVX '14, June 25 - 27, 2014, Newcastle Upon Tyne, United Kingdom

ACM 978-1-4503-2838-8/14/06…$15.00.
http://dx.doi.org/10.1145/2602299.2602316

Figure 1: Experiment setup for eliciting leap gestures for iTV.

We now witness and explore inhabited, interactive, and internet television systems [1,3,12,30,32] that are augmented by audio surround systems, ambient effects [33,39], and secondary-screen devices [6]. However, interacting with the TV set has remained virtually unchanged, because interactions are still bound to the use of the standard TV remote control. In the context in which researchers see television as a concept on a converging path to interactivity [5], better designs of input devices are required to meet users' expected level of experience unencumbered by the interaction problems frequently reported for standard TV remote controls [2].

We are interested in this work in understanding people's preferences for interacting with TV with free-hand gestures, and we investigate, for the first time in the context of the interactive TV, a new gesture acquisition scenario, *i.e.*, short-range hand pose and 3-D finger movements that we capture with the Leap Motion controller [23] (see Figure 1). We are thus able to provide insights on the use of fine-grained finger gestures for TV interfaces that complement existing research on gesture interaction for TV [4,8,14,17,40,44].

Our contributions are as follows: (1) we collect people's preferences for interacting with the TV set with 3-D finger movements and hand poses that we acquire with the Leap Motion controller (which are referred in this work as *leap gestures*); (2) we contribute a set of design guidelines for using such type of gestures for controlling various functions of the TV set, and release our collected dataset of gestures to the community for further studies. We see our exploration as a first step toward designing low-effort finger movements and hand pose gestures for lean-back control of the TV set.

RELATED WORK

We place our effort in a larger body of work originating from both academia [4,8,10,17,22,40,44] and industry [24,34,37] interested in designing gestural interfaces for the interactive TV, but also in the even larger community of researchers exploring people's preferences for gesture commands [19,21,25,27-29,36,42]. However, we direct our attention toward understanding the use of fine-grained finger movements and hand poses for executing tasks for the TV set. Such gesture types have not been explored before in the interactive TV context, probably because of the lack of accessible technology to capture them. However, the dexterity and multi-functionality of the human hand has been thoroughly studied in psychology [16]. It is our belief that such gesture types are more appropriate in the context of lean-back versus lean-forward interaction paradigms with television than are large body movements [38,40,44].

We connect our work to previous explorations of gesture interfaces for TV. For example, Freeman and Weissman [14] proposed the first TV gesture interface that mapped hand movements to a cursor displayed on the screen; Bobeth et al. [4] investigated the way older adults employ free-hand gestures for controlling TV functions; Dias et al. [10] were interested in designing gesture interfaces for specific applications running on TV; and Vatavu [39] introduced augmented TV spaces for the control of which they proposed pointing gestures captured with an augmented remote. Finally, the PalmRC work of Dezfuli et al. [8] approaches the most our rationale for exploring low-effort short-range gestures for TV. The authors employed the palm of the hand as a supporting surface for finger touch to enable eyes-free control of the TV set.

We conduct this work in the tradition opened by the guessability methodology of Wobbrock et al. [43] that has been successfully applied for gesture elicitation studies in various application domains [21,22,28,29,36,42] including the interactive TV [38,40,44]. For example, Vatavu [40] collected and analyzed free-hand gestures captured at coarse level with the Kinect sensor in what constituted the first gesture elicitation user study for the interactive TV. A follow-up exploration [38] extended the initial findings and gesture set with more discussion centered on people's preferences of gestures versus TV remotes. Wu and Wang [44] were also interested in hand and body gestures that they captured at the same coarse level of detail. Beyond this previous work, we believe there is much opportunity in leveraging the fine-grained dexterous 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 that we capture with the Leap Motion controller [23]. In doing so, we deliver the community with further insight on designing gestural interfaces with this so far unexplored category of gestures. It is our strong belief 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 [8].

EXPERIMENT

We conducted an elicitation experiment [7,42,43] to collect people's preferences for leap gestures in the context of iTV.

Participants

Eighteen (18) volunteers (4 females) participated in the study (mean age 25.0 years, $SD = 3.1$). All participants were right-handed. Ten participants had no previous experience with gestural interfaces, whereas the other eight had used Nintendo Wii and Microsoft Kinect controllers for games. All participants owned touch-screen phones, but touch gestures are different in nature from the mid-air finger movements and hand poses investigated in this work.

Apparatus

A 40-inch (102 cm) Sony TV was connected to a laptop running Microsoft Windows 8.1 and our custom gesture acquisition software collecting Leap Motion gesture data. The Leap Motion controller is a 3-D tracking device that is able to detect and track targets with a precision of up to 0.01 mm in a 3-D space of .227 cubic meters with a 150° field of view, and can report tracked data (*i.e.*, position, direction, and velocity coordinates) for up to 10 fingers at a rate of over 200 frames per second [23]. The Leap controller was conveniently placed for our participants at comfortable arm reach (see Figure 1 on the previous page).

Referents

We selected 21 referents[1] common for television watching, but that also include new functions recently made available on Smart TVs, *e.g.,* open browser. The referents list was divided into four categories: (a) nine basic TV commands (BASIC): open, close, next and previous channel, volume adjustments, and menu commands; (b) three generic commands (GENERIC): yes, no, and ask system for help; (c) six channel query commands (QUICK-CHANNEL): go to favorite and second favorite channels, access random channel, go back to last channel, and go to specific channels identified by their numbers, such as channels #7 and #27; and (d) three feature related commands (TV-FEATURE), such as show TV guide, channels list, and open web browser. Table 1 lists all referents. Our set of referents is similar to those used in previous studies, *e.g.*, Vatavu [40] employed 12 referents (our BASIC and GENERIC); Wu et al. [44] used 18 referents (out of which 9 are our BASIC ones, while they focus more on content play, such as "fast forward" or "play song" functions); and Morris [29] used 15 referents (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 to understand the opportunity to employ gestures for other tasks. For example, we decided to include functions to quickly access important channels (*i.e.*, the favorite channel), but also two referents to understand how participants will refer to channel numbers with hand gestures (*i.e.*, "Go to channel #7" and "Go to channel #27").

[1]We follow the terminology of Wobbrock et al. [42] that used the word *referent* to denote the effect of a gesture command.

Table 1. Set of referents used for the 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	Increasing sound volume
6	Volume down	Decreasing 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 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 #7
17	Go to channel #27	Quick access to channel #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

Task

Participants were seated comfortably at approximately 2 meters from the TV set. The experimenter was present during the entire duration of the study with the role to introduce participants to the features of the Leap Motion controller and to supervise the data collection procedure. Before running the study, participants were given some time to familiarize with the equipment and discover its active sensing area, *i.e.*, the 3-D volume above the device in which the hand is detectable by the device. The elicitation experiment consisted in presenting each referent (Table 1) with a text message on screen followed by an instruction to suggest a suitable gesture command. Participants took as much time as they needed to propose gestures. Once they were confident about their gesture proposals, the experimenter asked participants to reproduce the gesture one more time so that it could be recorded by the Leap Motion controller and annotated by our software.

Referents were presented in a random order, resulting in 21 trials, one trial per referent. At the end of the experiment, participants filled in a questionnaire in which they went through all referents one more time and tried to recall their gesture proposals. For each proposed gesture, participants rated how well it fit to the referent on a 5-point Likert scale, with 1 denoting "no fit at all" and 5 "very well fit". During this process, participants were also asked to perform the gesture one more time so that the experimenter could 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"). If participants were not able to recall the previously proposed gesture, they were asked to propose a new one. Participants also rated their likeliness to remember gestures 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 minutes to complete per participant.

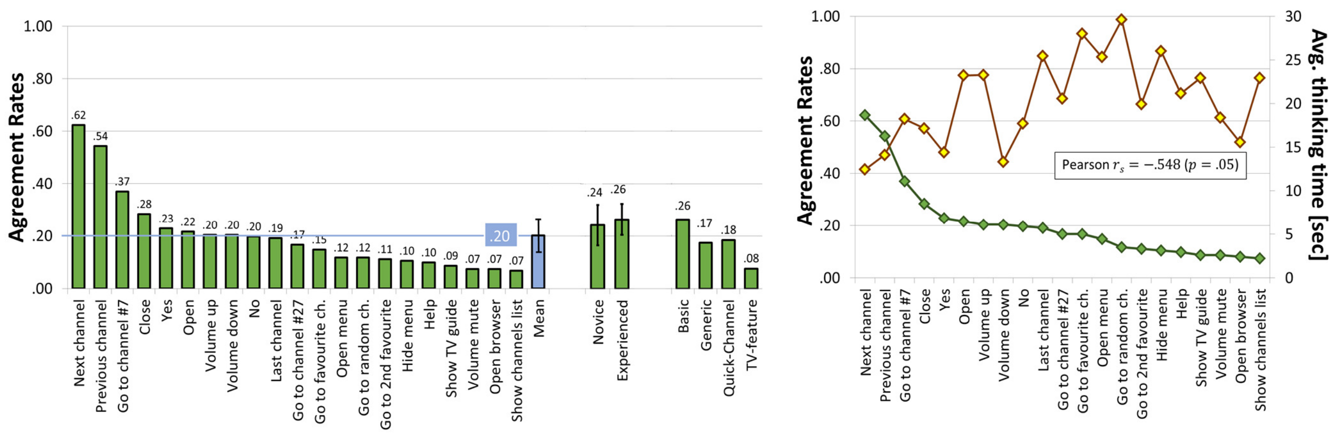
RESULTS

Consensus between participants

We measured consensus by calculating *individual* agreement rates for each referent with the methodology of Wobbrock et al. [42,43], but also by computing *overall* Kendall's W coefficients of concordance [20]. In our case, agreement rates vary between $1/18 = .055$ (corresponding to the case with each participant proposing a distinct gesture for a given referent) and a maximum of 1 (perfect consensus between participants, all suggesting the same gesture for a given referent). We refer the reader to Wobbrock et al. [42,43] for the formula to calculate agreement rates and run-through examples. Kendall's coefficient of concordance [20] 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).

The mean agreement rate across all referents was $.20$ ($SD = .15$), see Figure 2a. This result was confirmed by the value of the Kendall's W coefficient which 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 [20] (p. 276), we can interpret the magnitude of its effect as medium (*i.e.*, less than $.30$, but greater than $.10$) according to Cohen's suggested limits for appreciating effect size. The highest agreement rate was obtained for the "Next" and "Previous channel" commands ($.62$ and $.54$ respectively), for which participants proposed hand movements to left and right. The lowest agreement rates (.07) were obtained for abstract tasks, such as "Volume mute", "Open browser", and "Show channels list" (Figure 2a).

Figure 2c shows the average agreement rates calculated for each of the four categories of referents. The highest agreement rate was $.26$ for BASIC (Kendall's $W = .310$, $\chi^2(8) = 44.584, p < .001$), followed by $.17$ for GENERIC ($W = .082$, $\chi^2(2) = 2.943, n.s.$), .18 for QUICK-CHANNEL ($W = .161$, $\chi^2(5) = 14.504, p < .05$), and $.08$ 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.,* up and down, next and previous, etc.), while TV-FEATURE includes abstract tasks. For reference, we list all participants' gesture proposals for the entire set of 21 referents under the Appendix section.



(a) (b) (c) (d)

Figure 2: Agreement rate results for the 21 referents in our set (a), comparison between novice and experienced users (b), between referent categories (c), and correlation between agreement and average thinking times (d).

Experienced versus novice users

Eight participants had previously used gestures for video games. To understand the effect of previous experience on elicited gestures, we calculated agreement rates distinctly for the two groups. Results showed higher agreement for the experienced group, $.26$ versus $.24$, see Figure 2b. However, the difference was not significant ($U = 172.000, Z = -1.223, n.s$), showing that previous freehand and body gesture practice (from other application domains, such as gaming) had no influence on consensus for TV gestures.

Agreement rate and thinking time

Participants spent in average 20.5 seconds ($SD = 5.0$) to search suitable gesture commands. We found a significant negative correlation between agreement rates and thinking time (Pearson $r_{(N=21)} = -.548, p = .05$), see Figure 2d. This result is surprising, showing that the more time participants took to think about gestures, the less agreement resulted. This finding can be interpreted in two ways. First, the more time participants allocated to the task, the more creative they wanted to be generating gesture commands less likely to be proposed by others. Second, participants' first choice (*i.e.*, the gesture choice after a minimum thinking time) was likely to be found by other participants as well, probably due to some internal mechanism of understanding referent actions, *e.g.,* move hand to left and right for moving to the next or previous item in a list.

Gesture goodness

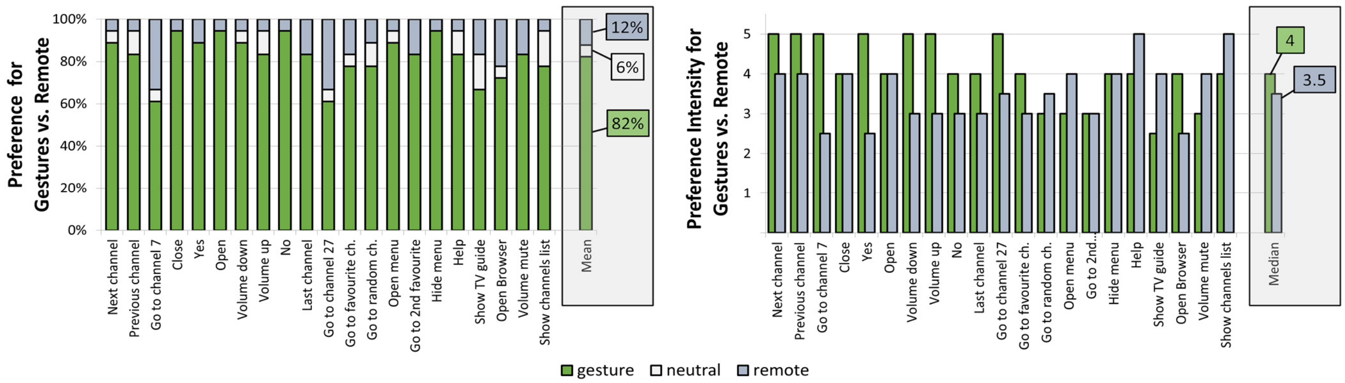
Participants used a 5-point Likert scale to rate how fit their gesture proposals were for each referent (gesture goodness), with 1 denoting "no fit at all", 2 "less fit", 3 "moderate", 4 "good fit", and 5 "very good fit". Overall, the median rating was 4 showing good confidence in proposed gestures.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

Preference for gestures versus the TV remote

We were also interested in participants' preferences for gestures versus the TV remote control. Participants rated their preferences using an 11-point Likert scale going from 5 to 0 and back to 5 with the left-most 5 levels encoding preference for gestures, 0 a neutral state, and the right-most 5 preference for the remote. Results were in favor of gestures that were preferred for 82% of all ratings versus 12% for the remote, while 6% were neutral (Figure 3a). The intensity of the preferences measured on a 5-point scale showed a median score of 4 for gestures and 3.5 for the remote (Figure 3b).

Recall rate

Participants were also asked how easy they found recalling gestures they had just proposed (recall likeliness), which they answered using a 5-point Likert scale from 1 "very easy" to 5 "very difficult". The median rating across all participants and referents was 3, "moderate difficulty". At the same time, while running through the questionnaire, participants had to perform once more each gesture, and the experimenter observed their reaction time and encoded it on a 5-point Likert scale as well, with 1 corresponding to "instantaneous recall" and 5 being "no recall at all". The experimenter's median rating was 1 as the majority of the participants recalled their gestures instantly. However, when further analyzing this data, we found that only 72.8% of the participants' replay of gestures were correct (out of all 18×21=378 gestures), while in 11.4% cases participants could not remember their gesture proposal, and in 15.8% of all cases they "recalled" the wrong gesture (a gesture that we referred to as a *false positive*). Figure 4 illustrates the recall results for each referent. A Friedman test showed a significant effect of referent type on the experimenters' rating ($\chi^2(20) = 53.391, p < .001$). Best recalled gestures were found for "Next channel" and "Volume up", while lowest recall rates occurred for "Open menu" and "Open browser". 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.



(a) (b)

Figure 3: Participants' self-reported preferences for using leap gestures versus the TV remote, shown overall as percentages (a) and intensity of preference (b). NOTE: referents are listed in descending order of their agreement rate (as in Figure 2a).

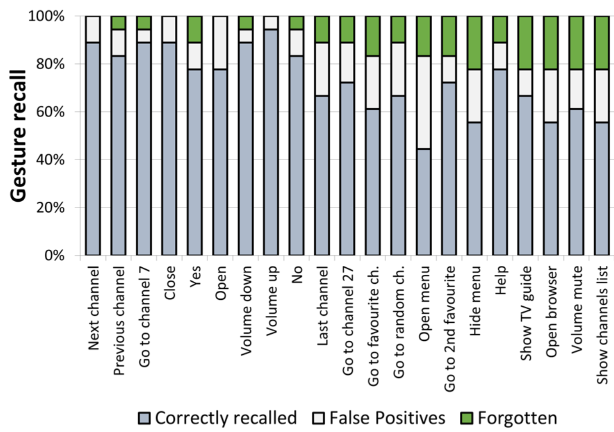


Figure 4. Leap gesture recall. NOTE: referents are listed in descending order of their agreement rate (as in Figure 2a).

GESTURE SET

We collected 378 gestures (=18 participants × 21 referents) with corresponding fit-to-function ratings. Based on the agreement rate results (Figure 2), we assigned each referent with the gesture that received the highest agreement. For references with low or no consensus at all, we selected 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, where we highlighted gestures that made it into the final set. Please note that this gesture set is by no means the definite set to use. Its main goal is rather to inspire gesture interface designs for the TV than to stand as a standard. For example, practitioners may opt for a combination of fine-grained finger movements and large arm gestures as in previous work [38,40,44], in which case they would only use some of our findings. However, besides the "winning" gesture for each referent, we believe there is also value in the other, lower rated gestures as well. For this reason, we decided to list in the Appendix all participants' gesture proposals. We also make available our set of recorded leap gestures to the research community for further studies, such as developing gesture recognizers to support such interfaces (`http://www.eed.usv.ro/~vatavu`).

IMPLICATIONS FOR DESIGN

Our results give insights on the way people define, rate and evaluate, and later recall fine-grained resolution 3-D finger and hand pose gestures. By using our findings, we are able to provide several guidelines for practitioners interested in designing TV interfaces employing such gesture types.

1. Finger and hand pose gestures are preferred to remotes, but there is low agreement between users. We found an overwhelming preference for using gestures instead of the TV remote control. For 82% of all responses, participants preferred gestures over the remote. The result is surprising given the low agreement rate we found for gesture preferences (.20). This finding shows that finger gestures tend to be highly personalized and suggests user-dependent training in order to avoid poorly designed interfaces [26,31] with less intuitive mappings between gestures and functions.

2. Users fall back on previously acquired gesture interaction models. During the experiment we observed one interesting behavior emerging when participants were thinking about gestures. When having to execute a more difficult task, users tended to propose gestures using a strategy that appeared as design iteration until they reached a simple and familiar gesture command. For example, users sometimes noted the similarity of the gesture they executed with touch gestures, such as directional swipes for "Next" and "Previous channel". This behavior originates from previous practice with touch-screen devices [18].

3. Preference for 2-D gestures. We found that users mostly employed the 3-D gesture-sensing device to articulate 2-D gestures. Most of the gestures we collected can be executed in 2-D without any major loss. For example, directional movements of the hand and drawing letters and symbols occurred mostly in a vertical plane in front of the user. For some gestures, users imagined a 2-D plane above Leap Motion that they used as a support for drawing gestures.

4. Users prefer either motion or hand pose gestures, and combinations of these two are less likely. To find out more about our participants' gesture preferences, we analyzed the resulted gesture dataset by classifying gestures into four classes as per the taxonomy of Vatavu and Pentiuc [41]:

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). 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 5 shows the distribution of gestures.

5. Users associate gestures and commands in a way that helps maximize recall rate. This behavior was revealed by the recall percentages (see Figure 4) that show similar values for dichotomous leap gestures. When encountering referents with opposite effects (*e.g.* "Next" and "Previous channel", "Volume up" and "Volume down"), most participants considered gestures should also be similar.

7. Preference for culture-specific leap gestures. We observed many such gestures, *e.g.*, thumbs-up, hand wave, fingers closing in shut-up gesture, etc. The gestures 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" are also probably connected with the left to right reading order.

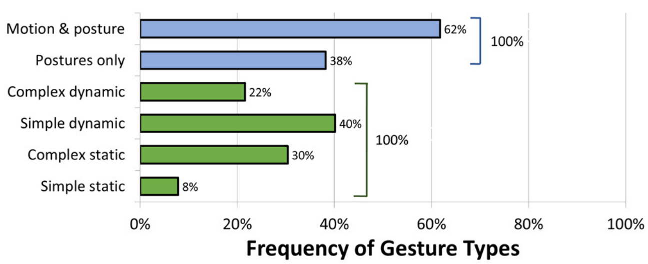


Figure 5. Frequency distribution of participants' leap gestures according to the gesture taxonomy of [41].

With our results we also confirm findings reported in a previous study focusing on large free-hand gestures [40]:

8. Exploit hand pose to distinguish between different commands. Hand pose is important to differentiate between gestures with similar motion. For instance, the colloquial gesture "come to me" was executed with all fingers to denote "Help" but only with three fingers to "Go to the favorite channel". Vatavu [40] also reported the importance of hand pose to differentiate between free-hand gesture commands.

9. Users show preference for drawing letters in mid-air to execute tasks whose names start with those letters. In many cases participants suggested letters to identify tasks, especially abstract ones, such as "Open Menu" (letter "M"), "TV Guide" (letter "G"), "Open Browser" ("B"), etc. There were multiple letter suggestions for the same command, such as both "C" and "L" for the list of channels. We also encountered the use of symbols, such as drawing "@" to open the browser, or the universal quantifier symbol "∀" to access a random channel. Users drew digits to specify channels by their numbers, *e.g.*, "Go to channel #27". We believe this behavior is also explained by participants' previous experience with touch-screens on which they can easily produce letters and symbols with stroke gestures.

We also witnessed cases in which users performed gestures with the support of the non-dominant hand, approaching the idea of the PalmRC prototype of Dezfuli et al. [8], which allows us to derive yet another design guideline:

10. Make use of concrete or imaginary support surfaces to assist users to articulate gestures. We observed situations in which participants employed part of their hands as an active sensing area (*i.e.*, a button). In other cases, participants performed gestures in a vertical plane in front of the body. Yet in other situations, participants imagined a horizontal plane above the Leap Motion controller that they used as a reference for their gestures.

CONCLUSION

We presented results of the first study on fine-resolution gestures for the interactive TV. We delivered guidelines for working with such fine-resolution gesture types for iTV scenarios employing a Leap Motion or similar device. We compared our results with previous studies on free-hand gestures [40] and complemented their findings. To encourage further exploration of such gesture types for iTV, including recognition and interaction techniques, we provide our user-defined dataset composed of 378 gestures with recorded position, direction, and velocity coordinates for hand and fingers. We hope that this first exploration on fine-resolution gestures will attract the community attention toward designing viable gesture alternatives for the remote in the context of lean-back TV interaction.

ACKNOWLEDGMENTS

This research was supported by the project liFe-StaGE, "Multimodal Feedback for Supporting Gestural Interaction in Smart Environments", project no. 740/2014, co-funded by UEFISCDI, Romania and OeAD, Austria.

REFERENCES

1. Barkhuus, L. 2009. Television on the internet: new practices, new viewers. In *Proc. of CHI '09*, 2479-2488
2. 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. *Comp. in Entertainment* 6(1), Art 5
3. Benford, S., Greenhalgh, C., Craven, M., 2000. Inhabited television: broadcasting interaction from within collaborative virtual environments. *ACM TOCHI* 7(4), December 2000, 510-547
4. 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 *Proc. of EuroiTV '12*, 35-44
5. Cesar, P., Chorianopoulos, K. 2009. The Evolution of TV Systems, Content, and Users Toward Interactivity *Found. and Trends in HCI* 2(4), April 2009, 373-95
6. Cesar, P, Bulterman, D.C., Jansen, A.J. 2008. Usages of the Secondary Screen in an Interactive Television Environment: Control, Enrich, Share, and Transfer Television Content. *Proc. of EuroITV '08*, 168-177

7. Cooke, N.J. 1994. Varieties of knowledge elicitation techniques. *Int. Journ. Hum.-Comput. St.* 41(6), 801-849
8. Dezfuli, N., Khalilbeigi, M., Huber, J., Müller, F., Mühlhäuser, M. 2012. PalmRC: imaginary palm-based remote control for eyes-free television interaction. In *Proc. of EuroiTV '12*. ACM, NY, USA, 27-34
9. Dezfuli, N., Khalilbeigi, M., Mühlhäuser, M., Geerts, D. 2011. A study on interpersonal relationships for social interactive television. In *Proc. of EuroITV '11*, 21-24
10. Dias, T., Variz, M., Jorge, P., Jesus, R. 2013. Gesture interaction system for social web applications on smart TVs. In *Proc. of OAIR '13*,. Paris, France, 225-226
11. Elsweiler, D., Mandl, S., Lunn, B.K. 2010. Understanding casual-leisure information needs: a diary study in the context of television viewing. In *Proc. of IIiX '10*, 25-34
12. Fanciulli, M. 2008. Principles of entertainment in inhabited television. In *Proc. of AVI '08*, 5-12
13. Freeman, D., Vennelakanti, R., Madhvanath, S., Freehand pose-based Gestural Interaction: Studies and implications for interface design. In *Proc.of IHCI'12*,1-6
14. Freeman, W.T., Weissman, C.D. 1995. Television Control by Hand Gestures. In *Proc. of the IEEE Int. Workshop on Automatic Face and Gesture Recognition*
15. Jeong, S., Song, T., Kwon, K., Jeon, J.W. 2012. TV remote control using human hand motion based on optical flow system. In *Proc. of ICCSA'12*, Springer-Verlag, Berlin, Heidelberg, 311-323
16. Jones, L.A., and Lederman, S.J. 2006. *Human Hand Function.* Oxford University Press, Inc., New York
17. Juhlin, O., Önnevall, E. 2013. On the relation of ordinary gestures to TV screens: general lessons for the design of collaborative interactive techniques. In *Proc. of CHI '13*. ACM, New York, NY, USA, 919-930
18. Holzinger, A., Finger instead of mouse: touch screens as a means of enhancing universal access. In *Proc. of ERCIM'02*, No.235. 387-397
19. Kane, S.K., Wobbrock, J.O., Ladner, R.E. 2011. Usable gestures for blind people: understanding preference and performance. In *Proc. of CHI '11*, 413-422
20. Kendall, M.G., Babington Smith, B. 1939. The Problem of m Rankings. *Annals of Math. Statistics* 10(3), 275–287
21. Kray, C., Nesbitt, D., Dawson, J., Rohs, M. 2010. User-defined gestures for connecting mobile phones, public displays, and tabletops. *Proc. of MobileHCI '10*, 239-248
22. 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. Journ. Hum.-Comput. St. 69* (11), 693-704
23. Leap Motion, https://www.leapmotion.com/
24. Magic Remote, http://www.lg.com/global/magicremote/
25. Lim, J.H., Jo, C., Kim, D-H. 2012. Analysis on User Variability in Gesture Interaction. In *Proc. of ICHIT'12*. Springer LNCS 7425, 295-302
26. Malizia, A., Bellucci, A. 2012. The artificiality of natural user interfaces. *Commun. ACM* 55(3), 36-38
27. Mauney, D., Howarth, J., Wirtanen, A., Capra, M. 2010. Cultural similarities and differences in user-defined gestures for touchscreen user interfaces. In *Proc. of CHI EA '10*. ACM, New York, NY, USA, 4015-4020
28. Morris, M.R., Wobbrock, J.O., Wilson, A.D. 2010. Understanding users' preferences for surface gestures. In *Proc. of GI '10*, 261-268
29. Morris, M.R. 2012. Web on the wall: insights from a multimodal interaction elicitation study. In *Proc. of ITS'12*. ACM, New York, NY, USA, 95-104
30. Nadamoto, A., Tanaka, K. 2005. Complementing your TV-viewing by web content automatically-transformed into TV-program-type content. In *Proc. of MULTIMEDIA '05.* ACM, New York, NY, USA, 41-50
31. Norman, D.A. 2010. Natural user interfaces are not natural. *Interactions* 17(3), May 2010, 6-10
32. Olsen, D.R., Partridge, B., Lynn, S. 2010. Time warp sports for internet television. *ACM TOCHI* 17(4), Art 16
33. Phillips Ambilight, http://www.philips.co.uk/c/televisions/33092/cat/#/hue
34. Philips uWand, http://www.uwand.com/
35. Ren, G., O'Neill, E. 2013. Freehand gestural text entry for interactive TV. In *Proc. of EuroITV '13*, 121-130
36. Ruiz, J., Li, Y., Lank, E. 2011. User-defined motion gestures for mobile interaction. *Proc. of CHI '11*, 197-206
37. Samsung Smart TV: TV Gesture Book, http://www.samsung.com/global/microsite/tv/common/guide_book_5p_vi/waving.html
38. Vatavu, R.D. 2013. A Comparative Study of User-Defined Handheld vs. Freehand Gestures for Home Entertainment Environments. *Journ. Ambient Intell. Smart Environments 5*(2). IOS Press, 187-211
39. Vatavu, R.D. 2013. There's a World outside Your TV: Exploring Interactions beyond the Physical TV Screen. In *Proc. of EuroITV '13*. ACM, NY, USA, 143-152
40. Vatavu, R.D. 2012. User-defined gestures for free-hand TV control. In *Proc. of EuroiTV '12*, 45-48
41. Vatavu, R.D., Pentiuc, S.G. 2008. Multi-Level Representation of Gesture as Command for Human-Computer Interaction. *Computing and Informatics 27*(6). Slovak Academy, 837-851
42. Wobbrock, J.O., Morris, M.R., Wilson, A.D. 2009. User-defined gestures for surface computing. In *Proc. of CHI '09.* ACM, New York, NY, USA, 1083-1092

43. Wobbrock, J.O., Aung, H.H., Rothrock, B., and Myers, B.A. 2005. Maximizing the guessability of symbolic input. In *Proc. of CHI EA '05*, 1869-1872

44. Wu, H., Wang, J. 2012. User-Defined Body Gestures for TV-based Applications. In *Proc. of ICDH '12*. IEEE Comp. Society, Washington, DC, USA, 415-420

APPENDIX A. COMPLETE SET OF PARTICIPANTS' GESTURE PROPOSALS

We present the full list of gestures proposed by our participants for all the referents. The first highlighted gesture in for each referent was the one that received the highest agreement rate. However, some referents received low consensus, and for these cases we selected one gesture from all proposals as the authors' choice for that referent, which we highlighted as well.

No.	Referent	Participants' gesture proposals
1	**Open**	**Open palm**, hand waving, move hand upward, move fist upward followed by wrist rotation, move hand downward, move palm away from the body, thumbs-up, move hand downward and backward then and forward and downward
2	**Close**	**Close palm**, hand waving, closing into a pinch all (all finger tips touching), move hand downward, fist moving up and down twice, move fist upward followed by wrist rotation, hand performing the "go away" cultural gesture, draw "X", perform click in mid-air with the index finger
3	**Next channel**	**Move hand right to left**, **Move hand left to right**, thumbs-up with moving to the right, "ok" cultural pose with thumb and index fingers, moving the index finger from left to right and then clicking in mid-air. Please note the two options in terms of movement direction (left-to-right and right-to-left) that correspond to two metaphors: moving the viewing window (as it happens with scrolling actions and traditional GUI) and moving the items themselves [31].
4	**Previous channel**	**Move hand left to right**, **Move hand right to left**, thumbs-up with moving 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 [31].
5	**Volume up**	**Move hand upward**, hand in pinch pose expanding fingers, thumbs-up with moving to the right, thumbs up moving upwards twice, move hand upward, rotate imaginary button to the right, draw "+", opening hand from thumb-index pinch, draw triangle pointing up, open palm, draw circle clockwise, hand performing the "go away" cultural gesture
6	**Volume down**	**Move hand downward**, from open palm to index-thumb pinch, thumbs-up with moving to the left, thumbs up moving downwards twice, rotate imaginary button to the left, move hand from left to right, draw triangle pointing down, closing into a pinch all (all finger tips touching), draw circle counter-clockwise, hand performing "come closer" cultural gesture
7	**Volume mute**	**Closing fingers into pinch**, fist followed by extending little finger, open palm, thumbs-down to thumbs-up, draw "X", close fist, open palm to index-thumb pinch, thumb-little finger pinch, open palm facing down move left to right, draw crossed zero, draw circle counter-clockwise, move hand downward, move hand left to right, move hand right to left, move palm away from body
8	**Open menu**	**Draw letter "M"**, draw small "o" small "m", rotate wrist down to up, move hand up, opening hand from thumb-index pinch, perform click in mid-air with the index finger, hand in pinch pose open fingers, hand waving, move hand down.
9	**Hide menu**	**Hand wave**, closed fist open index and little finger, close fist, open hand rotate upward to downward, from open palm to index-thumb pinch, move hand upward, hand performing the "go away" cultural gesture, rotate palm left to right, hand performing "come closer" cultural gesture.
10	**Help**	**Draw letter "H" or symbol "?" in mid-air**, hand performing "come closer" cultural gesture, hand in "peace sign" cultural gesture, wave fingers, finger snapping, move hand toward body, wave, pinch and wave
11	**Yes**	**Thumbs-up hand pose**, draw "check" sign, three fingers down, draw letter "Y", 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 "X", move hand down
13	**Go to favorite channel**	**Show index finger**, thumbs-up, perform "come closer" cultural gesture, pinch between the thumb and little finger, move index finger down and upward, draw "check" sign, open palm, fingers snapping, draw star, thumbs-up rotated to right, move hand downward and backward then and forward and downward
14	**Go to 2nd favorite channel**	**Show two fingers (index and middle)**, thumbs-up rotated 90 degrees to right, thumb ring finger pinch, two fingers up, peace sign left to right, fingers snapping twice, draw digit "2", two fingers up followed by thumbs-up, move hand in front two times
15	**Go to random channel**	**Rotate palm facing down to palm up**, draw circle clockwise, draw circle anti-clockwise, draw circle anti-clockwise for three times, close fingers, open palm rotating down to upward, wave with fingers, perform a click in mid-air with the index finger, perform "so so" cultural gesture, open palm, draw letter "R", hand waving, drawing universal quantifier symbol (∀)
16	**Go to channel #7**	**Draw digit "7"**, various finger configurations to indicate "7" as a preferred channel
17	**Go to channel #27**	**Draw number "27" in mid-air**, various finger configurations to indicate "27" as a preferred channel
18	**Show TV guide**	**Draw letter "G"**, move hand upward, thumbs-up, draw letter "M", move palm in front of the body, move hand downward, perform "so so" cultural gesture, move hand left to right, show index finger, pinch followed by wave, open palm
19	**Show channels list**	**Draw square**, four fingers move down, open palm, draw letters "C" and "L", move hand downward, hand performing "come closer" cultural gesture, wave, move hand upward, pinch followed by wave, perform double click in mid-air with the index finger, move hand toward the body, draw letter "L", draw circle anti-clockwise
20	**Last channel**	**Draw circle anti-clockwise**, index and middle fingers moved to the left, fingers snapping, two fingers up, thumbs-up moving back, hand rotating imaginary button to the left, wave
21	**Open Browser**	**Draw symbol "@"**, move palm downward twice, close fist, draw letter "W" three times as in "WWW", hand performing "go away" cultural gesture, pinch twice, pinch followed by thumb ok, moving index finger from left to right, move hand from right to left, move hand downward, move palm in front of the body, show three fingers.