Scenario-based Exploration of Integrating Radar Sensing into Everyday Objects for Free-Hand Television Control

Alexandru-Ionuţ Şiean
MintViz Lab, MANSiD Center
Ştefan cel Mare University of Suceava
Suceava, Romania
alexandru.siean@usm.ro

Cristian Pamparău
MintViz Lab, MANSiD Center
Ştefan cel Mare University of Suceava
Suceava, Romania
cristian.pamparau@usm.ro

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

ABSTRACT

We address gesture input for TV control, for which we examine mid-air free-hand interactions that can be detected via radar sensing. We adopt a scenario-based design approach to explore possible locations from the living room where to integrate radar sensors, e.g., in the TV set, the couch armrest, or the user's smartphone, and we contribute a four-level taxonomy of locations relative to the TV set, the user, personal robot assistants, and the living room environment, respectively. We also present preliminary results about an interactive system using a 15-antenna ultra-wideband 3D radar, for which we implemented a dictionary of six directional swipe gestures for the control of dichotomous TV system functions.

CCS CONCEPTS

• **Human-centered computing** → **Gestural input**; **Interface design prototyping**; **Scenario-based design**.

KEYWORDS

Gesture input, free-hand input, mid-air gestures, TV, remote control

ACM Reference Format:
Alexandru-Ionuţ Şiean, Cristian Pamparău, and Radu-Daniel Vatavu. 2022. Scenario-based Exploration of Integrating Radar Sensing into Everyday Objects for Free-Hand Television Control. In *ACM International Conference on Interactive Media Experiences (IMX '22), June 22–24, 2022, Aveiro, JB, Portugal.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3505284.3532982

1 INTRODUCTION

Smart TVs offer a wide range of functionality to enhance home entertainment experiences in the living room, from access to online content [20,30] to augmented and mixed reality television [33,48] to cross-device input and second-screen TV watching [14] to employing the interactive TV as a medium for applications of various kinds [13,21,47]. Beyond the conventional TV remote control, technical solutions for interacting with the TV set have included smartphones [25], video game controllers [42], augmented remote controls [6], and gesture input using a wide range of sensing technology [17,38,43,46,53]. Among these input modalities and devices, we focus on gesture input due to its many advantages, such as intuitiveness, flexibility, versatility, and shared control over the TV for multi-user interaction. Specifically, we focus on free-hand gestures that can be detected by radars, for which a growing interest has been manifested in the HCI community [52].

Radar sensors provide many opportunities to detect gestures for interactive applications. Radars are available for mobile devices [22, 27], integration with work surfaces [4], and clothes [26], among others. It is worthwhile thus to systematically examine possible locations for placing and integrating radar sensing in the living room in the context of TV watching and control. To this end, we adopt the scenario-based design [12,34] approach, and propose a four-level taxonomy of locations from the living room where radar sensors can be placed, dissimulated, affixed, or integrated to enable mid-air gesture interaction with the TV; see Figure 1 on the next page for an overview. This taxonomy can be used to inform the design of interactive applications for the living room centered on the TV that employ radar sensing to implement user input. We also present preliminary results regarding a system implemented using a 15-antenna ultra-wideband 3D radar and a dictionary of six mid-air directional swipe gestures that can be used to control a variety of dichotomous TV system functions.

2 RELATED WORK

We relate to prior work on gesture-based TV control and radar-based gesture sensing and recognition, respectively.

2.1 Gesture-based TV Control

Gesture input for TV control has been largely examined in the context where usability and accessibility problems have been frequently reported for conventional TV remote controls. Specifically, TV remote controls require visual attention, replacing batteries, can get lost, and may be in the possession of other users; see [8,9,16,51] for such examples. Also, their form factors, tiny buttons, and many buttons represent accessibility challenges for users with disabilities, such as people with motor impairments [40]. In this context, technical solutions for gesture input to control the TV have been proposed and evaluated. An early prototype of Freeman and Weissman [17] employed a video camera to detect the location and pose of the viewer's hand for click-like input for the TV. Vatavu and Pentiuc [46] introduced an interaction technique where hand gestures were recognized above a coffee table. Vatavu [43] conducted the first end-user gesture elicitation study to document users' preferences for free-hand gesture control of the TV, which was followed by other elicitation studies, e.g., gestures to control a movable TV [38], gestures preferred by people who are blind [15], users' preferences

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
IMX '22, June 22–24, 2022, Aveiro, JB, Portugal
© 2022 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9212-9/22/06.
https://doi.org/10.1145/3505284.3532982

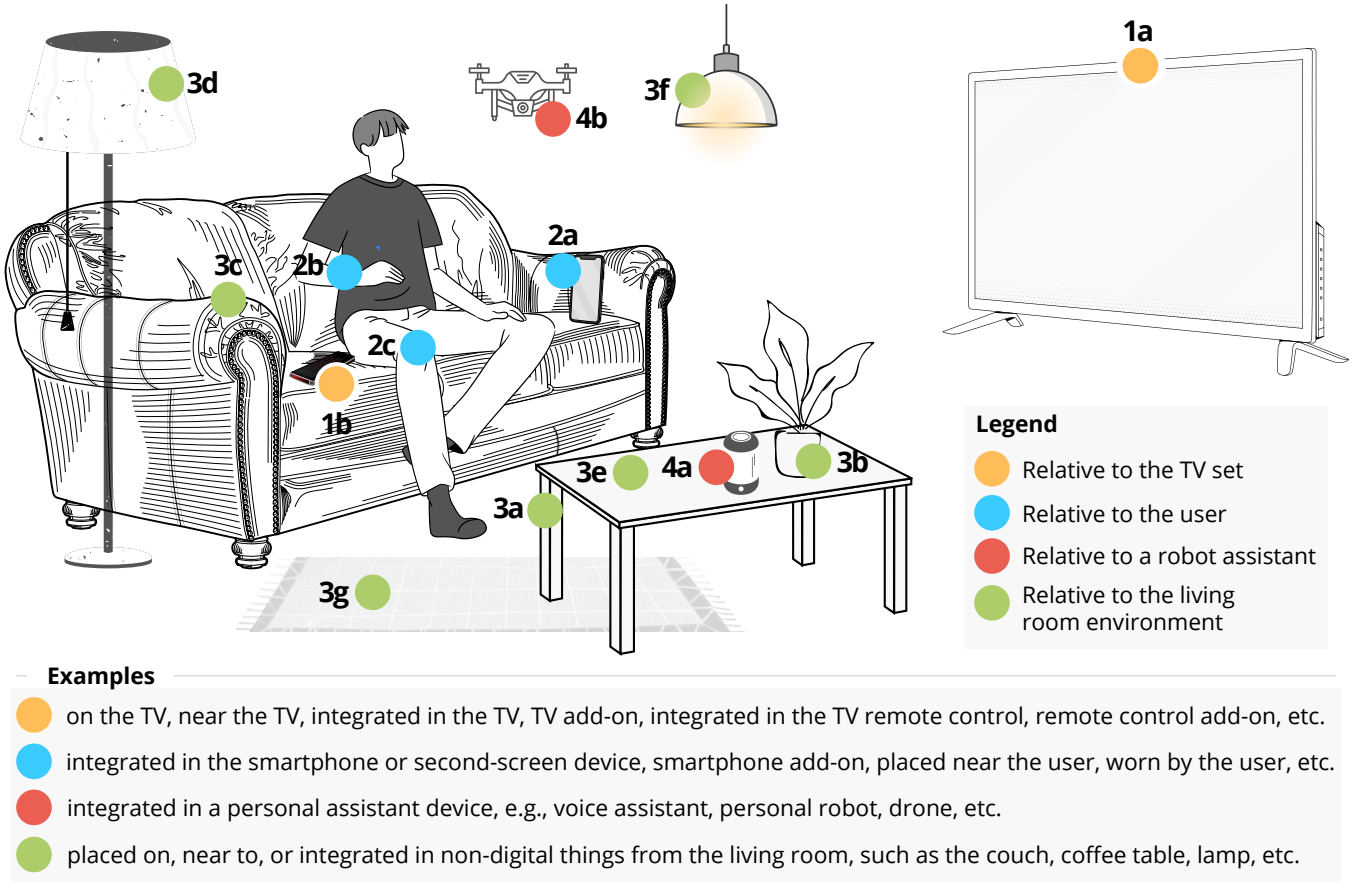

Figure 1: Illustration of various locations for placing, dissimulating, and integrating radar sensing in the living room, structured along four dimensions: locations relative to the TV set, the user, robot assistants, and the living room environment, respectively. For example, location 2b specifies integration of radar sensing into a wearable device, such as a smartwatch, and location 3c specifies integration of radar sensing in the couch armrest.

for bimanual gestures [49], and gestures that allows moderators to control content in Augmented Reality TV [35]. Other prototypes have employed mobile and wearable devices, such as Popovici *et al.* [33], who linked smart pockets to the smart TV with pointing gestures detected by a smart armband with an integrated IMU. These examples demonstrate a variety of sensing technology and corresponding interaction techniques for gesture-based TV control. Next, we focus on gestures detected using radars, which represents one key application of radar sensing in HCI [52].

2.2 Radar-based Gesture Recognition

Radars are convenient for human sensing because they function in environments with high light intensity, low light and darkness, when occluded by a surface or other objects [4,31], or under various weather conditions [52]. Prior work has introduced and evaluated several recognition techniques for radar gestures. For example, mHomeGes [28] is a system that detects gestures with 95.3% accuracy for smart home interactions; Soli [27] is a miniature gesture sensing technology based on millimeter-wave radar that tracks gestures with sub-millimeter accuracy at over 10,000 frames per second; RadarNet [22] is an efficient recognition technique for radar gestures that employs a Convolutional Neural Network and runs efficiently on battery-powered, computationally constrained processors; and Pantomime [31] is a a deep learning architecture delivering 95% accuracy for gestures performed in mid-air. For more examples, we refer to Ahmed *et al.* [2] for a review of hand gesture recognition with radar sensors.

3 A TAXONOMY OF LOCATIONS FROM THE LIVING ROOM FOR RADAR SENSING

We introduce a taxonomy of locations from the living room where a radar sensor can be integrated, dissimulated, affixed to, or simply placed in plain sight to pick up mid-air gestures for TV remote control. To this end, we employ scenario-based design [11], an approach that capitalizes on the flexibility of using scenarios to manage the fluidity of possible design situations, multiple views of the interaction, and flexible amount of detailing [12], respectively. In our case, the scenarios are represented by possible integration

of radar sensing in the living room (which acts as the setting for our scenarios) to enable gesture-based input (i.e., our input modality of choice) to control the TV (i.e., our task). A scenario-based approach is convenient at this stage because it puts the focus on how the system can be used to accomplish the task (e.g., What kind of interactions are possible when the radar is integrated in the TV *vs.* the coffee table *vs.* the user's smartwatch?), rather than on functional specifications of the system (e.g., What recognition techniques are best suited for radar gestures?). To describe such possible design situations and corresponding scenarios, we consider a design space centered on the TV set, the user's body, the living room environment, and objects and devices that transcend the living room, such as autonomous robot companions. By positioning scenarios in this space, we arrive at several possible design solutions for the placement of radar sensors (see Figure 1 for illustrations).

3.1 Radar Placement Relative to the TV Set

In this scenario, the radar is part of the TV set, as follows:

(1a) The radar is integrated in the TV or is connected to the TV, e.g., via one of the available USB ports, as a hardware add-on that extends the built-in functionality of the TV set. The perspective and field of view leveraged by such a placement enable a variety of gesture types performed with the hand and arm in front of the TV set, from hand poses [17] to pointing at the TV screen [42] and the space around it [45] to mid-air gestures [35,53] to arm and whole-body gesture input [38,43]. Such a design solution also enables multi-user interaction, e.g., gestures performed by multiple viewers that share the control of the TV [32].

(1b) The radar is integrated in the TV remote control, which results in new designs of TV remote controls, complementary to prior solutions [6,42], that can sense gesture input. The field of view of the radar sensor does not need to be as wide as in scenario 1a, since the TV remote control is close to the user. Moreover, this design scenario of an augmented TV remote control enables mixed input in the form of buttons and gestures [44], respectively.

3.2 Radar Placement Relative to the User's Personal Computer Devices

In this scenario, the radar sensor is located near the user or integrated in the user's personal computer devices. We identify several design possibilities, as follows:

(2a) The radar is integrated in the smartphone, a prevalent personal device, for which technical solutions are already possible for the integration; see [22,27].

(2b) The radar is integrated in a wearable, such as a smartwatch, smart ring, etc. Unlike design scenario 2a that involves a reasonably large device and corresponding battery power, this option regards devices with various form factors that may impose technical constraints on the size of the radar and the resources it may require for operation. Although we could not find examples in the scientific literature for integrating radars in wearable devices, small radar chips, such as Soli [27], afford such an integration.

(2c) The radar sensor is integrated in the user's clothes, for example in the sleeve, pocket, or in clothing accessories. For example, Leiva *et al.* [26] evaluated radar gesture sensing through wool, cotton, and leather fabrics to support technical implementation of wearable radars.

3.3 Radar Placement Relative to the Living Room Environment

In this scenario, radars are placed in various locations from the living room, including everyday non-digital objects, which turn into remote controls for the TV set. Examples of possible design options at this level of the taxonomy include:

(3a) The radar is located on, under, or is integrated in the coffee table from the living room. This scenario enables mid-air gestures performed with the hand, which mimic use case scenarios demonstrated for TV control [46], but also extension to use feet gestures in the active area of the sensor.

(3b) The radar is integrated in a decorative object from the coffee table. The same types of gestures as in design 3a are possible, except that the object is mobile and, thus, enables flexible placement, orientation, and use. Such a design relates to TV control scenarios involving everyday objects from the living room, such as tangible cubes [10,39] or even plush toys [24].

(3c) The radar is integrated in the couch, e.g., in the armrest, enabling a physical space for gesture articulation that is comfortable for the user [41].

(3d) The radar is integrated in a physical object located in the proximity of the couch, such as the floor lamp illustrated in Figure 1. Depending on the location of the user with respect to the sensor, mid-air gestures of the hand and whole-body gestures may be supported with sensing from that location.

(3e) The radar is integrated in the floor, which offers a distinct perspective on the user's gestures and, correspondingly, the opportunity to employ specific gesture types [5].

(3f) The radar is integrated in the ceiling or in objects close to the ceiling, which offer a complementary perspective and gesture types [18] to the design solution 3e.

3.4 Radar Placement Relative to a Robot Assistant

In this scenario, the radar sensor is placed in an object or device that is not bound to the physical space of the living room environment, e.g., personal robot assistants are relevant examples. We identify the following design options:

(4a) The radar is integrated into a voice assistant, which represents a category of devices that have become widespread for smart homes [3].

(4b) The radar is integrated into a moving robot assistant. Unlike design 4a, the robot assistant automatically follows the user, e.g., the "follow me" functionality implemented by personal drones that respond to gesture input [1].

More examples are enumerated at the bottom of Figure 1. Our taxonomy is useful to identify locations for integrating radar sensing for practical applications in the living room. Next, we present a preliminary system prototype built with a commercial radar.

Figure 2: By combining directional left-right swipes and the distance from the sensor at which the swipes are performed, a set of six mid-air gestures results for effecting dichotomous yes-no, up-down, and next-previous functions common for TV.

4 PRELIMINARY PROTOTYPE, FINDINGS, AND FUTURE WORK

To understand the technical feasibility of the various locations from our taxonomy, we implemented a prototype using the 15-antenna Walabot Creator device and the Walabot API.[1] We acquired the 3D trajectory of the hand detected above the radar, and used it to recognize two directional swipe gestures—*swipe left* and *swipe right*—on the y axis. Also, we leveraged the distance from the sensor on the z axis at which these two gestures are performed to specify three active zones—*near*, *close*, and *far*—above the radar; see Figure 2. A set of six gestures results from the combination of two directions and three zones, which can be mapped to three types of dichotomous functions for TV control: *yes-no* to confirm and reject selections, *up-down* to manipulate the value of a control, such as volume, brightness, etc., and *next-previous* to implement navigation in a list, e.g., the list of TV channels. The radar was placed on a table (Figure 2), which corresponds to the coffee table (3a), decorative object (3b), and the couch armrest (3c) scenarios illustrated in Figure 1. Other locations may need adaptations of our simple gesture recognition pipeline, including special preprocessing of the raw signal and recognition techniques [37].

Figure 3 presents θ–R images obtained from the Walabot radar when placed in various locations corresponding to the scenarios from Figure 1 and various types of occlusion. For example, the radar was occluded in Figures 3b-3e when placed under a table, in a box, under the TV remote control, and in the trousers right front pocket, respectively. Figures 3f-3h illustrate scenarios where the user is at a distance of several meters from the radar. These images suggest that high discriminability is expected from such locations, for which we leave the technical evaluation for future work. Also, other types of radars will result in different types of data. Thus, it is useful to examine as part of future work other characteristics of radar sensors, such as resolution or field of view, needed to implement the various design solutions identified in our taxonomy. For instance, depending on the modulation technique, various types of information can be obtained from a radar, e.g., 1D, continuous-wave modulation separates objects by their velocity, but 3D, frequency-modulated continuous wave multiple input/multiple output modulation separates objects by velocity, distance, and angle. We leave such examinations of radar technology for future work.

Another interesting direction for future work is connecting the locations of our taxonomy with interaction concepts and techniques for smart environments, such as interactions based on proximity [19] or peripheral interactions [7], and to corresponding tools designed to support such interactions [29,36]. Also, placing the radar at the various locations from our taxonomy determines different possibilities for sensing gesture types, from low-scale gestures above the smartphone or the TV remote control to large-scale arm-level gestures performed in front of the TV screen. Thus, connecting our taxonomy of locations to gesture taxonomies and design spaces for gesture input [23,50,53] represents interesting future work.

[1] https://api.walabot.com/sample.html

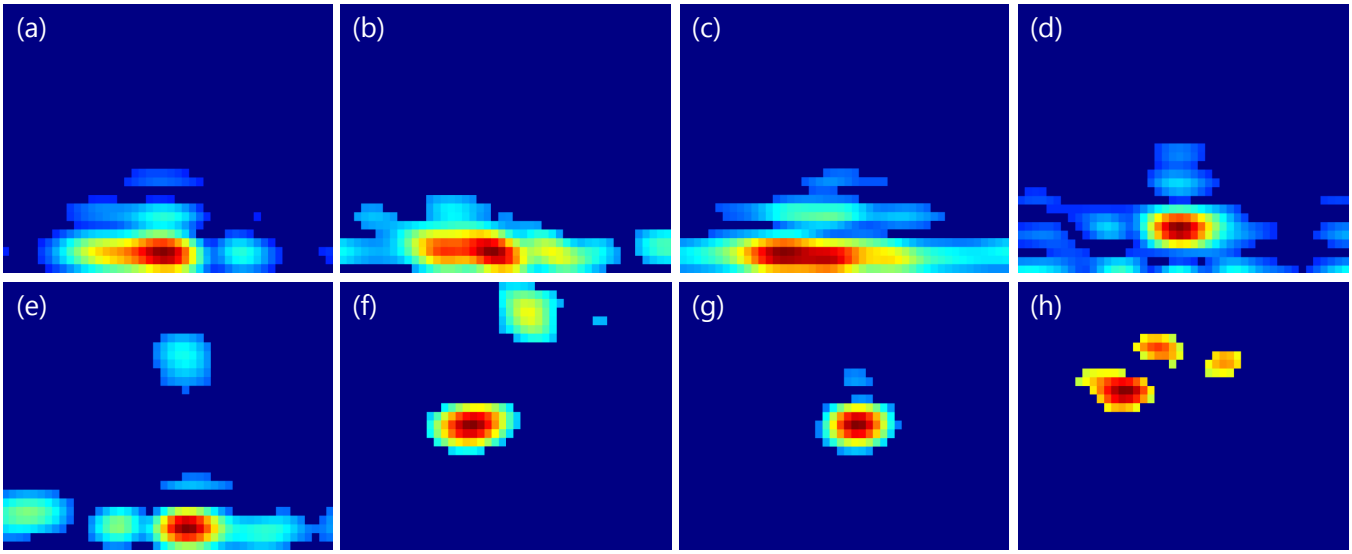

Figure 3: θ–R images acquired under various placements of the radar sensor: (a) radar uncovered, placed on a table, (b) radar under the table, (c) in a box, (d) under the TV remote control, (e) in the pocket, (f) next to the TV set facing the user, (g) on the floor, (h) two meters above the ground simulating placement on a drone. Each image shows the open hand pose.

5 CONCLUSION

We presented in this paper an exploration of possible design solutions for integrating radar sensing in the living room with the goal of implementing gesture-based TV remote control. We capitalized in our taxonomy on the versatility of radar sensing under various conditions, e.g., low light or occlusion, and the variety of gesture types that can be detected with radars. Our preliminary system with a commercial radar sensor showed the feasibility of several such locations, while other technical examinations are left for future work. Also, although we centered our discussion on the TV, our taxonomy of locations for radar sensing is useful to be extended to interactions with other types of digital devices and home appliances from the living room in the context of smart home environments.

ACKNOWLEDGMENTS

This work was supported by grants of the Ministry of Research, Innovation and Digitization, CNCS/CCCDI-UEFISCDI, project numbers PN-III-CEI-BIM-PBE-2020-0001 (1BM/2021) and PN-III-P3-3.6-H2020-2020-0034 (12/2021).

REFERENCES

[1] Parastoo Abtahi, David Y. Zhao, Jane L. E., and James A. Landay. 2017. Drone Near Me: Exploring Touch-Based Human-Drone Interaction. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 34 (sep 2017), 8 pages. https://doi.org/10.1145/3130899

[2] Shahzad Ahmed, Karam Dad Kallu, Sarfaraz Ahmed, and Sung Ho Cho. 2021. Hand Gestures Recognition Using Radar Sensors for Human-Computer-Interaction: A Review. *Remote Sensing* 13, 3, Article 527 (2021), 24 pages. https://doi.org/10.3390/rs13030527

[3] Tawfiq Ammari, Jofish Kaye, Janice Y. Tsai, and Frank Bentley. 2019. Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Trans. Comput.-Hum. Interact.* 26, 3, Article 17 (apr 2019), 28 pages. https://doi.org/10.1145/3311956

[4] Daniel Avrahami, Mitesh Patel, Yusuke Yamaura, and Sven Kratz. 2018. Below the Surface: Unobtrusive Activity Recognition for Work Surfaces Using RF-Radar Sensing. In *Proceedings of the 23rd Int. Conf. on Intelligent User Interfaces (IUI '18).* ACM, New York, NY, USA, 439–451. https://doi.org/10.1145/3172944.3172962

[5] Gilles Bailly, Jörg Müller, Michael Rohs, Daniel Wigdor, and Sven Kratz. 2012. ShoeSense: A New Perspective on Gestural Interaction and Wearable Applications. In *Proceedings of the CHI Conf. on Human Factors in Computing Systems (CHI '12).* ACM, New York, NY, USA, 1239–1248. https://doi.org/10.1145/2207676.2208576

[6] Gilles Bailly, Dong-Bach Vo, Eric Lecolinet, and Yves Guiard. 2011. Gesture-Aware Remote Controls: Guidelines and Interaction Technique. In *Proceedings of the ACM International Conference on Multimodal Interaction (ICMI '11).* ACM, New York, NY, USA, 263–270. https://doi.org/10.1145/2070481.2070530

[7] Saskia Bakker, Elise Hoven, and Berry Eggen. 2015. Peripheral Interaction: Characteristics and Considerations. *Personal Ubiquitous Comput.* 19, 1 (jan 2015), 239–254. https://doi.org/10.1007/s00779-014-0775-2

[8] Regina Bernhaupt, Antoine Desnos, Michael Pirker, and Daniel Schwaiger. 2015. TV Interaction Beyond the Button Press. In *Human-Computer Interaction – INTERACT 2015.* Springer-Verlag, Berlin, Heidelberg, 412–419. https://doi.org/10.1007/978-3-319-22668-2$_3$1

[9] Regina Bernhaupt, Marianna Obrist, Astrid Weiss, Elke Beck, and Manfred Tscheligi. 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. https://doi.org/10.1145/1350843.1350848

[10] Florian Block, Albrecht Schmidt, Nicolas Villar, and Hans W. Gellersen. 2004. Towards a Playful User Interface for Home Entertainment Systems. In *Ambient Intelligence*, Panos Markopoulos, Berry Eggen, Emile Aarts, and James L. Crowley (Eds.). Springer, Berlin, 207–217. https://doi.org/10.1007/978-3-540-30473-9$_2$0

[11] John M. Carroll (Ed.). 1995. *Scenario-Based Design: Envisioning Work and Technology in System Development.* John Wiley & Sons, Inc., USA.

[12] John M. Carroll. 2000. Making Use: Scenarios and Scenario-Based Design. In *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS '00).* ACM, New York, NY, USA, 4. https://doi.org/10.1145/347642.347652

[13] Konstantinos Chorianopoulos and George Lekakos. 2007. Learn and Play with Interactive TV. *Comput. Entertain.* 5, 2, Article 4 (apr 2007), 9 pages. https://doi.org/10.1145/1279540.1279544

[14] Cédric Courtois and Evelien D'heer. 2012. Second Screen Applications and Tablet Users: Constellation, Awareness, Experience, and Interest. In *Proceedings of the 10th European Conference on Interactive TV and Video (EuroITV '12).* ACM, New York, NY, USA, 153–156. https://doi.org/10.1145/2325616.2325646

[15] 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

[16] Douglas Ferguson. 1994. Measurement of Mundane TV Behaviors: Remote Control Device Flipping Frequency. *Journal of Broadcasting & Electronic Media* 38 (01 1994), 35–47. https://doi.org/10.1080/08838159409364244

[17] William T. Freeman and Craig D. Weissman. 1995. Television Control by Hand Gestures. In *Proceedings of the IEEE International Workshop on Automatic Face and Gesture Recognition*. IEEE, Washington, D.C., USA, 179–183.

[18] Kentaro Fukuchi, Toshiki Sato, Haruko Mamiya, and Hideki Koike. 2010. Pac-Pac: Pinching Gesture Recognition for Tabletop Entertainment System. In *Proceedings of the International Conference on Advanced Visual Interfaces (AVI '10)*. ACM, New York, NY, USA, 267–273. https://doi.org/10.1145/1842993.1843040

[19] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino, and Miaosen Wang. 2011. Proxemic Interactions: The New Ubicomp? *Interactions* 18, 1 (jan 2011), 42–50. https://doi.org/10.1145/1897239.1897250

[20] Tam Hanna. 2012. *Apps on TV* (1st ed.). Apress, USA.

[21] Gunnar Harboe, Noel Massey, Crysta Metcalf, David Wheatley, and Guy Romano. 2008. The Uses of Social Television. *Comput. Entertain.* 6, 1, Article 8 (may 2008), 15 pages. https://doi.org/10.1145/1350843.1350851

[22] Eiji Hayashi, Jaime Lien, Nicholas Gillian, Leonardo Giusti, Dave Weber, Jin Yamanaka, Lauren Bedal, and Ivan Poupyrev. 2021. RadarNet: Efficient Gesture Recognition Technique Utilizing a Miniature Radar Sensor. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA, Article 5, 14 pages. https://doi.org/10.1145/3411764.3445367

[23] Maria Karam and m.c. schraefel. 2005. *A Taxonomy of Gestures in Human Computer Interactions*. Technical Report. University of Southampton. https://eprints.soton.ac.uk/261149/

[24] Yoshinori Kawasaki, Takeo Igarashi, Takashi Ajioka, and Ikuji Honda. 2005. Vision-based Gestural Interaction Using Plush Toys. In *Proceedings of ACM UIST '05*. ACM, New York, NY, USA, 1–2. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.107.804

[25] Christine Kühnel, Tilo Westermann, Fabian Hemmert, Sven Kratz, Alexander Müller, and Sebastian Möller. 2011. I'm Home: Defining and Evaluating a Gesture Set for Smart-Home Control. *International Journal of Human-Computer Studies* 69, 11 (2011), 693–704. https://doi.org/10.1016/j.ijhcs.2011.04.005

[26] Luis A. Leiva, Matjaz Kljun, Christian Sandor, and Klen Copic Pucihar. 2020. The Wearable Radar: Sensing Gestures Through Fabrics. In *Proceedings of the 22nd International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '20)*. ACM, New York, NY, USA, Article 17, 4 pages. https://doi.org/10.1145/3406324.3410720

[27] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* 35, 4, Article 142 (jul 2016), 19 pages. https://doi.org/10.1145/2897824.2925953

[28] Haipeng Liu, Yuheng Wang, Anfu Zhou, Hanyue He, Wei Wang, Kunpeng Wang, Peilin Pan, Yixuan Lu, Liang Liu, and Huadong Ma. 2020. Real-Time Arm Gesture Recognition in Smart Home Scenarios via Millimeter Wave Sensing. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 4, Article 140 (dec 2020), 28 pages. https://doi.org/10.1145/3432235

[29] Nicolai Marquardt, Robert Diaz-Marino, Sebastian Boring, and Saul Greenberg. 2011. The Proximity Toolkit: Prototyping Proxemic Interactions in Ubiquitous Computing Ecologies. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 315–326. https://doi.org/10.1145/2047196.2047238

[30] Marianna Obrist, Pablo Cesar, David Geerts, Tom Bartindale, and Elizabeth F. Churchill. 2015. Online Video and Interactive TV Experiences. *Interactions* 22, 5 (aug 2015), 32–37. https://doi.org/10.1145/2799629

[31] Sameera Palipana, Dariush Salami, Luis A. Leiva, and Stephan Sigg. 2021. Pantomime: Mid-Air Gesture Recognition with Sparse Millimeter-Wave Radar Point Clouds. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 27 (mar 2021), 27 pages. https://doi.org/10.1145/3448110

[32] 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

[33] Irina Popovici and Radu-Daniel Vatavu. 2019. Understanding Users' Preferences for Augmented Reality Television. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR '19)*. IEEE, Washington, D.C., USA, 269–278. https://doi.org/10.1109/ISMAR.2019.00024

[34] Mary Beth Rosson and John M. Carroll. 2002. *Scenario-Based Design*. L. Erlbaum Associates Inc., USA, 1032–1050.

[35] Niloofar Samimi, Simon von der Au, Florian Weidner, and Wolfgang Broll. 2021. AR in TV: Design and Evaluation of Mid-Air Gestures for Moderators to Control Augmented Reality Applications in TV. In *Proceedings of the 20th International Conference on Mobile and Ubiquitous Multimedia (MUM '21)*. ACM, New York, NY, USA, 137–147. https://doi.org/10.1145/3490632.3490668

[36] Ovidiu-Andrei Schipor, Radu-Daniel Vatavu, and Wenjun Wu. 2019. SAPIENS: Towards Software Architecture to Support Peripheral Interaction in Smart Environments. *Proc. ACM Hum.-Comput. Interact.* 3, EICS, Article 11 (jun 2019), 24 pages. https://doi.org/10.1145/3331153

[37] Arthur Sluÿters, Sébastien Lambot, and Jean Vanderdonckt. 2022. Hand Gesture Recognition for an Off-the-Shelf Radar by Electromagnetic Modeling and Inversion. In *Proc. of the 27th Int. Conference on Intelligent User Interfaces (IUI '22)*. ACM, New York, NY, USA, 506–522. https://doi.org/10.1145/3490099.3511107

[38] Kashmiri Stec and Lars Bo Larsen. 2018. Gestures for Controlling a Moveable TV. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video (TVX '18)*. ACM, New York, NY, USA, 5–14. https://doi.org/10.1145/3210825.3210831

[39] Muhammad Tahir, Gilles Bailly, and Eric Lecolinet. 2007. ARemote: A Tangible Interface for Selecting TV Channels. In *Proc. of the 17th Int. Conference on Artificial Reality and Telexistence (ICAT'07)*. IEEE, Washington, D.C., USA, 298–299. https://doi.org/10.1109/ICAT.2007.11

[40] Ovidiu-Ciprian Ungurean and Radu-Daniel Vatavu. 2021. Coping, Hacking, and DIY: Reframing the Accessibility of Interactions with Television for People with Motor Impairments. In *Proceedings of the ACM International Conference on Interactive Media Experiences (IMX '21)*. ACM, New York, NY, USA, 37–49. https://doi.org/10.1145/3452918.3458802

[41] Jeroen Vanattenhoven, David Geerts, Jean Vanderdonckt, and Jorge-Luis Pérez-Medina. 2019. The Impact of Comfortable Viewing Positions on Smart TV Gestures. In *Proceedings of the International Conference on Information Systems and Computer Science (INCISCOS '19)*. IEEE, Washington, D.C., USA, 296–303. https://doi.org/10.1109/INCISCOS49368.2019.00054

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

[43] 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

[44] Radu-Daniel Vatavu. 2013. A Comparative Study of User-Defined Handheld vs. Freehand Gestures for Home Entertainment Environments. *J. Ambient Intell. Smart Environ.* 5, 2 (mar 2013), 187–211. https://doi.org/10.3233/AIS-130200

[45] 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

[46] Radu-Daniel Vatavu and Stefan Pentiuc. 2008. Interactive Coffee Tables: Interfacing TV within an Intuitive, Fun and Shared Experience. In *Proceedings of the European Conference on Interactive Television (EuroITV '08, Vol. 5066)*. Springer, Berlin, Heidelberg, 183–187. https://doi.org/10.1007/978-3-540-69478-6_24

[47] Carlos Velasco, Yunwen Tu, and Marianna Obrist. 2018. Towards Multisensory Storytelling with Taste and Flavor. In *Proceedings of the 3rd International Workshop on Multisensory Approaches to Human-Food Interaction (MHFI'18)*. ACM, New York, NY, USA, Article 2, 7 pages. https://doi.org/10.1145/3279954.3279956

[48] Vinoba Vinayagamoorthy, Maxine Glancy, Christoph Ziegler, and Richard Schäffer. 2019. Personalising the TV Experience Using Augmented Reality: An Exploratory Study on Delivering Synchronised Sign Language Interpretation. In *Proc. of the CHI Conf. on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 532, 12 pages. https://doi.org/10.1145/3290605.3300762

[49] Panagiotis Vogiatzidakis and Panayiotis Koutsabasis. 2022. "Address and Command": Two-Handed Mid-Air Interactions with Multiple Home Devices. *International Journal of Human-Computer Studies* 159 (2022), 102755. https://doi.org/10.1016/j.ijhcs.2021.102755

[50] Julie Wagner, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, and Wendy E. Mackay. 2013. Body-Centric Design Space for Multi-Surface Interaction. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1299–1308. https://doi.org/10.1145/2470654.2466170

[51] Alexis Joan Walker. 1996. Couples Watching Television: Gender, Power, and the Remote Control. *Journal of Marriage and Family* 58 (1996), 813–823.

[52] Hui-Shyong Yeo and Aaron Quigley. 2017. Radar Sensing in Human-Computer Interaction. *Interactions* 25, 1 (dec 2017), 70–73. https://doi.org/10.1145/3159651

[53] Ionut-Alexandru Zaiti, Stefan-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