Perceived Usability, Desirability, and Workload of Mid-Air Gesture Control for Smart TVs

Irina Popovici
MintViz Lab | MANSiD Research Center
University Ștefan cel Mare of Suceava
Suceava 720229, Romania
i.popovici.irina@gmail.com

Radu-Daniel Vatavu
MintViz Lab | MANSiD Research Center
University Ștefan cel Mare of Suceava
Suceava 720229, Romania
radu.vatavu@usm.ro

ABSTRACT
We explore in this work users' perceived workload, desirability, and usability of selecting mid-air targets representing TV menu options anchored to physical loci in 3-D space. Toward this end, we introduce a gesture-based selection technique and a spatial user interface for Smart TVs that consists of shortcuts to TV channels located in mid-air in front of the user's body. Target selection is implemented with pointing and hand gestures recognized with the Myo armband. Ten participants evaluated our prototype and were elicited for feedback. We report empirical results about perceived usability (SUS=77.8), desirability (high frequency of positive connotation words), and workload (TLX=43.9) of our gesture-based selection technique and spatial user interface and discuss future work directions.

Author Keywords
Mid-air gestures; pointing; menu selection; usability; NASA TLX; SUS; desirability; study; Smart TVs; spatial user interfaces; Myo.

ACM Classification Keywords
H.5.2. User interfaces: Haptic I/O; Input devices and strategies; Artificial, augmented, and virtual realities.

INTRODUCTION
Interactive television and Smart TVs have witnessed remarkable progress facilitated by advances and developments in display technology [28,29,44], smart devices and implementation of second-screen TV watching scenarios [13,15,36], connecting the TV to the Internet and to Internet-of-Things (IoT) devices and smart spaces [27, 32,41,53], new user interface designs for television powered by Augmented Reality (AR) technology [19,20, 49], and new human sensing techniques [8,50,51] that enabled rich interactive experiences for users. Controlling the TV set has benefited from all these developments. Prior work has demonstrated diverse input modalities for viewers to operate their TV sets with smart mobile devices, such as tablets and smartphones [5,6,23,25], using voice input [39,40], free hand movements and whole-body gestures [12, 47,52], or by operating augmented TV remote controls [1,46]. Some of these innovative input devices and techniques have already found their way from research laboratories and scientific publications to the industry,

Figure 1: Snapshot of our spatial user interface and prototype for Smart TVs. In this picture, the user changes the current channel by pointing to an active locus in mid-air. Pointing and hand gestures are detected by a Myo device worn on the dominant hand.

being considered by Smart TV manufacturers for their top products [26,34,42].

However, current input techniques for operating the TV set have their shortcomings. For example, voice input may be affected by miss-recognition [18]; gesture input may require a training period for users to learn to execute commands correctly and accurately, and the gesture set design must take into account many usability criteria, such as gestures that are ergonomically easy to perform and that bear an intuitive mapping to the TV functions they execute [47,52]; physical remote controls can get lost or need maintenance, such as the need to periodically cleanse them or change their batteries. Therefore, there still is room for exploring new input devices and techniques to control the Smart TV effectively.

In this paper, we focus on mid-air gesture control for operating the TV set. Within this specific application context, we collect users' feedback to understand usability aspects about target selection from invisible menus anchored in mid-air. Specifically, we investigate users' *perceived usability*, *desirability*, and *workload* of our gesture user interface. Our contributions are as follows:

1. We introduce a gesture-based user interface for television control operated by an invisible menu located in mid-air in front of the user's body. We

present the technical implementation of our prototype (Figure 1) using the Myo armband [24].

2. We present empirical results from a user study with N=10 participants conducted to understand the perceived usability, desirability, and workload of our midair gesture user interface. Our findings reveal a good level of perceived usability (SUS = 77.8), many positive appreciations and feedback, such as a system that is *easy to use*, *friendly*, and *intuitive* (reflected in the 84 words selected by our participants with the Microsoft Desirability Toolkit), but also high physical demand (NASA TLX = 64 on the 100 scale).

RELATED WORK

In this section, we present an overview of previous work on mid-air user interfaces, including gesture-based input, and discuss interactive prototypes that used the Myo armband, while focusing on applications for Smart TVs.

Gesture Input with the Myo Armband

The Myo armband is a wearable gesture input device that reports the electrical activity of forearm muscles and hand orientation and acceleration. With the embedded 9-axis IMU, Myo can be used to recognize 3-D gestures in mid-air, while the electromyography measurements (EMG) at forearm level are used to detect hand poses and gestures. The default setup of Myo can recognize five gesture types: *double tap*, *fingers spread*, *wave right*, *wave left*, and *fist*; see Thalmic Labs [24].

Myo has found applications in a variety of domains, ranging from healthcare [22,33,43] to virtual and augmented reality [33,45] and gesture user interfaces [11,21]. In healthcare, for instance, one immediate use of Myo is to collect patient data. In this direction, Koskimäki *et al*. [22] developed "MyoGym," an application for monitoring user activity during training. MyoGym was validated with a controlled user study that evaluated 10 participants performing 30 gym exercises. Another use for Myo in healthcare has been to record and report electromyography readings as an alternative to expensive clinical equipment. The system developed by Tabor *et al*. [43] implemented EMG recording and analysis to help patients train their muscles to accommodate easier to prosthetics. Training was achieved with a survival style game, "The Falling of Momo," where the user had to navigate a monkey on moving platforms and avoid dangers along the way. To make the monkey advance, muscle activation was required, sensed through Myo's array of electrodes. The assistive system developed by Munroe *et al*. [33] for children with cerebral palsy also employed a computer game: children performed squeeze gestures on objects displayed on AR glasses. Other applications of Myo include virtual reality or enhancing human performance. For example, Tsai *et al*. [45] preferred the Myo armband for their virtual reality system over other gesture input devices, such as Leap Motion [17], as Myo proved more flexible in terms of the input space in which users performed gestures. Dalmazzo *et al*. [11] used Myo to record the electrical activity exerted by the left hand of a user playing the violin. Based on data from Myo, machine learning models were implemented to help users learn and perfect their violin playing skills.

Kerber *et al*. [21] conducted an experiment to evaluate the recognition accuracy of the five default gestures provided by Myo's software development kit, which was reported at 68%. The authors proposed an improved recognition algorithm and extended the original set to a total of 40 gestures, for which they reported a recognition accuracy rate of 95% [21].

Gesture User Interfaces for Smart TVs

Gesture-based user interfaces for controlling the TV set proposed in prior work addressed a wide range of TV functions to control, from standard tasks, such as changing channels and adjusting the audio volume [47,52] to operating complex functions specific to multi-screen television systems [46,49].

Bailly *et al*. [1] observed that users perform yaw and pitch movements naturally when operating the TV remote control and re-purposed such movements into actual commands with their "gesture-aware" remote control. Other studies also focused on enhancing or even replacing the standard TV remote. Devices such as the Wii Nintendo remote [46, 47,49], the Microsoft Kinect sensor [47,50], or the Leap Motion controller [51] have been examined to design new, augmented TV remotes that can sense users' actions, gesture commands included. For example, the "RemoteTouch" system of Choi *et al*. [9] employed a touchpad; Vatavu [48,49] used the Wii remote control; and the elicitation study of Zaiţi *et al*. [52,54] reported users' preferences for mid-air hand gestures performed with the Leap Motion controller.

New user interfaces and input devices have been proposed for multi-screen television [48,49]. For example, the Nintendo Wii controller was re-purposed for detecting pointing movements and recognizing motion gestures to enable users to control multiple TV screens [49]. Plaumann *et al*. [38] examined mid-air interactions for multiple users to control the TV at once. Results showed that gestures performed by multiple users may cause system contradictions that, when handled improperly, lead to suboptimal user experience.

A MID-AIR GESTURE USER INTERFACE FOR TV

We designed and implemented a gesture-based user interface for controlling a Smart TV using the Myo armband [24].

Apparatus and Development Tools

The user interface was implemented using HTML 5, CSS 3, and JavaScript 1.7, and was tested under Google Chrome (v66.0.3359.139) on a laptop PC connected to a large, 55-inch Smart TV (Samsung UE55D). The communication between the user interface running in the web browser and the Myo armband was implemented with Myo's JavaScript SDK available to developers from the Myo web page [24]. For the purposes of our evaluation to collect user feedback for our gesture-based user interface, we simulated television control by playing video content streamed from YouTube using HTML 5 controls and JavaScript API.

User Interface Design and Implementation

Our interface enables control of TV channels by detecting pointing and hand gestures. Pointing is performed to physical loci in mid-air that constitute an invisible menu of TV channels located in front of the user's body. For our study, we designed the spatial menu with nine options or shortcuts to TV channels, which we positioned in space following a 3×3 matrix-like arrangement. The number of channels was informed by the upper limit of Miller's "magical" number 7 ± 2 that reflects "*the span of absolute judgment and the span of immediate memory [that] impose severe limitations on the amount of information that we are able to receive, process, and remember*"; see Miller [31]. The matrix arrangement was chosen to mimic the placement of numerical keys on familiar devices, such as TV remote controls or the T9 keyboard. Channel locations were registered before the experiment.

The user interface shown on the TV screen displays the current channel in the top left corner. To implement selection of a new channel, we required two interactive gestures, for which we chose Myo's *double tap* and *fist* due to their ease of execution; see Figure 2 for visual illustrations of these gestures. When the user performs the *double-tap* gesture, the "search for channels" mode becomes active. In search mode, the orientation of the hand in front of the body is used to locate channels mapped to physical loci in space. The closest channel in 3-D is identified and relevant information is displayed at the bottom part of the TV screen: the channel's number, name, and a short description of its contents; see Figure 1. Performing the *fist* gesture confirms the selection.

Pointing with the Myo Armband

Myo reports its orientation in the form of a unit quaternion $q_t = (w_t, x_t, y_t, z_t) \in [0,1]^4$. To use Myo in pointing mode, quaternions need to be corrected by applying an offset with respect to a known, fixed location in space. The offset $q_{offset} = (w_{offset}, x_{offset}, y_{offset}, z_{offset})$ is user-dependent and we determined it during a short calibration phase by asking each user to point their arm towards the TV. To apply the offset, the quaternion q_t reported by Myo at time t is multiplied with q_{offset}, which corresponds to the rotation $q_r = (w_r, x_r, y_r, z_r)$ between q_t and q_{offset}, as follows:[1]

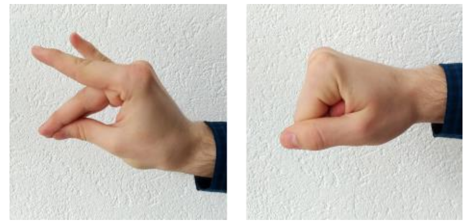

Figure 2: The two Myo gestures [24] used in our prototype: *double-tap* (left) and the *fist* gesture (right).

$$\begin{aligned} w_r &= w_{offset} \cdot w_t - x_{offset} \cdot x_t - y_{offset} \cdot y_t - z_{offset} \cdot z_t \\ x_r &= w_{offset} \cdot x_t + x_{offset} \cdot w_t + y_{offset} \cdot z_t - z_{offset} \cdot y_t \\ y_r &= w_{offset} \cdot y_t - x_{offset} \cdot z_t + y_{offset} \cdot w_t + z_{offset} \cdot x_t \\ z_r &= w_{offset} \cdot z_t + x_{offset} \cdot y_t - y_{offset} \cdot x_t + z_{offset} \cdot w_t \end{aligned} \tag{1}$$

While in search mode, the orientation of the hand is used to identify the most likely channel in 3-D to which the user is pointing. To this end, we compute a measure of distance between the orientation of the hand (q_{hand}) and the quaternion corresponding to a given channel location/orientation in space (q_{ch}) using the following formula [16]:

$$d(q_{hand}, q_{ch}) = 1 - \langle q_{hand}, q_{ch} \rangle^2 \tag{2}$$

where $\langle q_{hand}, q_{ch} \rangle$ denotes the inner product:

$$\langle q_{hand}, q_{ch} \rangle = w_{hand} \cdot w_{ch} + x_{hand} \cdot x_{ch} + y_{hand} \cdot y_{ch} + z_{hand} \cdot z_{ch}$$

The result is converted to an angle measurement:

$$\theta(q_{hand}, q_{ch}) = acos\left(1 - 2\left(1 - \langle q_{hand}, q_{ch} \rangle^2\right)\right) \tag{3}$$

The following pseudocode describes the Nearest-Neighbor classification algorithm that we use to identify the TV channel pointed by the user in search mode. The pseudocode assumes that qchannels[..] represents an array of quaternions corresponding to the orientations of the TV channels in the vertical plane in front of the user's body, while q describes the orientation of the hand at time t.

```
function findChannelPointedByHand(Quaternion q):
    minimum = maxfloat
    selectedChannel = na
    for i, channels, i += 1:
        d = quaternionDistance(q, qchannels[i])
        if d < minimum:
            minimum = d
            selectedChannel = qchannels[i]
        endif
    endfor
    return (selectedChannel)
```

[1] See https://developer.thalmic.com/docs/api_reference/platform/classmyo_ 1_1_quaternion.html for details

USER STUDY

We conducted a controlled experiment to understand the perceived usability, desirability, and workload of our mid-air gesture-based user interface for Smart TVs.

Participants

Ten (10) participants (5 male, 5 female) volunteered for the experiment, ages between 22 and 28 years old (M=24.6, SD=2.3 years). Nine participants were right handed. Participants reported spending on average 2.4 hours on a daily basis (SD=0.95 hours) watching TV or streaming online videos.

Task

The experiment started with an accommodation phase, during which participants followed a tutorial on how to use Myo and practiced the *double tap* and *fist* gestures. The next step consisted in trying out our prototype, accommodating to the user interface and the matrix-like arrangement of TV channels in mid-air. Once participants were confident about their performance, they were asked to change channels by following the instructions of the experimenter, e.g., "*Please change the TV channel to channel #7*." The order of channels was randomized across participants. At the end of the experiment, participants were asked to fill in a questionnaire with demographic information, take a NASA TLX test [35] using an on-line version,[2] fill in the SUS questionnaire [7], and select as many words as they wished from a Microsoft Desirability Toolkit sheet [4] to describe their experience with our gesture-based spatial user interface for Smart TVs.

RESULTS

In this section, we report our participants' subjective perceptions regarding usability, desirability, and workload.

Usability

The SUS tool [7] consists of ten 5-point Likert scale questions, for which the answers are aggregated into a score from 0 to 100, with 100 denoting a perfect usability result. The average SUS score for our prototype was 77.8 (SE = 5.3).[3] Based on prior work that analyzed the distribution of SUS scores [2,3], our result is above average near the "good" threshold of 70 suggested by Bangor *et al*. [2] and it falls inside the "high acceptability" range, according to the scale of Bangor *et al*. [3]. As participants experienced problems with their gestures not always detected by Myo, we have reasons to believe that the SUS score is likely to increase with more accurate recognition algorithms [21].

Desirability

Our participants selected an average of 8.4 words (SE = 1.4) from the Microsoft Desirability Toolkit sheets [4] to describe their experience with our prototype.

Figure 3: Word cloud generated using the 84 words selected by our participants with the Microsoft Desirability Toolkit [4] to describe their experience with our prototype. *Note*: world cloud generated with https://worditout.com**.**

Figure 3 shows a word cloud generated using all the 84 words selected by our participants. We found that the words with the highest frequency were positive, such as *easy to use* (frequency = 6, more than half of the participants considered the user interface to be easy to use), *friendly* (f=5), *attractive* (f=5), *satisfying* (f=4), *intuitive* (f=4), *creative* (f=4), and *innovative* (f=4).

Perceived Workload

The NASA TLX test measures the subjective rating of perceived workload on a scale from 1 (low perceived workload) to 100 (high workload) along six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Figure 4 shows the perceived workload, averaged across all participants, for each dimension.

The mental demand scale reports on the amount of cognitive effort required to use our prototype. Results showed a mental demand of 42 (SE=6.8), showing a medium level of cognitive effort to locate and select TV channels in space. This result is explained by the fact that channels were arranged in a matrix-like structure, easy to visualize and operate in space. Also, the experiment included an accommodation phase, during which participants practiced Myo gestures and selected TV channels. We estimate the value of the mental demand to decrease with further practice.

Physical demand measures the amount of physical effort required to use our prototype, for which we found an average value of 64 (SE=7.2). This high level of physical demand is a direct effect of using the Myo armband in our prototype, which required firm gestures to deliver accurate recognition results from electromyography measurements. As reported by Kerber *et al*. [21], the standard gesture recognition algorithm delivered with the Myo's software development kit [24] has an accuracy rate of just 68%. Therefore, there were times when our participants had to perform a gesture repeatedly until it was detected. We believe that this aspect affected the perceived level of physical demand. This result was complemented by a

[2]http://www.keithv.com/software/nasatlx/

[3]We report standard error values (SE) instead of the standard deviation of the sample (SD) to show how far our sample mean, 77.8 in this case, is likely to be from the population mean.

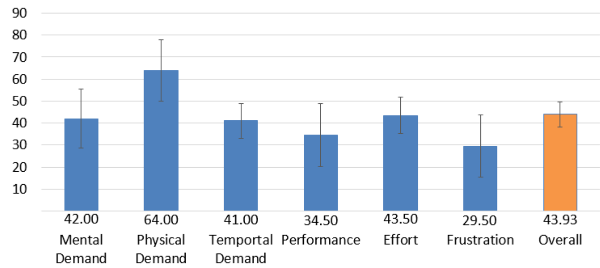

Figure 4: Mean results for each dimension of the NASA TLX test. *Note*: error bars show 95% CIs.

frustration level of 29.5 (SE=7.2) and a perceived effort of 43.5 (SE=4.3) − the effort scale indicates how hard participants had to work on both mental and physical levels to accomplish the tasks of the experiment.

Temporal demand evaluates the pressure that participants felt during the experiment. On average, temporal demand was 41 (SE=4.1). During the experiment, participants were asked to change channels as fast as possible to stimulate performance. Thus, the NASA TLX test reflected the pressure induced by the experimental setup, which we expect to be much less during normal everyday use. As participants had to perform firm gestures for proper recognition, this fact led to the perception of high effort.

The performance scale measures how successful participants thought they were with our prototype. Overall, our participants scored an average of 34.5 (SE=7.4). The reason for a low perceived performance may be due to Myo not detecting hand gestures from time to time.

CONCLUSION AND FUTURE WORK

We presented empirical results on the perceived usability, desirability, and workload of a gesture-based user interface for Smart TVs consisting in pointing to active loci in mid-air. The magnitude of the SUS usability measure and the high frequency of positive words (the Microsoft Desirability Toolkit) employed by participants to describe their experience with our prototype showed good usability and high desirability, despite medium to high workload indicated by the TLX test.

These results recommend future work directions. Firstly, our sample of participants was too small to run statistical tests,[4] such as to understand the effect of gender or spatial orientation skills on user performance. Secondly, other gesture sensing devices might alleviate the problems of Myo not detecting gestures effectively. Examples include the Microsoft Kinect sensor [30,47,50], the Leap Motion controller [17,52,54], or wearable devices, such as smartwatches [10], smart rings [14], or networks or body sensors [37]. Another direction of interesting work is understanding users' preferences for customization of mid-air menus for Smart TVs with both custom gestures and custom channel locations in 3-D space.

[4] For example, it would have been interesting to contrast the performance of men and women with our user interface, or of different groups of participants with different spatial orientation skills, but the power of the t-test was estimated at about 20% for a δ -level of 0.8.

ACKNOWLEDGEMENTS

This work was supported by project no. 101BM/2017 funded by UEFISCDI, Romania and WBI, Belgium.

REFERENCES

1. Gilles Bailly, Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2011. Gesture-aware Remote Controls: Guidelines and Interaction Technique. *In Proceedings of the 13th International Conference on Multimodal Interfaces (ICMI '11).* ACM, New York, NY, USA, 263–270. https://doi.org/10.1145/2070481.2070530
2. Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *J. Usability Studies* 4, 3 (May 2009), 114–123. http://dl.acm.org/citation.cfm?id=2835587.2835589
3. Aaron Bangor, Philip T. Kortum, and James T. Miller. 2008. An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction* 24, 6(2008), 574–594. https://doi.org/10.1080/10447310802205776
4. Joey Benedek and Trish Miner. 2002. Measuring Desirability: New methods for evaluating desirability in a usability lab setting. (2002). http://www.microsoft.com/usability/uepostings/desirabilitytoolkit.doc
5. Regina Bernhaupt and Michael M. Pirker. 2014. User Interface Guidelines for the Control of Interactive Television Systems via Smart Phone Applications. *Behav. Inf. Technol*. 33, 8 (Aug. 2014), 784–799. https://doi.org/10.1080/0144929X.2013.810782
6. Jan Bobeth, Johann Schrammel, Stephanie Deutsch, Michael Klein, Mario Drobics, Christina Hochleitner, and Manfred Tscheligi. 2014. Tablet, Gestures, Remote Control?: Influence of Age on Performance and User Experience with iTV Applications. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14).* ACM, New York, NY, USA, 139–146. https://doi.org/10.1145/2602299.2602315
7. John Brooke. 1996. *SUS-A quick and dirty usability scale. Usability evaluation in industry*. CRC Press. https://www.crcpress.com/product/isbn/9780748404605
8. Andy Brown, Rhia Jones, Mike Crabb, James Sandford, Matthew Brooks, Mike Armstrong, and Caroline Jay. 2015. Dynamic Subtitles: The User Experience. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15)*. ACM, New York, NY, USA, 103–112. https://doi.org/10.1145/2745197.2745204

9. Sangwon Choi, Jaehyun Han, Geehyuk Lee, Narae Lee, and Woohun Lee. 2011. RemoteTouch: Touch-screen-like Interaction in the Tv Viewing Environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 393–402. https://doi.org/10.1145/1978942.1978999
10. Petru-Vasile Cioată and Radu-Daniel Vatavu. 2018. In Tandem: Exploring Interactive Opportunities for Dual Input and Output on Two Smartwatches. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion (IUI '18 Companion)*. ACM, New York, NY, USA, Article 60, 2 pages. https://doi.org/10.1145/3180308.3180369
11. David Dalmazzo and Rafael Ramirez. 2017. Air Violin: A Machine Learning Approach to Fingering Gesture Recognition. In *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education (MIE 2017).* ACM, New York, NY, USA, 63–66. https://doi.org/10.1145/3139513.3139526
12. Nem Khan Dim, Chaklam Silpasuwanchai, Sayan Sarcar, and Xiangshi Ren. 2016. Designing Mid-Air TV Gestures for Blind People Using User-and Choice-Based Elicitation Approaches. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, New York, NY, USA, 204–214. https://doi.org/10.1145/2901790.2901834
13. David Geerts, Rinze Leenheer, Dirk De Grooff, Joost Negenman, and Susanne Heijstraten. 2014. In Front of and Behind the Second Screen: Viewer and Producer Perspectives on a Companion App. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14)*. ACM, New York, NY, USA, 95–102. https://doi.org/10.1145/2602299.2602312
14. Bogdan-Florin Gheran, Jean Vanderdonckt, and Radu-Daniel Vatavu. 2018. Gestures for Smart Rings: Empirical Results, Insights, and Design Implications. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 623–635. https: //doi.org/10.1145/3196709.3196741
15. Christian Holz, Frank Bentley, Karen Church, and Mitesh Patel. 2015. "I'M Just on My Phone and They'Re Watching TV": Quantifying Mobile Device Use While Watching Television. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15).* ACM, New York, NY, USA, 93–102. https://doi.org/10.1145/2745197.2745210
16. Du Q. Huynh. 2009. Metrics for 3D Rotations: Comparison and Analysis. *Journal of Mathematical Imaging and Vision* 35, 2 (Oct 2009), 155–164. https://doi.org/10.1007/s10851-009-0161-2
17. Leap Motion Inc. 2018. Leap Motion. https://www.leapmotion.com/
18. Jiepu Jiang, Wei Jeng, and Daqing He. 2013. How Do Users Respond to Voice Input Errors?: Lexical and Phonetic Query Reformulation in Voice Search. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13).* ACM, New York, NY, USA, 143–152. https://doi.org/10.1145/2484028.2484092
19. Brett Jones, Rajinder Sodhi, Michael Murdock, Ravish Mehra, Hrvoje Benko, Andrew Wilson, Eyal Ofek, Blair MacIntyre, Nikunj Raghuvanshi, and Lior Shapira. 2014. RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-camera Units. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14).* ACM, New York, NY, USA, 637–644. https://doi.org/10.1145/2642918.2647383
20. Brett R. Jones, Hrvoje Benko, Eyal Ofek, and Andrew D. Wilson. 2013. IllumiRoom: Peripheral Projected Illusions for Interactive Experiences. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13).* ACM, New York, NY, USA, 869–878. https://doi.org/10.1145/2470654.2466112
21. Frederic Kerber, Michael Puhl, and Antonio Krüger. 2017. User-independent Real-time Hand Gesture Recognition Based on Surface Electromyography. *In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17).* ACM, New York, NY, USA, Article 36, 7 pages. https://doi.org/10.1145/3098279.3098553
22. Heli Koskimäki, Pekka Siirtola, and Juha Röning. 2017. MyoGym: Introducing an Open Gym Data Set for Activity Recognition Collected Using Myo Armband. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers (UbiComp '17).* ACM, New York, NY, USA, 537–546. https://doi.org/10.1145/3123024.3124400
23. Artur Henrique Kronbauer, Fabio Fonseca Gomes, and Brunno Brito Araujo. 2015. Remote Home: A Universal Control for Residential Environments. In *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web (WebMedia '15).* ACM, New York, NY, USA, 213–216. https://doi.org/10.1145/2820426.2820451
24. Thalmic Labs. 2018. Myo Gesture Control Armband | Wearable Technology by Thalmic Labs. https://www.myo.com/
25. Dong-Hoon Lee, Jung-Hyun Kim, Ho-Youn Kim, and Dong-Young Park. 2016. Remote Application Control

Technology and Implementation of HTML5-based Smart TV Platform. In *Proceedings of the 14th International Conference on Advances in Mobile Computing and Multi Media (MoMM '16).* ACM, New York, NY, USA, 208–211. https://doi.org/10.1145/3007120.3007159

26. LG. 2018. LG Magic Remove: TV Remotes for Smart TVs. http://www.lg.com/us/magic-remote

27. Yihua Lou, Wenjun Wu, Radu-Daniel Vatavu, and Wei-Tek Tsai. 2017. Personalized Gesture Interactions for Cyber-Physical Smart-Home Environments. *Science China Information Sciences* 60, 7 (2017), 072104:1–15. http://dx.doi.org/10.1007/s11432-015-1014-7

28. Shao-Ping Lu, Ruxandra-Marina Florea, Pablo Cesar, Peter Schelkens, and Adrian Munteanu. 2017. Efficient Depth-aware Image Deformation Adaptation for Curved Screen Displays. In *Proceedings of the on Thematic Workshops of ACM Multimedia 2017 (Thematic Workshops '17).* ACM, New York, NY, USA, 442–450. https://doi.org/10.1145/3126686.3126692

29. Mark McGill, John H. Williamson, and Stephen Brewster. 2016. Examining The Role of Smart TVs and VR HMDs in Synchronous Ata-Distance Media Consumption. *ACM Transactions on ComputerHuman Interaction* 23, 5, Article 33 (Nov. 2016), 57 pages. https://doi.org/10.1145/2983530

30. Microsoft. 2018. Kinect - Windows app development. https://developer.microsoft.com/en-us/windows/kinect

31. George A. Miller. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63, 2 (1956), 81. http://psycnet.apa.org/doi/10.1037/h0043158

32. Michael Miller. 2015. *The Internet of Things: How Smart TVs, Smart Cars, Smart Homes, and Smart Cities Are Changing the World.* Que Publishing, USA.

33. Christopher Munroe, Yuanliang Meng, Holly Yanco, and Momotaz Begum. 2016. Augmented Reality Eyeglasses for Promoting HomeBased Rehabilitation for Children with Cerebral Palsy. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction (HRI '16).* IEEE Press, Piscataway, NJ, USA, 565–565. http://dl.acm.org/citation.cfm?id=2906831.2906981

34. Kevin Murnane. 2018. Google's Assistant Has Found A New Home In LG's OLED And Super UHD TVs. https://www.forbes.com/sites/kevinmurnane/2018/01/03/googles-assistant-has-found-a-new-home-in-lgs-oled-and-super-uhd-tvs

35. NASA. 2018. NASA TLX | Task Load Index. https://humansystems.arc.nasa.gov/groups/tlx/

36. Timothy Neate, Michael Evans, and Matt Jones. 2017. Enhancing Interaction with Dual-Screen Television Through Display Commonalities. In *Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '17).* ACM, New York, NY, USA, 91–103. https://doi.org/10.1145/3077548.3077549

37. Noitom. 2018. Perception Neuron by Noitom. https://neuronmocap.com/

38. Katrin Plaumann, David Lehr, and Enrico Rukzio. 2016. Who Has the Force?: Solving Conflicts for Multi User Mid-Air Gestures for TVs. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '16).* ACM, New York, NY, USA, 25–29. https://doi.org/10.1145/2932206.2932208

39. Jinfeng Rao, Ferhan Ture, Hua He, Oliver Jojic, and Jimmy Lin. 2017. Talking to Your TV: Context-Aware Voice Search with Hierarchical Recurrent Neural Networks. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17).* ACM, New York, NY, USA, 557–566. https://doi.org/10.1145/3132847.3132893

40. Bernard Renger, Junlan Feng, Ovidiu Dan, Harry Chang, and Luciano Barbosa. 2011. VoiSTV: Voice-enabled Social TV. *In Proceedings of the 20th International Conference Companion on World Wide Web (WWW '11).* ACM, New York, NY, USA, 253–256. https://doi.org/10.1145/1963192.1963302

41. Richard L. Rutledge, Aaron K. Massey, and Annie I. Antón. 2016. Privacy Impacts of IoT Devices: A SmartTV Case Study. In *Proceedings of the 24th IEEE International Requirements Engineering Conference Workshops (REW '16).* 261–270. https://doi.org/10.1109/REW.2016.050

42. Samsung. 2018. Samsung Smart TV: TV Gesture Book. http://www.samsung.com/ph/smarttv/common/guide_book_3p_si/waving.html

43. Aaron Tabor, Scott Bateman, Erik Scheme, David R. Flatla, and Kathrin Gerling. 2017. Designing Game-Based Myoelectric Prosthesis Training. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17).* ACM, New York, NY, USA, 1352–1363. https://doi.org/10.1145/3025453.3025676

44. Yutaka Tokuda, Mohd Adili Norasikin, Sriram Subramanian, and Diego Martinez Plasencia. 2017. MistForm: Adaptive Shape Changing Fog Screens. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17).* ACM, New York, NY, USA, 4383–4395. https://doi.org/10.1145/3025453.3025608

45. Wan-Lun Tsai, You-Lun Hsu, Chi-Po Lin, Chen-Yu Zhu, Yu-Cheng Chen, and Min-Chun Hu. 2016. Immersive Virtual Reality with Multimodal Interaction and Streaming Technology. In *Proceedings of the 18th ACM International Conference on Multimodal*

Interaction (ICMI 2016). ACM, New York, NY, USA, 416–416. https://doi.org/10.1145/2993148.2998526

46. Radu-Daniel Vatavu. 2012. Point & Click Mediated Interactions for Large Home Entertainment Displays. *Multimedia Tools and Applications* 59, 1 (2012), 113–128. https://doi.org/10.1007/s11042-010-0698-5

47. Radu-Daniel Vatavu. 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, 45–48. https://doi.org/10.1145/2325616.2325626

48. Radu-Daniel Vatavu. 2013. A Comparative Study of User-Defined Handheld vs. Freehand Gestures for Home Entertainment Environments. *Journal of Ambient Intelligence and Smart Environments* 5, 2(2013), 187–211. http://dx.doi.org/10.3233/AIS-130200

49. Radu-Daniel Vatavu. 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, 143–152. https://doi.org/10.1145/2465958.2465972

50. Radu-Daniel Vatavu. 2015. Audience Silhouettes: Peripheral Awareness of Synchronous Audience Kinesics for Social Television. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '15).* ACM, New York, NY, USA, 13–22. https://doi.org/10.1145/2745197.2745207

51. Radu-Daniel Vatavu and Matei Mancas. 2014. Visual Attention Measures for Multi-screen TV. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14).* ACM, New York, NY, USA, 111–118. https://doi.org/10.1145/2602299.2602305

52. Radu-Daniel Vatavu and Ionuţ-Alexandru Zaiţi. 2014. Leap Gestures for TV: Insights from an Elicitation Study. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video (TVX '14).* ACM, New York, NY, USA, 131–138. https://doi.org/10.1145/2602299.2602316

53. Murad Yusufov and Ivan Kornilov. 2013. Roles of Smart TV in IoTenvironments: A survey. In *Proceedings of the 13th Conference of Open Innovations Association (FRUCT '13).* 163–168. https://doi.org/10.23919/FRUCT.2013.8124240

54. Ionuţ-Alexandru Zaiţi, Ştefan Gheorghe Pentiuc, and Radu-Daniel Vatavu. 2015. On Free-Hand TV Control: Experimental Results on User-Elicited Gestures with Leap Motion. *Personal and Ubiquitous Computing* 19, 5-6 (2015), 821–838. http://dx.doi.org/10.1007/s00779-015-0863-y