IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 27 December 2023, accepted 23 January 2024, date of publication 16 February 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3366667

RESEARCH ARTICLE

Analysis of User-Defined Radar-Based Hand Gestures Sensed Through Multiple Materials

ARTHUR SLUŸTERS[1], (Senior Member, IEEE), SÉBASTIEN LAMBOT[2], (Member, IEEE), JEAN VANDERDONCKT[1], (Senior Member, IEEE), AND SANTIAGO VILLARREAL-NARVAEZ[3]

[1]Louvain Research Institute in Management and Organizations, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium
[2]Earth and Life Institute, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium
[3]NADI, Université de Namur, 5000 Namur, Belgium

Corresponding author: Arthur Sluÿters (arthur.sluyters@uclouvain.be)

The work of Arthur Sluÿters was supported by Fonds de la Recherche Scientifique (FNRS) under Grant 40001931 and Grant 40011629. The work of Jean Vanderdonckt was supported by the European Innovation Council of European Union (EU EIC) Pathfinder-Awareness Inside Challenge "Symbiotik" Project (October 2022–December 2026) under Grant 101071147. The work of Santiago Villarreal-Narvaez was supported by the OPTIMIS Project by Pôle MecaTech, until December 2022, under Grant 8564.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Commission d'éthique en matière de recherche avec des personnes humaines of Institut de Recherche en Sciences Psychologiques (IPSY) under Application No. 2022-60.

ABSTRACT Radar sensing can penetrate non-conducting materials, such as glass, wood, and plastic, which makes it appropriate for recognizing gestures in environments with poor visibility, limited accessibility, and privacy sensitivity. While the performance of radar-based gesture recognition in these environments has been extensively researched, the preferences that users express for these gestures are less known. To analyze such gestures simultaneously according to their user preference and their system recognition performance, we conducted three gesture elicitation studies each with n_1=30 participants to identify user-defined, radar-based gestures sensed through three distinct materials: the glass of a shop window, the wood of an office door, and polyvinyl chloride in an emergency. On this basis, we created a new dataset of nine selected gesture classes for n_2=20 participants repeating twice the same gesture captured by radar through three materials, *i.e.*, glass, wood, and polyvinyl chloride. To uniformly compare recognition rates in these conditions with sensing variations, a specifically tailored procedure was defined and conducted with one-shot radar calibration to train and evaluate a gesture recognizer. 'Wood' achieved the best recognition rate (96.44%), followed by 'Polyvinyl chloride' and 'Glass'. We perform a preference-performance analysis of the gestures by combining the agreement rate from the elicitation studies and the recognition rate from the evaluation.

INDEX TERMS Gesture elicitation study, gesture sensing through materials, hand gesture recognition, new datasets, one-shot radar calibration, radar-based gesture recognition, user-defined gestures.

I. INTRODUCTION

Radio Detection And Ranging (Radar) technologies detect objects and movements via electromagnetic or radio waves [1] that can penetrate non-conducting materials, such as glass, wood, plastic, paper, cardboard, and rubber [2], [3]. This capability makes radars appropriate for sensing human gestures in conditions usually not covered by other sensing technologies such as computer vision: environments with poor visibility (*e.g.,* dark places, conditions with fog, rain, dust, smoke, or adverse light) [4], [5], limited accessibility (*e.g.,* hidden or indirect locations, places behind or below rough surfaces) [6], [7], and privacy-sensitivity (*e.g.,* when the user does not want to be filmed or recognized) [8], [9].

The associate editor coordinating the review of this manuscript and approving it for publication was M. Sabarimalai Manikandan.

On one hand, radar-based gesture recognition systems have demonstrated their effectiveness [10], reliability, and robustness [11] despite being influenced by environmental conditions [12]. Most studies [12], [13], [14], [15], [16], [17], [18] have reported their *performance* in terms of high recognition accuracy rates, making dynamic [16], [17], real-time [19], [20], [21] interaction feasible. High-frequency, wide-bandwidth radars detect even the finest movements [22], such as the fingers of a hand [23], at a reasonable distance [24]. However, the wide variety of radars used [25] and the particularity of their models, methods, and tools used to recognize gestures limit reuse in another context [26]. There are almost as many different methods as there are different radars, all reporting recognition rates worthy of direct real-time interaction. The same method is never used twice for two different radars. Different Machine Learning (ML) and Deep Learning (DL) methods achieve excellent results, provided they are optimized for the radar and dataset in question, at the risk of overfitting. While most of these studies optimize recognition performance [25], they give little or no consideration to the preferences expressed by end-users regarding the actual use of well-recognized gestures [27].

On the other hand, the *preference* of users for this or that gesture is the subject of an increasing number of studies [28], but they too sometimes ignore performance. Many Gesture Elicitation Studies (GESs) [29] identify a vocabulary [30] of common gestures, such as for 3D travel in mid-air [31], but this type of study is rare for radars: Siean et al. [27] deplore the scarcity of such studies, the lack of replication, and the difficulty of transposing the results. A GES [29] is a participatory design method where participants propose gestures they would like to use to execute system functions, such as snapping fingers to open the menu of a smart TV [32]. The experimenter analyzes the elicited gestures and calculates an agreement rate among participants' gesture proposals to understand which gestures are most preferred. The gesture vocabularies for existing sensors are not necessarily transposable to radars, because their contexts of use are different and because the radar assumes potentially different gestures.

In short, radar gesture performance studies deal little or nothing with user preferences, and preference studies deal little or nothing with performance. This paper reconciles performance with preference in the context of gestures recognized by a commodity radar [5] through multiple materials [33]. In this paper, we will also use a commodity radar, which refers to a standardized, widely produced, available, and cheap, radar characterized by its unified specifications, and interchangeable components. Custom-made radars are challenging to reproduce and to reuse across contexts.

The contributions of this paper are manifold: after reviewing work related to radar-based gesture recognition through various materials (Section II), we define three scenarios that are specifically tailored to radar-based gestures sensed through three materials (Section III): glass, wood, and polyvinyl chloride (PVC). Using these scenarios, we conduct three gesture elicitation studies to collect preferred gestures from participants (Section III) from which we selected nine common gestures, leading to the acquisition of a new dataset of gestures sensed through multiple materials (Section IV). To uniformly compare recognition rates, we introduce a one-shot radar calibration procedure to train a template-based recognizer on this dataset (Section IV-F). Finally, we evaluate the recognition performance and perform a performance-preference analysis (Section V) to balance performance *vs.* preference of gestures sensed through multiple materials (Section VI) and to discuss their limitations (Section VII). Section VIII summarizes the contributions and presents some avenues to this work.

II. RELATED WORK

This section reviews existing work related to radar-based gesture interaction in general (Section II-A) and in particular when the radar senses through multiple materials (Section II-B). We define a radar gesture as follows [27]: "A radar gesture is any movement or pose of a body part or the whole body, performed in mid-air, around a physical object or digital device, or concerning the body, object, or device, which is detectable and recognizable by a radar sensor."

A. RADAR-BASED GESTURE INTERACTION

1) PERFORMANCE

Radar-based gesture recognition is recognized for its abundance of models, methods, and algorithms [25], [34]. There are almost as many different approaches as there are radars with different bandwidths and frequencies. Fortunately, some surveys [35] and reviews [26] have highlighted a few trends, but these are not easily transposable to other contexts of use.

The performance of radar-based systems has made great progress, in real-time recognition with high accuracy [34], especially in recent years compared to Cheng et al. [36]'s survey, where radars were almost absent. Current research is converging on a certain intelligence of the recognition process. Wang et al. [37] investigate gesture recognition applications based on Frequency-Modulated Continuous-Wave (FMCW) radar, summarize radar research, and classify radar gestures in coarse-grained and fine-grained categories.

A Systematic Literature Review (SLR) of radar gestures [27] reported that: (1) 67% of the papers focus on performance in terms of recognition accuracy rate; (2) 30% of the papers demonstrate the application through a case study or a user scenario; (3) only 2% are dedicated to user experience aspects, very rarely on a gesture elicitation study or any other type of study that would consider human factors. The most frequently used techniques are Convolutional Neural Networks (CNN: 33%), k-nn (18%), Long Short-Term Memory (LSTM: 14%), Support Vector Machines (SVM: 9%). A more extensive SLR [26] on 118 papers confirms these figures: 54% of the 151 algorithms come from DL, in which CNN are the most frequent, combined

with LSTM or SVM. This SLR unveils a large variety of radar technology of different operating frequencies and bandwidths, antenna configurations, but also various gesture recognition techniques. Although highly accurate (all papers report an accuracy often greater than or equal to 90%), these techniques require a large amount of training data that depends on the type of radar. Only 0.85% report some user study.

Radars are today used in both stationary and mobile contexts of use. Prior works on radar-based gesture recognition [14], [18], [38], [39] mostly relied on a fixed, custom radar and on advanced ML/DL algorithms to cope with the complexity of radar signals. More recent works support dynamic, real-time recognition in mobile contexts of use [15], [18], [40]. The Google Soli chip [41] is embedded in a smartphone for recognizing various classes of gestures. There are several radar-based hand tracking systems such as EtherPose [42], CW-Radar [43], the Magic Carpet [44] and its Gesture-Sensing Radars Project, GestureVLAD [10], Ingenious [15], Forte [45] RadarSense [26], RadarGesture [13]. Radar-based gesture interaction has been pioneered by the Magic Carpet [44], a Doppler radar used for sensing coarse body gestures by signal processing. Later on, Radar Categorization for Input & Interaction) (RadarCat) [46] used a random forest algorithm to differentiate 16 transparent materials and 10 body parts. Yeo et al. [47] used radar in tangible interaction for counting, ordering, and identifying objects for tracking their orientation, movement, and between-object distance. Radars are widely used in several domains of application, such as indoor human sensing with commodity radar [5], human activity recognition [8], human position estimation [38], motion classification [14].

GestureVLAD [10] uses unsupervised frame representation followed by supervised sequence representation to recognize 11 gestures from range-Doppler images with an accuracy of 91.38%, even with slight variations, which is always a challenge [48]. Pantomime [14] uses a 76-81 GHz radar and LSTM and Pointnet++ to recognize 21 gestures from 3D point clouds with an accuracy of 95%. Wang et al. [49] recognize 2D stroke gestures: their low-dimensionality, as opposed to 3D gestures, do not require more antennas. Short-range radar-based gestures could also be recognized using 3D CNNs with a triplet loss [24]. Most existing works exploit a custom radar built with specific, on-purpose features, hard to reproduce.

More recently, Choi et al. [19] recognize a set of 10 hand Google Soli gestures in real-time. It has been magnified in RadarNet [39], an algorithm optimized for efficiently recognizing five gesture classes on computationally constrained battery-powered devices. Attygale et al. [18] exploited the Soli via a three-dimensional convolutional neural network (Conv3D) and a spectrogram-based ConvNet to recognize on-object gestures (*e.g.,* 94% for a five-gesture set). Unlike embedding a radar in any object [47], an external radar enables any object to be tracked.

In view of this work, we can apply Deepl Learning, but we are keen to test a template-based [50] recognizer [51] because it customizes the gestures by modifying templates only, with a reasonable number of templates [12], and without re-training [52], [53]. Furthermore, we are siding with a commodity radar by taking the Walabot DIY 2 radar [3], [5], [15], [26], [45], [54], [55].

2) PREFERENCE

Siean et al.'s SLR [27] mentioned only two studies. Magrofuoco et al. [56] conducted a GES for controlling IoT devices that involved 25 participants and a confirmatory study with 20 participants. The authors analyzed micro-gestures performed with the hands and fingers and sensed with a small radar, and compiled a consensus set of 19 gestures using the agreement rate measure. Sluÿters et al. [57] conducted a multi-context GES for eliciting mid-air hand gestures to interact with multimedia content presented on a large display. The resulting vocabulary was transferred to the stage of gesture recognition to determine to what extent these user-defined gestures were also efficient as system-recognizable gestures. The first study only addressed the preference for such radar gestures, without mentioning their performance while the second transferred preferred gestures from a GES to a recognizer, but we still do not know the relationship between preference and performance. More studies are needed to understand end users' preferences regarding radar gestures, especially when opportunistically selected. Only a few studies [18], [40] justify this choice.

In view of all this work, we plan to conduct a series of gesture elicitation studies that are specifically aimed at identifying gestures in radar-based contexts of use to unveil users' preferences for gestures in these contexts.

B. RADAR-BASED GESTURE INTERACTION THROUGH MULTIPLE MATERIALS

Radar sensing [46] today appears as a viable alternative to sensing techniques that cannot operate in environments with poor visibility, limited accessibility, and privacy sensitivity. Radars can be operated below a surface [8], behind a wall [3], and beneath different materials, such as wallpaper, cardboard, and wood, without affecting the recognition [14]. Radars are also insensitive to weather and lighting conditions [2]. Radar technologies have been successfully applied in various domains such as virtual reality [58], activity recognition [8], [38], tangible interaction [47], detection of leaks in the ground [59], detection of pedestrian movements [60], detection of material [61], [62], detection of objects [3], [22], and detection of gestures through materials [33], [63], fabrics [40], [42], and liquids [64].

EtherPose [42] is a continuous hand pose tracker with two wrist-worn antennas, from which the real-time dielectric loading resulting from different hand poses measure to recognize gestures performed with the hand covered by fabrics. Solids on Soli [33] delivers an extensive catalog

of materials through which motion can be detected, which estimates the range of possible materials through which radar can detect movement. The 2D radar signal contains spatial information that is easier to interface with further pose or gesture recognition functions to extract useful information [65].

Given all this work, we test gesture recognition by a mobile radar through different materials, but in realistic usage contexts based on radar-specific scenarios with a view to high-performance, preferred, real-time interaction. For these reasons, we base our work on the results of three GESs, which are delivered in the next section. To avoid calibrating the radar in multiple contexts, we use the far-field full-wave radar equation [66] to calibrate the radar and make the data independent of the radar and/or material [67].

III. GESTURE ELICITATION STUDIES FOR RADARS

Chen et al. [54] regret that there is a gap between laboratory-tested scenarios and actual application scenarios. Wang et al. [22] define use scenarios with multiple participants in adapted contexts of use that are covered by radars only. Siean et al. [28] defined a scenario for controlling a smart television. Siean et al. [27] recommend that we should "Identify genuine application areas for radar gestures", "Design gesture types specific to radar sensing", "Identify new locations for radar sensor placement". To address these needs, we define three use scenarios (Section III-A) that are tailored to radar-based interaction in contexts of use that are not covered by existing studies. We conducted a separate gesture elicitation study for each scenario, named GES1, GES2, and GES3, respectively (Section III-B), the results are reported (Section III-C).

A. DEFINITION OF USE SCENARIOS

We envision three use scenarios for radar-based interaction:

- **Shop window display.** Retail store windows create a visually attractive first impression for customers. Beyond static visual merchandising, retail companies deploy interactive displays for presenting their products in a more interactive way to improve their customer experience [68]. We define a scenario inspired by [69] and [70]: *In a shopping mall, a customer interacts by gesture with an electronic catalog of clothes displayed behind a **glass** window in a supposedly smart environment that can be controlled for light and temperature. Clothes are presented in the form of photos and sound videos. The customer browses these contents by gesture and asks a sales assistant for help, either face-to-face or remotely by phone.*
- **Office Door Situated Display.** Hermès [71] is a display located on an office door to support interaction between visitors to the office owner, whether this person is in or out of the office. We define a scenario inspired by [71]: *A visitor rings the office **wooden** door to check whether the owner is present, to identify herself or himself, and to enter the office if the owner accepts the visit. If the owner is absent, the visitor can enter the office provided that accreditation is granted. If not, voicemails, notifications, and appointments management can be operated by gestures in front of the office door.*
- **Behind wall emergency.** Emergency situations, indoor falling [72], when vital signs are engaged [73], can be detected using a remote radar behind surfaces [3]. We define a scenario inspired by [72], [73], and [74]: *Someone is trapped in a room behind a door they can no longer open, and asks for help by gesture because they are unable to speak. The subject is lying down and describes the critical situation, both bodily (e.g., pain, temperature) and environmental, and calls emergency units, e.g., doctor, ambulance, police, and fire department.*

B. EXPERIMENTS

1) PARTICIPANTS

Thirty voluntary participants were recruited for each study via a contact list in different organizations and via social networks (GES1: 13 female, 17 male, Male/Female ratio=1.30, aged between 19 and 58 years old, M=27.43, SD=12.62 – GES2: 11 female, 19 male, Male/Female ratio=1.73, aged between 18 and 65 years old – GES3: 14 female, 16 male, Male/Female ratio=1.14; aged between 9 and 60 years old, M=32.07, SD=16.57). Their occupations varied. For GES1 for instance, $\frac{7}{30}$=23% come from information technologies, $\frac{5}{30}$=16% are civil or industrial engineers, $\frac{5}{30}$=16% come from linguistics, the other domains of activity including law, journalism, or education. All participants reported frequent use of computers and smartphones, no dexterity problems, no participation in such a study, and no use of the radar involved. Two GES1 participants knew the device before the experiment, even though they never used it.

2) STIMULI

Based on each scenario, we have searched for the potential functions most frequently used in such contexts: 19 for GES1 and GES3 (see first column of Fig. 2 and 4) and 16 for GES2 (see first column of Fig. 3). We created a Microsoft PowerPoint presentation showing a *referent* [29] for each function via a before/after action represented on one slide at a time. Each representation reproduced a simplified view of the screen with the cause and the effect.

3) SETUP AND APPARATUS

The experiment took place in a quiet office meeting room. A separate laptop was used as a display for showing the referents to the participants. All gestures were recorded by a capturing application and by a camera placed in the back of the participants to capture their hands and fingers without capturing their faces. To keep the study centered on the topic, participants were asked to limit their movements in this way.

4) PROCEDURE AND TASK

The participants were welcomed to the setup by a researcher and were first asked to sign an informed GDPR-compliant

FIGURE 1. Scenarios for radars: (a) Shop window display, (b) Office door situated display, (c) Behind wall emergency.

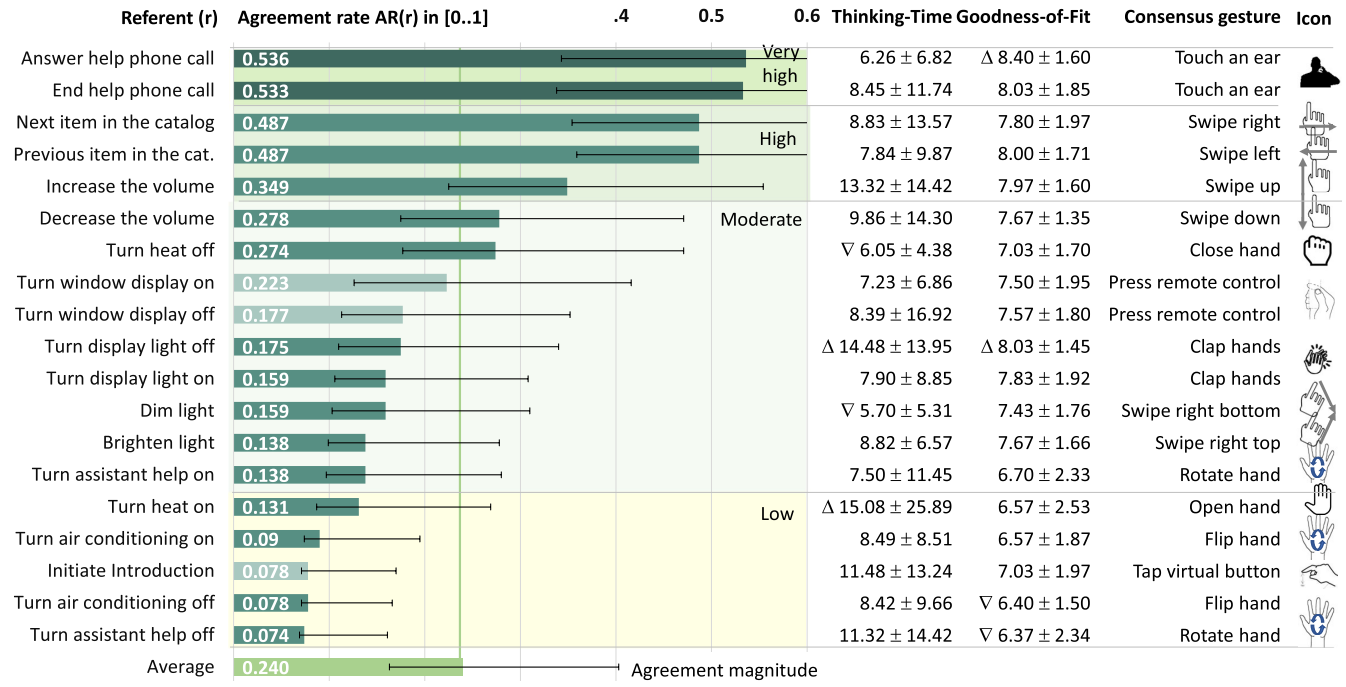


FIGURE 2. Consensus set of gestures resulting from the first gesture elicitation study "shop window display": referent, agreement rate (AR) and magnitude, thinking time, goodness-of-fit, consensus gesture, and icon. The two best and worst values of the two variables are depicted by a top-pointing arrow and a bottom-pointing arrow, respectively. Error bars show a confidence interval of 95%.

consent form. Then, they were given information about the GES and the general procedure of the experiment. They were also asked to fill out a sociodemographic questionnaire (*e.g.,* age, gender, handedness, their use of technologies based on a 7-point Likert scale ranging from 1=strongly disagree to 7=strongly agree) and to perform a Motor-skill test [75] consisted in pinching each finger with the thumb several times.

Each session implemented the GES original protocol [29], [76]: participants were presented with the *referents*, *i.e.*, actions to ensure the function in each context, for which they proposed suitable gestures to execute those referents, *i.e.*, gestures that fit referents well, are easy to produce and to remember. Participants were instructed to remain as natural as possible. The global order of referents was randomized per participant based on a real random generator based on weather temperatures. When a member of a pair of two related referents, such as "Dim light" and "Brighten light", was selected, the other member was presented just after so that participants could keep in mind the link between two related referents, thus minimizing inconsistency. Each session took approximately 40 minutes per participant. One researcher welcomed participants and checked the form filling and another experimenter presented referents, thus ensuring minimal interference from and among the researchers. The dependent variables for our within-subjects design are:

1) *Thinking-Time*: a real variable defined as the average of times elapsed between the first showing of the referent and the moment when the participant knew which gesture would be preferred, was measured in seconds.
2) *Goodness-of-fit*: a real variable defined as the average rating from 1 to 10 expressing to what extent participants thought their gesture was appropriate.
3) *Agreement rate* (AR): a real variable representing the agreement among participants for proposing a gesture

for a referent, as computed by AGATe [76] along with its magnitude (AR ≤0.1=low, 0.1< AR ≤0.3= moderate, 0.3< AR ≤0.5= high, 0.5< AR ≤1= very high).

C. RESULTS AND DISCUSSION

Figs. 2, 3, and 4 show the results for the three studies.

1) AGREEMENT RATE

Regarding **GES1**, the agreement rates are overall moderate in magnitude (M=0.24, SD=0.157), ranging between 0.074 (low) for ''Turn assistant help off'' and 0.536 (very high) for ''Answer help phone Call''. On the global sampling, $\frac{5}{19}$=26% of the rates belong to the low consensus category, $\frac{9}{19}$=47% of the rates belong to the moderate range, $\frac{3}{19}$=16% are high, and $\frac{2}{19}$=10% are very high. Apart from a few exceptions, most gestures received an agreement rate slightly higher or close to those reported in the GES literature ([76] that summarized agreement rates of 18 GESs). These values are due to two potential reasons: the design space of possible gestures with radar is so large that participants came up with many different gestures, sometimes repeatedly; the legacy bias [77] is not so present due to the perceived novelty of the radar. ''Answer help phone call'' (AR=0.436) and its symmetric ''End help phone call'' (AR=0.533) received the two best rates leading to the gesture class ''Touch an ear'', which is body-deictic (since pointing to one's ear) and pantomimic (since mimicking a physical phone reaching the ear). These two referents were assessed as the most familiar. Overall, classical gesture classes were observed (*e.g.,* ''Press remote control'', ''Swipe'', ''Rotate'', and ''Flip'' in many directions), some less frequent gestures were proposed (*e.g.,* ''Open hand'', ''Close hand'', ''Clap hands'').

Regarding **GES2**, the agreement rates are also moderate in magnitude (M=0.28, SD=0.224), ranging between 0.087 for ''Move appointment'' and 0.749 for ''Remove a notification''. The average rate falls inside the moderate category (<0.3) [76]. On the global sampling, $\frac{2}{16}$=13% of the rates belong to the very high consensus category, $\frac{4}{16}$=25% to the high category, $\frac{7}{16}$=44% to the moderate one, and $\frac{3}{16}$=18% to the low one. Despite the less usual scenario, many frequently proposed gestures are similar to touch gestures inherited from touch-based devices for ''Next screen'' and ''Previous screen'' and coincide with other GESs [78]. More original are the gestures ''Knock sign'' (AR=.301) proposed for the referent ''Call owner'', ''Hear sign'', ''Touch an ear'' (like in GES1), and ''Move fist''.

Regarding **GES3**, the agreement rates are moderate (M=0.181, SD=0.052), ranging between 0.087 for ''Call the police'' and 0.301 for ''I'm too hot'' (high magnitude). The average rate falls inside the moderate category (<.3) [76]. There are no referents with a very high rate of agreement, $\frac{1}{19}$=5% has a high agreement, $\frac{17}{19}$=90% has a moderate agreement, and $\frac{1}{19}$=5% receive a low agreement. The referents in this study are not linked to any environment or device, such as a smart home, smartphone, or smart car, therefore participants were not so subject to the legacy bias [77], which also explains the low magnitudes as they were unfamiliar with these functions. The referents with the two best rates can be associated with a device (*e.g.,* ''I'm too hot'' with air conditioning or ''I'm too cold'' with heating). This GES produced several unprecedented gestures such as ''Shake clothes'', ''Rub hands'', ''Push to threat'', ''Tap on the opposite arm'' while others were repeated from previous GES (*e.g.,* ''Clap hands'', ''Touch an ear'', ''Open hand'', and ''Close hand'').

2) THINKING TIME

This variable received quite high values (*e.g.,* M=9.25s, SD=12.43s for GES1, M=6.21s, SD=5.99s for GES3) and ranges from 3.37s for ''I had an accident'' to 9.38s for ''There is a gas leak) and large ranges (*e.g.,* from 5.70s for ''Dim light'' to 15.08s for ''Next item'' in GES1). The thinking time correlates negatively with the agreement rate (Pearson's $r_{n=19}$=−.525, R^2=.275, p=.0211): referents with a higher thinking time have a lower agreement rate. Similar figures are obtained for the two other GESs. This corresponds to Zaiţi et al. [79]' finding who found that the thinking time correlates negatively with the agreement rate.

3) GOODNESS OF FIT

This variable received a high value denoting subjective satisfaction of the participants in the gestures they proposed (*e.g.,* M=7.45, SD=0.58, Mdn=7.57 for GES1, M=7.51, SD=0.53, Mdn= 7.6 for GES2, M=7.50, SD= 0.41, Mdn=7.53 for GES3). The similarity of the values obtained for this variable in the three studies argues for a certain consistency in the way participants assessed the match between the gestures they proposed and the associated functions, despite the differences in the scenarios: the three scenarios have their level of familiarity increasing. A single-factor ANOVA did not return any significance (SS=1.81, df=3, F=.20, p>.1, *n.s.*), suggesting that the thinking time did not influence the goodness of fit. Participants were not particularly happier when they elicited a gesture faster or slower. However, it correlates significantly with the agreement rate (Pearson's $r_{n=19}$=.735, R^2=.540, p<.001***): the higher the goodness of fit, the higher the agreement rate.

IV. ACQUISITION OF GESTURE DATASETS THROUGH THREE MATERIALS

A. SELECTION OF GESTURE CLASSES

The main aim of the three gesture elicitation studies was to identify gesture classes in three different scenarios involving three different materials. Without knowing whether this influences the gestures proposed, we note that certain gesture classes recur from one study to the next, regardless of the materials involved. We will therefore proceed with a selection of different gesture categories, starting with those found in the studies, preferably more than once.

One can memorize only a limited number of gestures, preferably user-defined: Nacenta et al. [83] found that

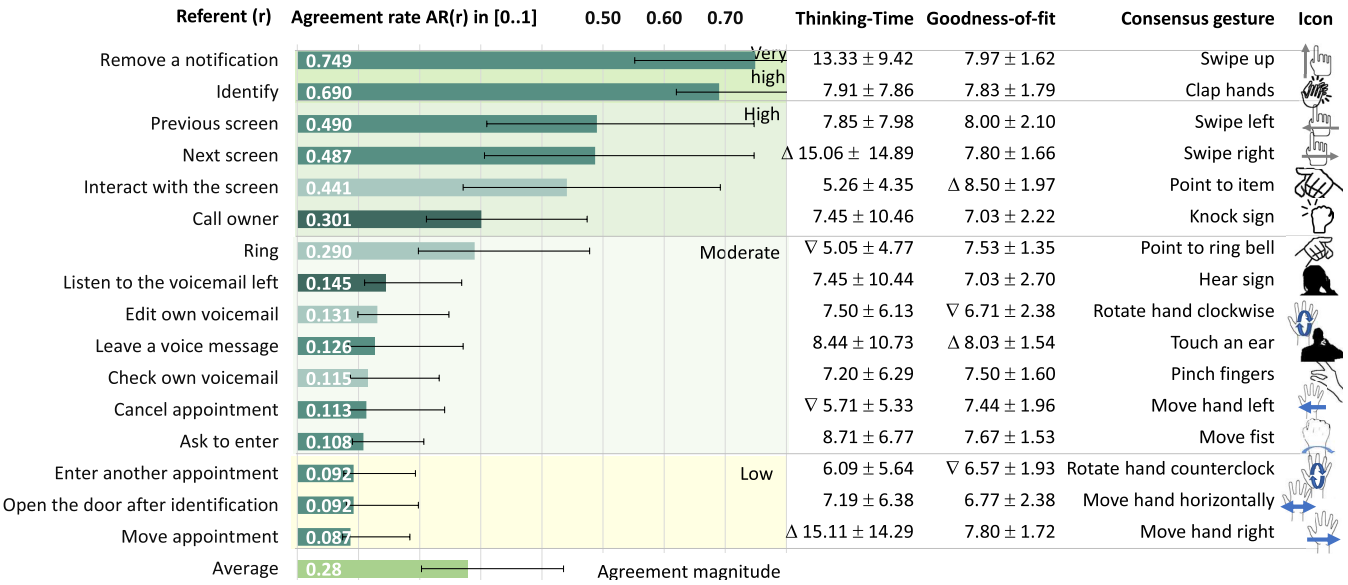


FIGURE 3. Consensus set of gestures resulting from the second elicitation study "office door situated display": referent, agreement rate (AR) and magnitude, thinking time, goodness-of-fit, consensus gesture, and icon. The two best and worst values of the two variables are depicted by a top-pointing arrow and a bottom-pointing arrow, respectively. Error bars show a confidence interval of 95%.

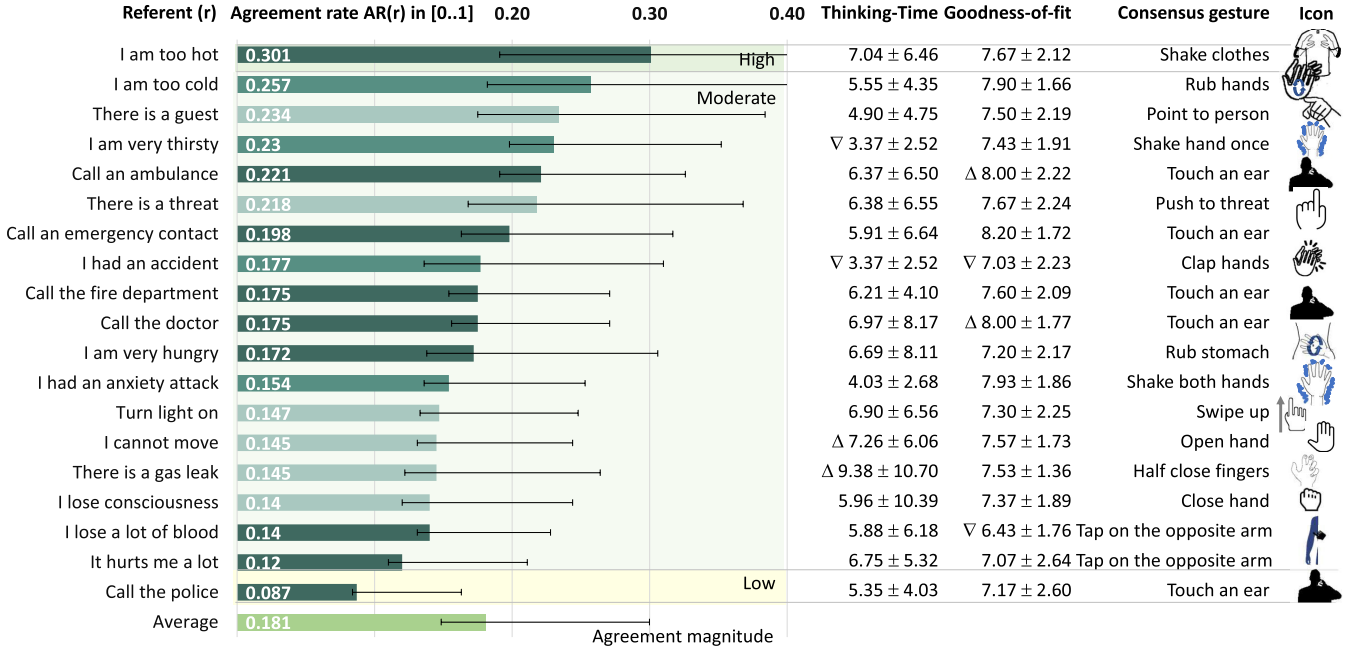


FIGURE 4. Consensus set of gestures resulting from the third gesture elicitation study for the "behind wall emergency": referent, agreement rate (AR) and magnitude, thinking time, goodness-of-fit, consensus gesture, and icon. The two best and worst values of the two variables are depicted by a top-pointing arrow and a bottom-pointing arrow, respectively. Error bars show a confidence interval of 95%.

user-defined gestures are easier to remember, both immediately after creation and on the next day with a 24% difference in recall rate compared to pre-designed gestures. Sluÿters et al. [57] reported also an excellent recall rate of a limited number of gestures when gestures were user-defined, even after one week. So, we wanted to select a limited number of classes from the user-defined gestures proposed in our three aforementioned studies. Our selection of nine gesture classes was based on the user-defined gestures proposed in the three aforementioned studies, on the review of successful radar-based gesture recognition as well as on the heuristic evaluation [84] considering the capabilities of a commodity radar. We selected nine gestures that cover the three levels of taxonomy (*i.e.*, hand, arm, and body - see Fig. 5) and that were identified in at least one GES. Table 1 shows our selection: "Open hand" from GES1 and GES3, "Close hand" from

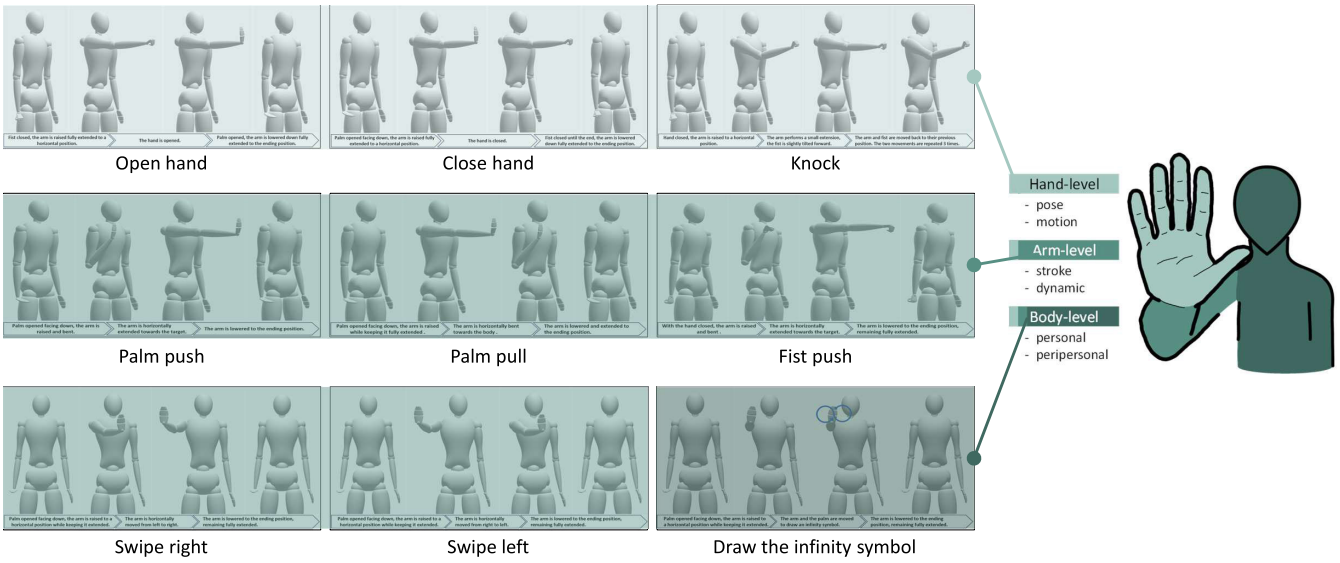


FIGURE 5. Selection of nine user-defined gestures classified according to the three-level taxonomy: hand (light green), arm (medium green), and body (dark green). Each gesture representation is decomposed into four steps for visual priming during the study.

TABLE 1. Selection of gesture classes and their characterization. Cat. refers to the category illustrated in Fig. 5: hand, arm, or body level. Class. refers to the location in Aigner et al.'s [80] classification such as M=metaphoric, D=dynamic, S=semaphoric, T=stroke, P=pantomimic and in wang et al. [37]'s classification such as C=coarse grained gesture and F=fine grained gesture. Corresp. denotes the correspondence of the selected gesture to gestures identified in radar-based studies when any. GES name refers to the presence of the selected gesture in any previously conducted GES (1 to 3), along with its Thinking-Time (TT), Goodness-of-Fit (GoF), and Agreement Rate (AR). The last column refers to the timed gesture image and its filtering in Fig. 7.

Name	Description	Cat.	Class.	Corresp.	GES name (ATT, GoF, AR)			Timed filtered
					GES 1	GES 2	GES 3	
Open hand	Raise the arm to a fully extended horizontal position and open the hand	Hand	DS, F	[27], [81]	Open hand (6.09, 6.57, 0.131)		Open hand (7.26, 7.57, 0.145)	Fig. 8a
Close hand	Raise the arm to a fully extended horizontal position and close the hand	Hand	DS, F	[24], [27], [81]	Close hand (6.05, 7.03, 0.274)		Close hand (5.96, 7.37, 0.140)	Fig. 8b
Knock	Raise the arm to a fully extended horizontal position and tilt the fist forward	Hand	P, F	[16], [23], [27]		Knock sign (7.45, 7.03, 0.301)		Fig. 8c
Palm push	Push the open palm hand forward	Arm	DS, C	[13], [22], [27], [82]			Push to threat (6.38, 7.67, 0.218)	Fig. 8d
Palm pull	Pull the open palm hand backward to the body	Arm	DS, C	[13], [22], [27], [82]		Move hand hor. (7.19, 6.77, 0.092)		Fig. 8e
Fist push	Extend the arm horizontally with the fist closed	Arm	DS, C	[45]		Move fist (8.82, 7.67, 0.108)		Fig. 7f
Swipe right	Extend the arm horizontally with the palm open and move it to the right	Arm	ST, C	[22], [27], [39], [49]	Swipe right (15.08, 7.80, 0.487)	Swipe right (15.08, 7.80, 0.487)		Fig. 7g
Swipe left	Extend the arm horizontally with the palm open and move it to the left	Arm	ST, C	[22], [27], [39], [49]	Swipe left (7.84, 8.00, 0.487)	Swipe left (7.84, 8.00, 0.490)		Fig. 7h
Infinity	Draw the infinity symbol in mid-air	Body	DI, C	[45]			Rub stomach (6.69, 7.20, 0.172)	Fig. 7i

GES1 and GES3, "Knock" from GES2, "Palm push" from GES3, "Palm pull" from GES2, "Fist push" from GES2, "Swipe right" from GES1 and GES2, "Swipe left" from GES1 and GES2, and "Infinity" from GES3. Our selection of nine gestures partially overlaps with other classifications, taxonomies, and other references (col. 4 and 5 of Table 1). Compared to [56], our "Swipe right" and "Swipe left" gestures can be mapped to functions that navigate, whereas our "Open hand" and "Close hand" gestures can be mapped to opening/closing functions.

Fig. 5 shows also the decomposition of each gesture into a sequence of steps materialized by an online mannequin that served for visual priming in the dataset acquisition. For example, "Palm push" is described as follows: from the starting position, with the palm opened facing down, the arm is raised and bent keeping the hand as close as possible to the body, and such that the palm is parallel to the surface of the plate, facing the target. Then, the arm is horizontally extended towards the target, while keeping the palm parallel to it, until the full extension. Finally, the arm is lowered to the ending

position, remaining fully extended and with the palm opened until the end.

B. MATERIAL SELECTION

Solids on Soli [33] delivered an extensive catalog of various materials, either as a single layer or with multiple layers (*e.g.,* multiple sheets of paper), which is very useful to determine the impact of non-conductive material on radar penetration. In contrast, we wanted to test radar gesture sensed through a limited number of materials that were selected based on the use scenarios (Section III-A) and on the following rationale:

- **Glass**: this material is frequently used in everyday life, in every house, as well as in public buildings like shops for their showcases. An interesting property that glass has over other materials is that it is transparent, thus enabling the end user to interact with a screen behind the glass.
- **Wood**: this common material is encountered very often in day-to-day life and can be utilized in a wide variety of contexts. Indeed, wood is used to make furniture such as desks and shelves, or for house elements such as doors and shutters. It can also be used for structural purposes. A wide variety of wood exists, all slightly varying in density and permittivity.
- **PVC**: this material is very popular in the piping industry, it is also used to make chairs or curtains, or even for flooring. The main interest of having PVC in the pool of materials is to have a representative of the polymers.

With the combination of these three materials, we cover a wide variety of real-world contexts of use and a large number of potential scenarios in which hand gesture recognition can be deployed. For this study, a wood plate made of plywood of 100 cm × 100.1 cm × 1.7 cm (W/H/D), a glass plate of 100 cm × 100 cm × 0.5 cm (W/H/D), and a PVC plate of 100.1 cm × 100.5 cm x 0.9 cm (W/H/D) were used (see Fig. 7–3, 4, 5 and the right part of Fig. 6).

C. APPARATUS

To ensure a cost-effective and accessible interaction, we selected the Walabot Developer Pack, a commodity ultra-wideband (UWB) FMCW radar whose dimensions are 5.67 in.×3.35 in.×0.71 in. We used the version operating over the narrower 6.3-8 GHz range, which is the most restrictive and challenging for gesture recognition. The device is equipped with an array of 18 antennas: 4 as transmitters and 14 as receivers. Depending on the configuration, it senses motion with up to 40 pairs of antennas. The Walabot has been proven efficient in various domains of applications, such as material identification [61] and activity recognition [8], [85]. Avrahami et al. [8] recognize human activities at a checkout counter and a typical office desk with a sensor deployed under the work surface. When the subject performs predefined activities, data is captured as RF samples. Zhang et al. [16] propose a deep neural network for continuous gesture recognition and evaluate it on a dataset of eight hand gestures (with 150 samples per class) performed very close to the radar. The network is trained with 120 samples per gesture while the 30 remaining samples are used for testing. Walabot can be used as a wall scanner to image studs or pipes behind a wall. Wang et al. [49] used this device for the early detection of nematodes in walnuts. Walabot generates a 2D heatmap image in spherical coordinates from the facing arena. The size of the arena can be modified within the range of θ=−45° to θ=45°, ϕ=−90° to ϕ=90° and R up to 10 m.

D. SETUP

The three different plates were mounted on easels in front of the participant and marked on the ground using tape to preserve the consistency of experimental conditions (Fig. 6). Thanks to these easels, the position of the plates can be adapted vertically and tilting can also be adjusted. The tilt angle of the easels is adjusted once, before the very first recording session, such that the plates would remain stable at a 90◦ with the ground, perfectly vertical, remaining untouched throughout the recording sessions. The height of each easel is adjusted vertically accordingly according to the participant during each recording session. A cardboard ruler is used to define the maximal position of the feet of the participant. Each radar is attached to the center of a plate using tape, in a horizontal position. The Walabot is connected via its micro-USB port to the USB port of a computer using a micro-USB to USB cable. The computer controlled by the experimenter is placed on a table behind the easels and the plates. A Walabot-specific recording software is used to register the raw data from the device. The software prompts the experimenter to enter a series of parameters in a configuration file (see Section VIII), valued as follows:

- *The name of the dataset*, the following format, which includes the anonymous code attributed to the participant, was used: anonymousCode_material, *e.g.,* 54_wood.
- *The profile of the Walabot*, always set to the value 2. It corresponds to the profile PROF_SENSOR_NARROW of the Walabot, which is available on both the EU/CE and the US/FCC versions. With this profile, the Walabot normally records 4096 fast-time samples per frame, but the software limits the recording to the first 1024 fast-time samples of each frame.
- *The user ID*, corresponding to the anonymous code.
- *The number of gesture classes*, is always set to the value 10 (the 9 gestures plus one corresponding to the recording of the background).
- *The number of repetitions of each gesture*, is always set to the value 5, even though the background is only recorded once, at the very end.
- *The automatic naming of the output files*, always set to ''yes''. This will name the file with the following format: WalabotSignal_User_ userID_Gesture_gesture Class_*repetition*. For example, the naming could produce something like WalabotSignal_User_54_ Gesture_6-3, which is useful for future computation.

The computer of the experimenter is also connected to a secondary screen placed on the ground, in between the

FIGURE 6. Setup for radar-based gesture acquisition through three materials: wood, glass, and PVC.

easels to show the gesture animations (Fig. 5) for visual priming.

E. PROCEDURE FOR GESTURE ACQUISITION

The following protocol was used to record the hand gestures:

1) The participant puts aside all objects from their pockets and removes any jewelry from their hand such as rings, bracelets and watches. Participants who were unable to remove some rings kept them during the experiment.
2) Starting with the glass plate, the participant is asked to stand straight facing the surface of the plate and to hold their arm horizontally fully extended in front of him. The experimenter adapts the height of the easel such that the target is aligned with the arm of the participant.
3) The participant, with their arm still in the horizontal position, is asked to open their hand such that their palm faces the surface of the plate. Then, the participant moves forward until their palm is 20 cm away from the

plate. The experimenter places a cardboard box on the ground against the feet of the participant.

4) The experimenter makes sure that the radar is plugged into the computer, starts the recording software, and configures it with the correct parameters. The experimenter verbally describes the gesture to perform to the participant and then shows its corresponding animation. Five repetitions are recorded:
   a) The participant stands in front of the plate, feet against the cardboard box, in the starting position.
   b) The experimenter starts recording and says "top".
   c) The participant performs the gesture.
   d) Once the participant reaches the ending position, the experimenter stops the recording.
   e) If the participant did not perform the gesture correctly, or if there was any issue during the recording, an extra repetition of the gesture is done.
5) For each gesture, step 4 is repeated.
6) Once all the gestures are achieved, the experimenter asks the participant to step aside.
7) Steps 2 to 6 are repeated for the wood and PVC plates.

We have now selected nine gestures to be acquired on the Walabot radar with glass, wood, and PVC. Given that future gestures will have to be recognized by a recognizer for these three materials, we introduce in the following section a novel procedure that avoids re-calibrating the radar each time the material is changed. A recognizer can be trained on one material at a time, but we want to consider all three materials at once.

F. ONE-SHOT RADAR CALIBRATION

A *one-shot radar calibration* is a process in which a radar is calibrated only once using a single measure concerning a single target. While classical calibration techniques are multi-shot when they perform multiple measures concerning multiple, different targets, thereby requiring sophisticated procedures, a one-shot calibration largely simplifies the process by relying on a single target.

During a one-shot radar calibration, the radar system collects data from the reference target, which could be a known reflective object or a calibration standard with precisely measured properties. The radar system then uses the acquired data to adjust its internal parameters, such as gain, timing, or frequency settings, to align with the expected performance. We used the far-field full-wave radar equation [66] to accurately model the radar signal. This equation allowed us to calibrate the radar system, mitigating the radar-antenna-material effects. It is expressed in the frequency domain as:

$$S(\omega) = R_i(\omega) + \frac{T(\omega)G_{xx}(\omega)}{1 - R_s(\omega)G_{xx}(\omega)} \tag{1}$$

where $S(\omega)$ represents the ratio between the scattered and incident field at the radar reference plane, ω is the angular frequency, $R_i(\omega)$ denotes the global reflection coefficient of the antenna for fields incident from the radar reference plane onto the source point (phase center), $T(\omega)$ represents the product of the global transmission coefficients for fields incident from the field point (phase center) onto the radar reference plane and vice versa, and $R_s(\omega)$ is the global reflection coefficient for fields incident from the target onto the field point. The term $G_{xx}(\omega)$ represents the Green's function for wave propagation in 3D layered media [66]. For the Walabot, after applying a Fast Fourier Transform, we measure $b(\omega)$ and $a(\omega)$ remains unknown. Consequently, Eq. (1) is rewritten as:

$$b(\omega) = H_i(\omega) + \frac{H(\omega)G_{xx}(\omega)}{1 - H_s(\omega)G_{xx}(\omega)} \tag{2}$$

where we define $H_i(\omega) = a(\omega)R_i(\omega)$, $H(\omega) = a(\omega)T(\omega)$, and $H_s(\omega) = R_s(\omega)$ as the three radar-antenna characteristic functions. To apply the full-wave radar equation, it is necessary to know these functions, denoted as $H_i(\omega)$, $H(\omega)$, and $H_s(\omega)$. These functions are determined through calibration measurements conducted at varying distances from a known medium, for which corresponding Green's functions can be computed. Typically, a perfect electrical conductor, such as a copper plane, is used for this purpose. Once these measurements are complete, and the corresponding Green's functions are calculated, they are used within a linear system of equations, which is solved analytically to derive the three unknown functions. Importantly, each frequency component operates independently.

Our calibration procedure was conducted at the lab facilities of GPRLouvain, which are equipped with an automated positioning scanner and a 3 m × 3 m copper plane (Fig. 7-1) designed specifically for radar calibration. Measurements were taken at five distinct distances from the copper plane thanks to a 3D movable arm to determine the three unknowns per frequency (Fig. 7-2). This results in an overdetermined system of equations, enhancing the calibration precision. The radar equation is theoretically applicable to any antenna geometry; therefore, we calibrated the Walabot mounted on three materials of interest: glass, wood, and PVC (Fig. 7-3–4). This assumes that these materials are integrated into the radar-antenna system, inherently accounting for all reflections and transmissions related to their presence and interactions with the radar unit itself. The radar equation can effectively remove radar-antenna-material effects:

$$G_{xx}(\omega) = \frac{S(\omega) - H_i(\omega)}{H(\omega) + S(\omega)H_s(\omega) - H_i(\omega)H_s(\omega)} \tag{3}$$

This equation yields measurements that are independent of material composition and thickness. Within a given frequency range, it provides a degree of independence for gesture recognition, mitigating the influence of both the specific radar device and the intervening material between the radar and the target individual. Notably, the far-field model effectively accounts for most factors, with the exception of the antenna radiation pattern. While the generalized radar equation [86] could emulate the antenna radiation pattern, it does so at the cost of significantly increased complexity. Depending

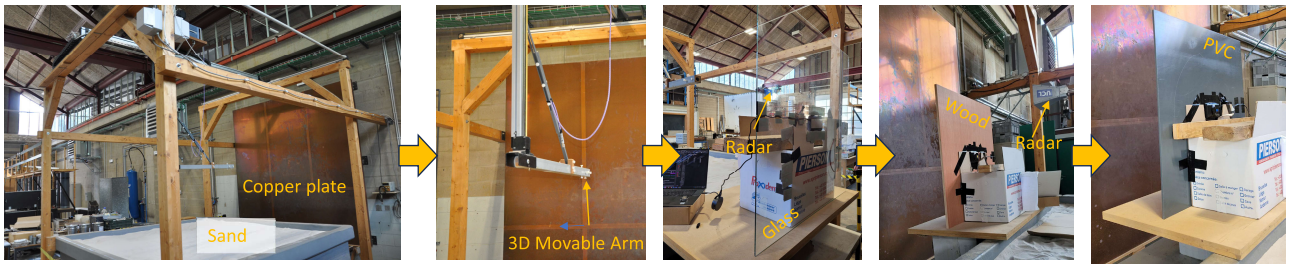


FIGURE 7. One-shot calibration of a radar: (1) preparation of the mobile arm in front of the copper plate above a sandbox, (2) positioning scanner, (3) radar calibration on glass, (4) on wood, and on (5) PVC.

on the impedance characteristics of the material, the field transmitted to the target can be lower than the emissions from the radar, resulting in a reduced Signal-to-Noise Ratio (SNR).

G. GESTURE DATASET ACQUISITION AND PROCESSING

1) ACQUISITION

Through the same channels used for the GES in Section III, we recruited 20 new participants (8 females, 12 males, and 0 identified as gender variant/non-conforming, aged between 19 and 81 years old, M=37.30, SD=19.18, 2 left-handed *vs.* 18 right-handed) who did not participate in these GES to avoid any carry-over effect. After performing the one-shot calibration (Section IV-F), installing the setup (Section IV-D), and running the procedure (Section IV-E), we collected a dataset of 20 participants × 9 gestures × 5 repetitions × 3 materials = 2,700 samples.

2) PROCESSING

We have sequentially executed the following signal-processing techniques on the dataset: the raw data are transformed into the time range by performing the Fast Fourier Transform (FFT) [87] along the fast time dimension; the background scene is removed from the resulting output [88]; an Inverse Fast Fourier Transform (IFFT) [89] is performed; a Time Gating [90] truncates the signal between 0 and 4 nanosec. to isolate the hand gesture from the body (*i.e.*, in this window, the signal corresponding to the reflections occurring between 0 and 60 centimeters from the radar is kept – see left part of gestures in Fig. 8); and final filtering (right part of gestures in Fig. 8) defines the permittivity value below which the estimated value of distance and permittivity are replaced by 75 cm and 1.1, respectively.

Fig. 8-left reproduces the output in the time domain of one antenna pair after the Time Gating with wood, where the X axis denotes the frame index whereas the Y axis is the time at which the signal is received to be interpreted as the distance between the hand and the radar. The color of each point corresponds to the signal amplitude. For example, the "Palm push" (Fig. 9d-left) is decomposed into a phase where the palm approaches the radar, a phase where the palm is stationary in front of the radar, and a phase where the signal disappears after the participant takes the hand out of the field of view. We observe a similar phenomenon for "Fist push" (Fig. 8f-left) and the opposite for "Palm pull" (Fig. 9e-left). Fig. 8-right reproduces the output of one antenna pair after the Filtering with wood, where the X axis represents the frame index and the Y axis represents the estimated distance and permittivity. The results confirm that they correctly represent the distance between the hand and the radar, but are more varying for the permittivity.

For example, the "Palm push" (Fig. 9d) and the "Palm pull" (Fig. 9e) see their permittivity rising when the palm is in front of the radar and staying low the rest of the time. However, for the other gestures where the configuration and the position of the hand is less consistent in front of the radar, the permittivity curves contain some spikes that do not always fit with the gestures. The amplitude of the reflection is also affected when there are slight variations in the orientation of the hand.

H. USER PRIVACY AND DATA SECURITY

For each GES and gesture acquisition, a random sampling was applied to select the corresponding number of participants from a list of candidates maintained in our organization. Participant occupations included various domains, such as management, social sciences, engineering, computer science, education, law, office, and finance. We checked that all participants owned smartphones and/or tablets, that they were used to touch and gesture input and they did not suffer from any particular physical restriction. They were invited to sign a GDPR-compliant consent form approved by our ethical committee specifying that their data is anonymized and that the dataset will only include an anonymous identifier, thus preserving their privacy. Data security is ensured by computer protection against external attacks and secure communication between the Walabot and the laptop running the recognizer.

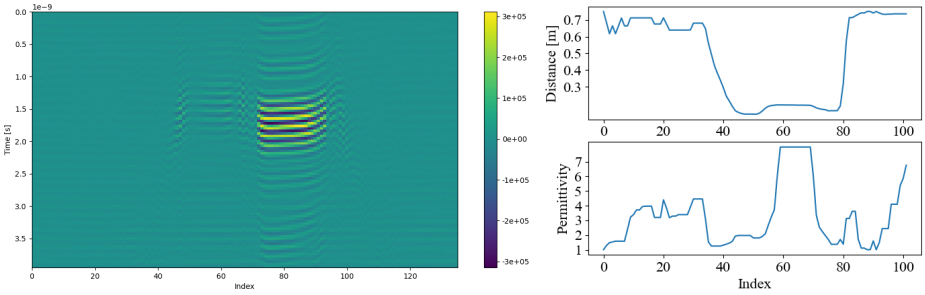
V. PERFORMANCE-PREFERENCE ANALYSIS

A. PERFORMANCE

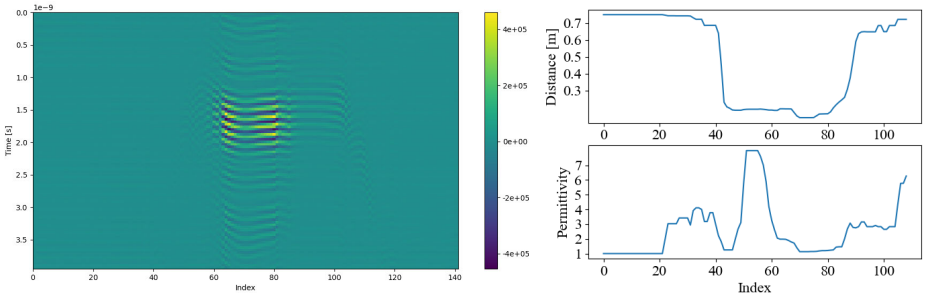
1) INDEPENDENT VARIABLES

We define the following independent variables:

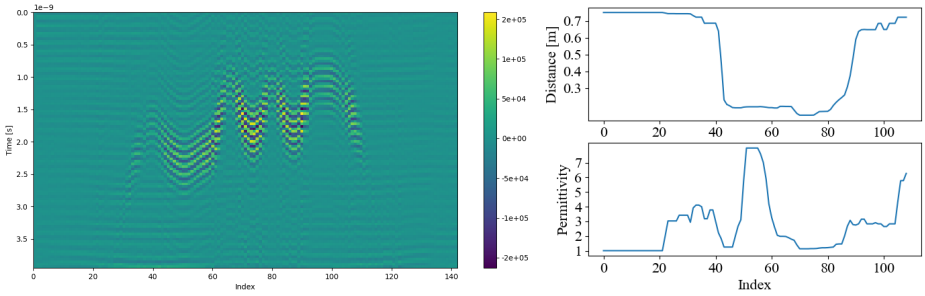
- *Number of training templates (T)*: represents the number of gesture samples used to train the recognizer. In the user-dependent scenario, it is assigned to T={1, 2, 4},



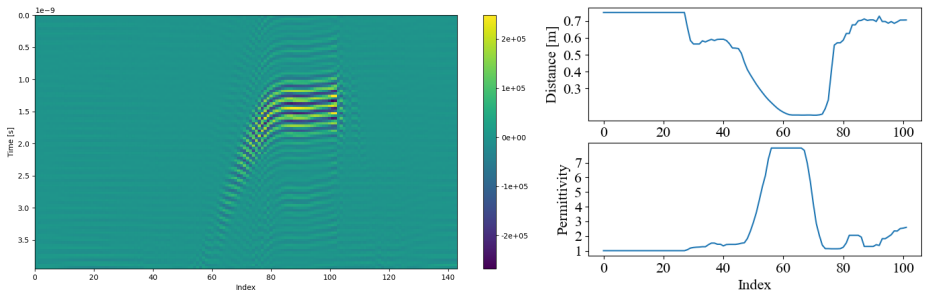
(a) Open hand.



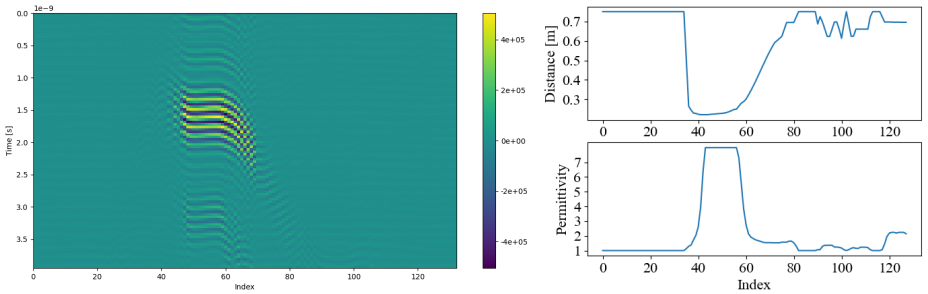
(b) Close hand.



(c) Knock.

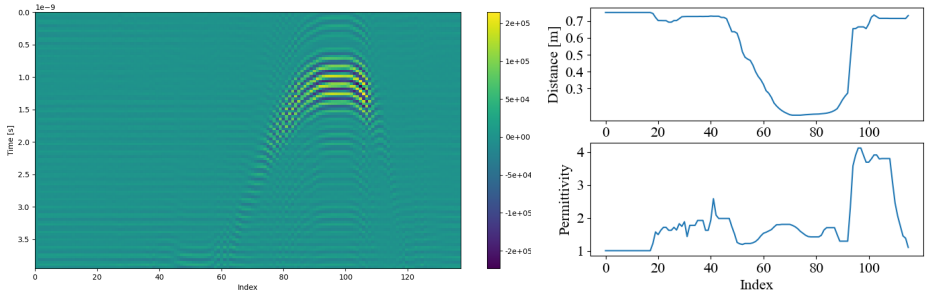


(d) Palm push.

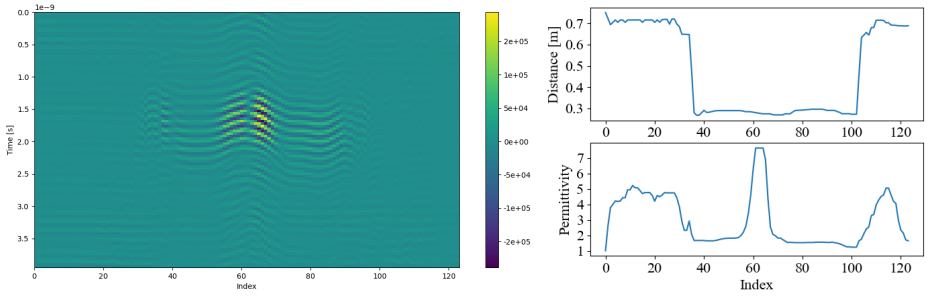


(e) Palm pull.

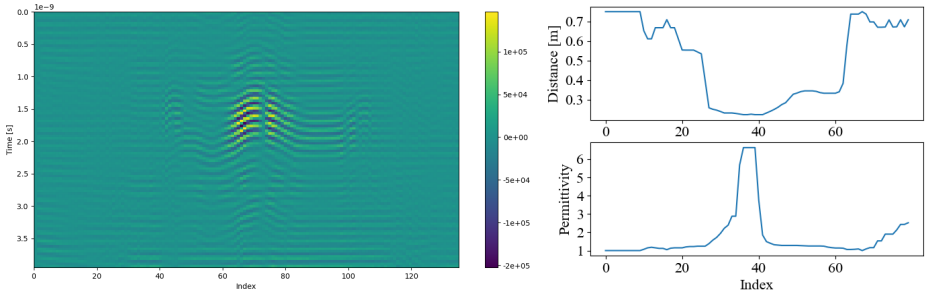
FIGURE 8. Signal of an antenna pair after the time gating (left) and filtering (right) for each gesture performed in front of the wood plate.



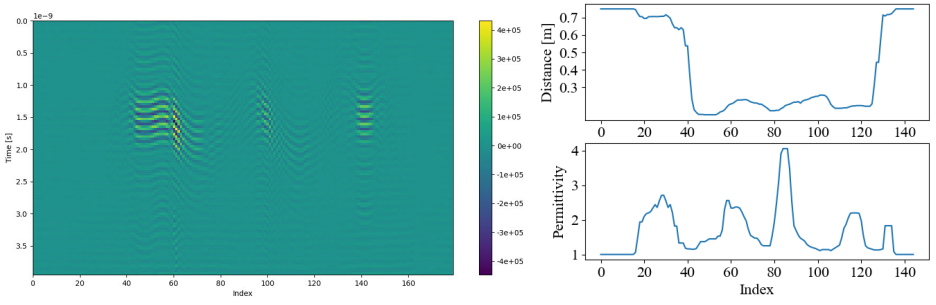
(f) Fist push.



(g) Swipe right.



(h) Swipe left.



(i) Draw infinity symbol.

FIGURE 8. ***(Continued.)*** **Signal of an antenna pair after the time gating (left) and filtering (right) for each gesture performed in front of the wood plate.**

as we have only 5 samples per gesture per user while in the user-independent scenario, it is assigned to T={1, 2, 4, 8, 16}.

- *Number of sampling points (N)*: represents the number of points per gesture template: N={$x \in \mathbb{N} | 4 \leq x \leq 40$}.
- *Antenna pairs (AP)*: defines the set of antenna pairs from which the data were used: AP={(1,2,3,6,8,9), (4,5,7,10,11,12), (1,2,3,4,5,6,7,8,9,10,11,12)}. These three sets demonstrated admissible results [26]. Furthermore, the two sets of 6 pairs are composed of non-redundant antenna pairs.
- *User dependence*: specifies whether the gesture recognition was evaluated in user-dependent (each repetition of the training and testing is done with only the gesture samples from one particular user) and/or user-independent (all the gesture samples from all the users

are mixed and used together in training and testing) scenarios [53].

2) GENERAL PROCEDURE

To compute the *Recognition Rate*, *i.e.*, the ratio of correctly recognized gestures divided by the number of trials, and the *Execution Time*, *i.e.*, the time for recognizing the class of a candidate gesture, for each combination of the independent variables, a typical procedure relies on a Leave-One-Out Cross-Validation (LOOCV) [91]. For each variable combination and gesture class, three steps are repeated 1000 times: (1) select a random testing sample from the gesture set; (2) train the recognizer using a set of randomly selected samples, produced by the same user as the testing sample; and (3) recognize the testing sample. We used the Jackknife recognizer [51] in all subsequent testings because this recognizer is efficient on multiple data types, is modality agnostic, and obtains excellent results with few samples. Furthermore, its template matching [50] enables the end user to customize the gesture set by editing templates without retraining everything, contrarily to ML/DL algorithms. All tests are run on a Dell laptop equipped with an Intel Core i7-10875H CPU at 2.30GHz with 31.8 Go of exploitable RAM.

The training and testing of the recognizer were achieved according to a train-test split methodology: one sample of a gesture class is selected for the testing set and a part or all the other samples are used for the training set. Once the training is complete, the recognizer assigns a label, corresponding to a gesture class, to the sample in the testing set. This procedure was repeated 200 times for the 9 gesture classes for each set of parameters, totaling 1,800 label assignments each time.

3) RESULTS AFTER THE BACKGROUND SUBTRACTION

In the user-dependent scenario, the best recognition rate with wood is 96.1% with the configuration T=4, N=33, and AP=(1,2,3,6,8,9), resulting in an execution time of 5.05 ms, thereby making it eligible for dynamic, real-time interaction. With the PVC, the best recognition rate is 83.7% with the configuration T=4, N=38, and AP=(4,5,7,10,11,12) for an execution time of 5.72 ms. Gestures performed in front of the glass are recognized with a rate of 67% with the configuration T=4, N=38, and AP=(1,2,3,4,5,6,7,8,9,10,11,12), leading to an execution time of 8.47 ms. Fig. 9 shows the evolution of the recognition rate for the best set of antenna pairs parameter for each plate. Independently of the plate, a higher number of training templates always gives better accuracy. Concerning the AP parameter, the best set of antenna pairs is different for the 3 plates, but the accuracy across all the sets for the same plate remains close. Regarding the number of points, the accuracy starts to increase with the number of sampling points increasing, but only up to a certain point. Around 25 to 30 sampling points, the accuracy starts to level off.

This remains the case for up to 40 sampling points. If we analyze the accuracy across the different plates, the wood plate gives overall better results, followed by the PVC plate, and the glass plate. This is expected, as the wood plate made of plywood has the lowest permittivity of approximately 2.8 [92], compared to a permittivity of approximately 6.0 [92] for the glass plate, which has a much greater impact on the radar signal. The PVC plate, with a permittivity of approximately 4 [72], gives results in between the two other plates.

In the user-independent scenario, the recognition rate is always low. For the gestures performed in front of the wood plate, the highest recognition rate is 31.5% with an execution time of 23.63 ms with the configuration T=16, N=40, and AP=(1,2,3,6,8,9). For the PVC plate, the best rate is 20.2% with an execution time of 21.45 ms with the configuration T=16, N=36, and AP=(4,5,7,10,11,12). Finally, we observe an accuracy of 19.8% at best with an execution time of 26.97 ms for the glass plate with the configuration T=16, N=29, and AP=(1,2,3,4,5,6,7,8,9,10,11,12). Fig. 10 shows the complete accuracy results for the best set of antenna pairs parameter for each plate.

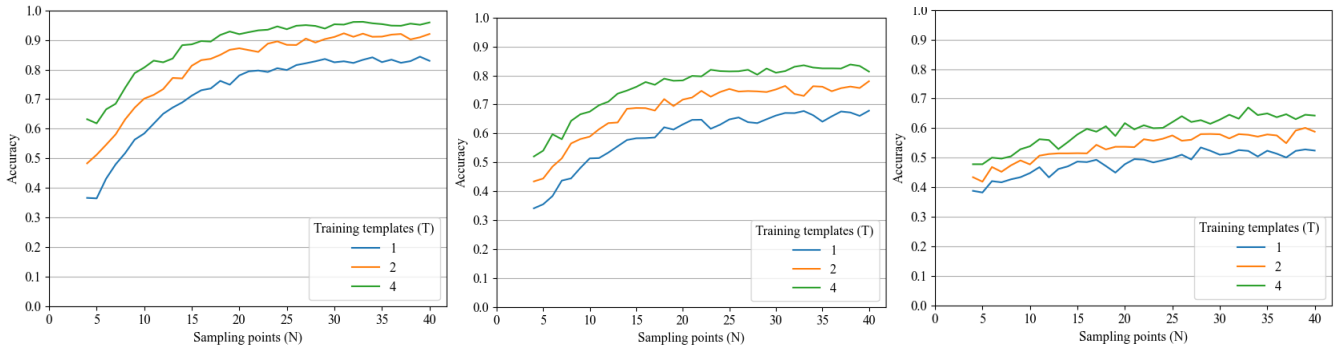
Compared to the user-dependent scenarios, the accuracy is much lower. The impact of the training templates and sampling point parameters remains the same for the wood plate and the PVC plate. The more training templates and sampling points, the higher the accuracy is, even though the impact is very small for the PVC plate. For the glass plate, increasing the sampling points doesn't significantly improve the accuracy. Concerning the set of antenna pairs parameter, the highest accuracy is obtained with the same sets as in the user-dependent scenarios for each material. Once again, the wood plate has the highest accuracy, followed by the PVC plate and then the glass plate.

4) RESULTS AFTER THE FILTERING STEP

Concerning the results with the filtering step data in the user-dependent scenarios, the highest accuracy obtained for the recognition of gestures done in front of the wood plate is 66.2% with an execution time of 0.38 ms with the configuration T=4, N=29, and AP=(1,2,3,4,5,6,7,8,9,10,11,12). For the gestures done in front of the PVC plate, the highest accuracy is 61.2% with an ART of 0.34 ms with the configuration T=4, N=27, and AP=(1,2,3,4,5,6,7,8,9,10,11,12). Finally, with the glass plate, the best accuracy is 66.3% with an ART of 0.26 ms with the configuration T=4, N=21, and AP=(1,2,3,4,5,6,7,8,9,10,11,12). Fig. 11 shows the complete accuracy results for the best set of antenna pairs for each plate.

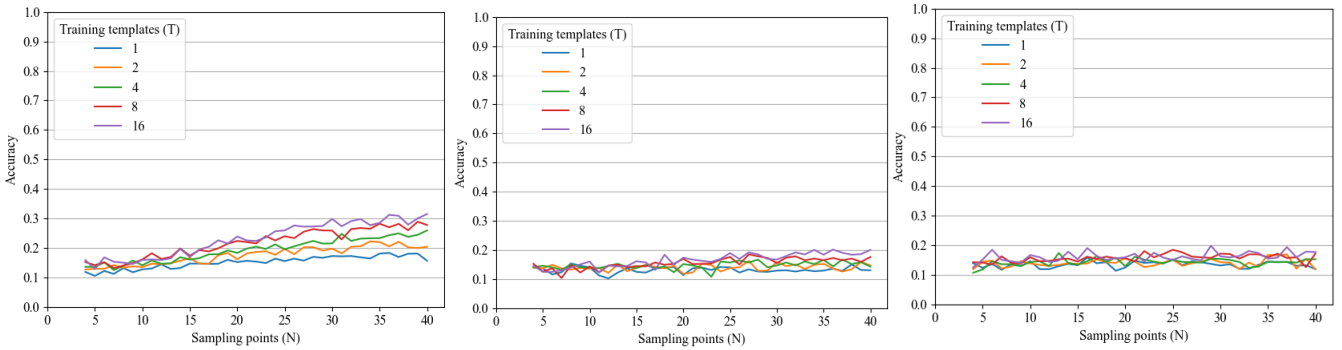
Independently of the plate, the higher number of training templates always gives a better accuracy. Concerning the AP parameter, the set with the 12 antenna pairs always gives better accuracy than the two other sets with 6 antenna pairs. The impact of the number of sampling points is also similar to what was observed with the background subtraction data. The accuracy starts to increase with the number of sampling points increasing, but only up to a certain point. Around 20 to 25 sampling points, the accuracy starts to level off. This remains the case for up to 40 sampling points.

After the filtering step, we obtained similar recognition rates across the 3 plates, which was not the case with the background subtraction data. If we compare the results for



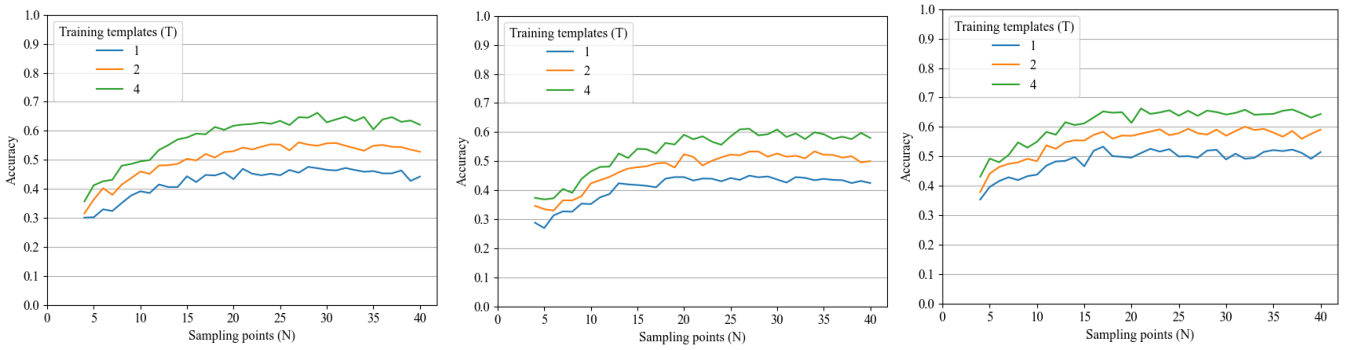
(a) Wood plate, AP=(1,2,3,6,8,9). (b) PVC plate, AP=(4,5,7,10,11,12). (c) Glass plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12).

FIGURE 9. Recognition rate after background subtraction in the user-dependent scenarios with the best AP for the plates.



(a) Wood plate, AP=(1,2,3,6,8,9). (b) PVC plate, AP=(4,5,7,10,11,12). (c) Glass plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12).

FIGURE 10. Recognition rate after the background subtraction in the user-independent scenarios with the best AP for the plates.

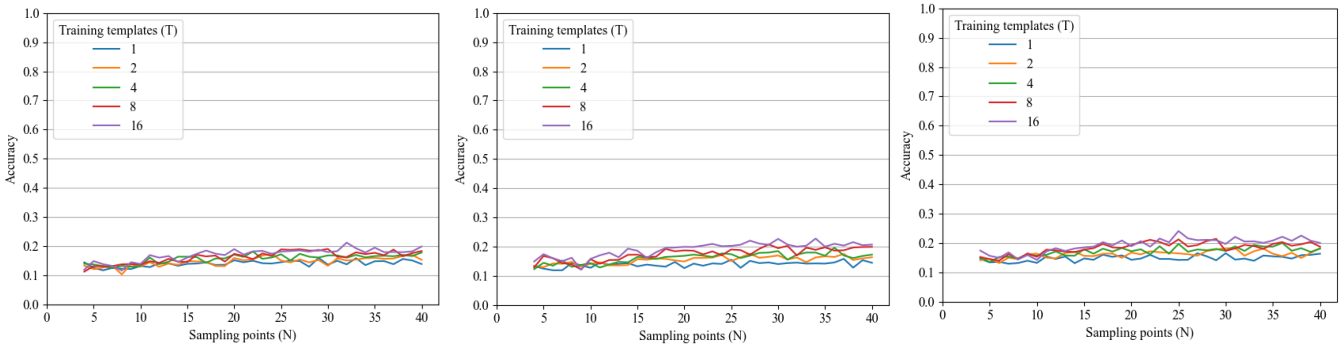


(a) Wood plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12). (b) PVC plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12). (c) Glass plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12).

FIGURE 11. Accuracy of the recognizers with the filtering step data in the user-dependent scenarios with the best *AP* value for gestures performed in front of the three plates.

each plate, the accuracy is lower with the filtering step data for the wood plate and the PVC plate than it was with the background subtraction data. This is due to the dimensionality reduction operated by the processing pipeline. However, for the glass plate, we obtained similar results, meaning that the processing pipeline was still able to extract relevant distance and permittivity values from a signal of lower quality. Regarding the execution time, it is faster for all the plates, again because of the dimensionality reduction.

In the user-independent scenario, the overall accuracy is once again low. The highest accuracy obtained for gestures performed in front of the wood plate is 21.3% with an execution time of 1.69 ms with the configuration T=16, N=32, and AP=(1,2,3,4,5,6,7,8,9,10,11,12). With the PVC plate, the highest accuracy is 22.8% with an ART of 1.35 ms with the configuration T=16, N=34, and AP=(1,2,3,6,8,9). With the glass plate, the highest accuracy is 24.2% with an execution time of 0.83 ms with the configuration T=16,



(a) Wood plate, AP=(1,2,3,4,5,6,7,8,9,10,11,12). (b) PVC plate, AP=(1,2,3,6,8,9). (c) Glass plate, AP=(4,5,7,10,11,12).

FIGURE 12. Accuracy of the recognizers with the filtering step data in the user-independent scenarios with the best AP value for gestures performed in front of the three plates.

N=25, and AP=(4,5,7,10,11,12). Fig. 10 shows the complete accuracy results for the best set of antenna pairs for each plate.

As expected, the accuracy in the user-independent scenario is lower than in the user-dependent one. However, we still see some similarities in the impact of the different parameters. The accuracy still increases on average with the number of training templates. Increasing the number of sampling points also has the same effect as in the user-dependent scenario, but the increase in accuracy is way slower. Concerning the set of antenna pairs parameter, as the resulting accuracy remains very similar for the 3 different AP values, it is not possible to confirm that one set has a significant advantage over the others. If we compare the results obtained with the background subtraction data in the user-independent scenario, we see that the accuracy with the filtering step data is lower for the wood plate but higher for the PVC plate and the glass plate.

B. PREFERENCE

To investigate the relationship between performance (Section V-A) and preference (Section III), we conducted a preference-performance analysis, an instance of the Importance-Performance Analysis (IPA) [93], in which we assign each of the nine gesture classes to four different quadrants according to a differentiation provided by the coordinate origin as regards the average value of each variable, which is represented by plain orange lines in Fig. 14. This figure shows the results of the analysis for the performance expressed by its recognition rate and the preference expressed by the corresponding agreement rate obtained from the GES. Fig. 14 shows the results for the wood plate according to the three measures obtained in the GES: by agreement rate, by thinking time, and by goodness-of-fit. The green areas contain the gestures that are both admissible for performance and preference. Similar figures can be obtained for the other conditions (see our companion website).

VI. DISCUSSION AND IMPLICATIONS FOR DESIGN

This section discusses the results of the testing presented in the previous section and their implications for the design of

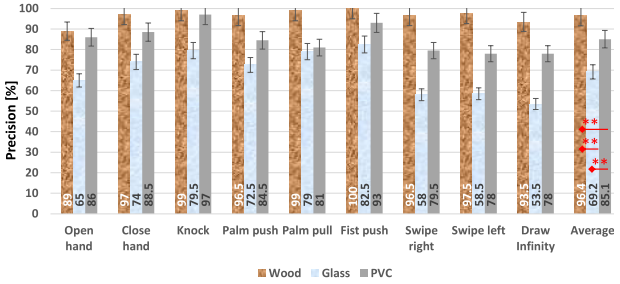


FIGURE 13. Precision of the gesture recognition per material: wood in brown, PVC in light gray, and glass in light blue. Error bars show a percentage of 5%.

radar-based gesture interaction. The results presented in the previous sections are summarized in Table 2 and Table 3 respectively. From the results in Table 2, we observe that performing hand gesture recognition through the wood plate with the Walabot in a user-dependent scenario and with the data from the background subtraction is a viable solution with appropriate Jackknife recognizer parameters. Fig. 13 shows the precision for nine gestures through the three materials. A Wilcoxon signed-ranked test for paired samples suggests that the wood condition is significantly more precise than the glass (z-score=2.75, p=.0029**) and than the PVC (z-score=2.74, p=.0059**) both with a large effect size (r=.61). Similarly, the PVC is significantly more precise than the glass (z-score=2.75, p=.0059**) with a large effect size (r=.6). Concerning the PVC and the glass plates, the accuracy obtained in the user-dependent scenario with the background subtraction is lower, and probably too low to ensure a good user experience in a real-life situation. Each radar gesture univocally appears in one of the four quadrants as follows:

1) *Quadrant 1* (Q1, in green) corresponds to high-performance values and high-preference values for radar gestures. It is labeled ''Keep up the good work'' because the gestures contained in this quadrant are positively assessed by the participants, who acknowledge both their preference and their performance. Designers

TABLE 2. Highest accuracy obtained for the 3 plates with the background subtraction step data in the user-dependent and the user-independent scenarios. Acc=accuracy [%], Exec. t.=average execution time [ms], *T*=number of training templates, *N*=number of sampling points, *AP*=set of antenna pairs.

	Material	Acc [%]	Exec. t. [ms]	*T*	*N*	*AP*
User-dependent	Wood	96.1	5.05	4	33	(1,2,3,6,8,9)
	PVC	83.7	5.72	4	38	(4,5,7,10,11,12)
	Glass	67.0	8.47	4	33	(1,2,3,4,5,6,7,8,9,10,11,12)
User-independent	Wood	31.5	23.63	16	40	(1,2,3,6,8,9)
	PVC	20.2	21.45	16	36	(4,5,7,10,11,12)
	Glass	19.8	26.97	16	29	(1,2,3,4,5,6,7,8,9,10,11,12)

TABLE 3. Highest accuracy obtained for the 3 plates with the filtering step data in the user-dependent and the user-independent scenarios. Acc=accuracy [%], Exec. t.=average execution time [ms], *T*=number of training templates, *N*=number of sampling points, *AP*=set of antenna pairs.

	Material	Acc [%]	Exec. t. [ms]	*T*	*N*	*AP*
User-dependent	Wood	66.2	0.38	4	29	(1,2,3,4,5,6,7,8,9,10,11,12)
	PVC	61.2	0.34	4	27	(1,2,3,4,5,6,7,8,9,10,11,12)
	Glass	66.3	0.26	4	21	(1,2,3,4,5,6,7,8,9,10,11,12)
User-independent	Wood	21.3	1.69	16	32	(1,2,3,4,5,6,7,8,9,10,11,12)
	PVC	22.8	1.35	16	34	(1,2,3,6,8,9)
	Glass	24.2	0.83	16	25	(4,5,7,10,11,12)

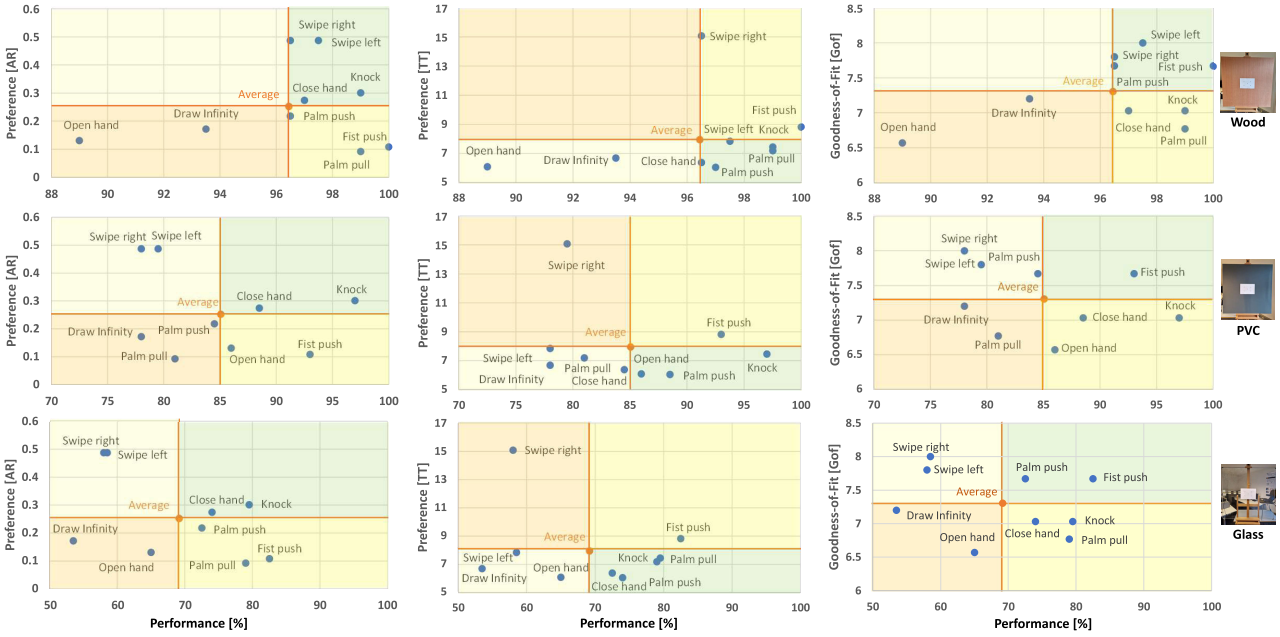


FIGURE 14. Preference-performance analysis for the three materials wood (top), PVC (middle), and glass (bottom) by agreement rate (left), by thinking time (center), and by goodness-of-fit (right).

are encouraged to maintain current strategies for these variables and retain these gestures as ideal candidates. For example, if we want to consider the agreement rate, then Fig. 14-a suggests "Swipe left", "Swipe right", "Close hand" and "Knock" as the gestures maximizing both criteria. However, if we consider the average thinking time, then Fig. 14-a suggests keeping the most performing gestures having the lowest thinking times, which are "Swipe left", Close hand", "Palm push", "Palm pull", and "Knock". These suggestions are not the same as for the agreement rate. If we consider the goodness-of-fit, Fig. 14-c suggests another set.

2) *Quadrant 2* (Q2, in yellow) corresponds to high-performance values and low-preference values for gestures. It is labeled "Concentrate here" because the gestures that fall into this quadrant should receive the highest priority. Action should be taken to address the challenges raised by these gestures because the participants performed well with them in principle but would not prefer to use them for their purposes. For example, the participants do not like to be forced to operate efficiently in a continuous manner.
3) *Quadrant 3* (Q3, in orange) corresponds to low-performance values and low-preference values for gestures. It is labeled "Low priority" because the

gestures located in this quadrant are not essential, do not perform well, and could be discontinued in the design process without any detriment.

4) *Quadrant 4* (Q4, in ivory) corresponds to low-performance values and high-preference values for gestures. It is labeled "Possible overkill" because the participants do not perform well with these gestures, yet perceive them to be highly preferred.

VII. LIMITATIONS

Our study has inevitably some limitations that are inherent to the various decisions made during the process. While three real-world scenarios initiated three gesture elicitation studies, the gestures collected from these studies were acquired in a simulated environment, which reflects only partially real-world environments. Considering other scenarios, especially for exceptional cases (*e.g.,* in hidden, dark, tense conditions), would probably lead to other gestures. Limitations are also inherent to the protocol of the gesture elicitation studies, which are subject to the legacy bias [77]. Replications and other protocols would certainly make the vocabulary of gestures more solid, perhaps also tested in other conditions than an elicitation study.

A. LIMITATIONS INDUCED BY THE RADAR USED

Although we selected the nine gesture classes from the elicitation studies to come up with their agreement rate, thinking time, and goodness-of-fit, we could have identified other sets of gestures, supersets, or sub-sets, and reproduced the rest of the study in the same way. The bandwidth of the Walabot version used in this paper is the narrowest of all Walabot versions, resulting in a lower resolution. Changing to a radar with a larger bandwidth should in principle improve the signal quality. Furthermore, while the Walabot features 18 antennas, three of them cannot be selected and the set of available antenna pairs is limited. Radars with a more flexible way of configuring the antennas as emitters and receivers would be appreciated. The Walabot suffers from a relatively low signal-to-noise ratio and interferences between the emitter and receiver antennas, but the radar equation removes these interferences.

In this study, we investigated the interaction of a high-frequency radar, specifically the Walabot operating in the C-band, with various non-metallic materials, including glass, PVC, and wood, to assess their impact on hand gesture recognition for human-machine interfaces. Our advanced electromagnetic modeling, supported by the full-wave radar equation, successfully mitigates the effects of these materials on the signal processing. However, it is important to note that the material's thickness and permittivity can still influence the signal-to-noise ratio, which is crucial for accurate gesture recognition.

While the detrimental effect of material characteristics on the radar signal increases with higher permittivity and thickness, most common non-metallic materials exhibit relatively low permittivity. This range typically falls between 2 and 8 for solid dry materials, making our results with glass, PVC, and wood broadly representative. However, materials with high water content or those that are intrinsically metallic should be avoided, as they can significantly attenuate or completely reflect radar waves, respectively. For instance, materials like drywall and certain plastics, with relative permittivities within the aforementioned range, should be compatible with our method. Liquids should be avoided as their permittivity is usually significant.

It is also crucial to consider the electrical conductivity and dielectric losses at C-band frequencies. Materials with high conductivity or significant dielectric losses can adversely affect the radar's performance due to wave attenuation. Thus, our methodology is most applicable to materials that are not only non-metallic but also have low to moderate permittivity and minimal dielectric losses in the relevant frequency range. For example, materials such as polystyrene (relative permittivity 2.6) or fiberglass (relative permittivity 4-6) could be suitable candidates for further exploration in similar applications

B. LIMITATIONS INDUCED BY THE RECOGNIZER

Then, important limitations are inherent to the choice of the recognizer and the number of templates: we selected Jacknife [51], one of the most powerful template-based recognizer that is modality agnostic and that gets excellent recognition rates in general, even with few templates. Other template-based recognizers [94], [95], [96] and other ML/DL algorithms should be used instead to determine the influence of their approaches on the recognition rate and execution time. The main reason why we kept a template-based recognizer was its ability to produce a customizable gesture-based interface: any modification of the templates does not affect the recognizer. So, user-defined gestures are welcome in this vein, which has been promoted by Mauro et al. [12].

VIII. CONCLUSION AND FUTURE WORK

Our initial mission statement was to determine to what extent a commodity radar, such as the Walabot, could serve as an effective and efficient sensing device for recognizing hand gestures across different materials, in our case glass, wood, and PVC. We also set ourselves the goal of establishing a certain balance between gesture recognition performance (which is widely reported in studies with radars) and the preference expressed by end users for these gestures (which is rarely investigated to date). To this end, we used three real-world scenarios involving these materials to demonstrate the radar's singular and original character: a glass pane can get dirty after intense tactile interaction and is subject to the elements, whereas a radar located behind the pane is not; a wooden office door allows the radar to be located behind the door to prevent vandalism, deterioration, and video use; someone may find themselves in a critical situation behind a PVC wall and be able to communicate their situation to somebody on the other side of the wall.

These three scenarios enabled us to conduct three gesture elicitation studies, each with 30 different participants, which determined a vocabulary of radar gestures specific to each

scenario and identified nine classes of gestures commonly used with radar. Creating a classification of radar gestures would be welcome, highlighting radar-specific gestures not found in any other study, and gestures that are also used with other devices. This classification is underway, showing the intersections between studies, and is available on our companion website. For each class of gesture, each study has produced its agreement rate, thinking time, and goodness-of-fit, enabling us to express the preference for these gestures.

After identifying nine gesture classes from these three studies, we built up a new dataset that is publicly available on our companion website. More and more datasets of this type are appearing, which would make it possible to establish comparative testing between different algorithms based on the same set of datasets, something that has not yet been achieved in this area. To rigorously acquire the gestures of 20 subjects with the 9 gestures repeated 5 times over 3 materials, which generated 2,700 samples, we introduced a unique one-shot radar calibration method that avoids, once and for all, re-calibrating the radar each time it is used for a particular material. Thanks to this method, it is no longer necessary to re-calibrate the radar after this procedure. We have shown theoretically and empirically how this unique one-shot calibration can be achieved. In future work, we could pursue this approach by demonstrating this benefit under more variable experimental conditions. In particular, it would be interesting to show that the recognizer training performed with one material could be transposed, to some extent, to recognition with another material. Indeed, Green's function resulting from the analysis of radar signals acquired with different materials remains the same, independently of each material. This would satisfy the property of material-independent gesture recognition.

Next, we trained a template-based recognizer on the novel dataset to identify the configurations that lead to the best recognition rates for each material according to stage and according to scenario, user-dependent or user-independent. We found that wood achieved the best results, especially in the user-dependent scenario. The user-independent scenario achieved low recognition rates, mainly due to variations in the participants' articulations, but also to the fact that the recognizer used only worked on the basis of 5 templates per person, which imposed a strict constraint. If we relax this constraint, we lose the property of customization: any modification of the dataset templates, leaving the recognizer used unchanged (Jacknife is template-based only), would result in retraining of any other ML/DL algorithm, unless near-real-time retraining is possible. The average execution times obtained in our study never exceed 30 ms, which fully authorizes direct manipulation interaction. It would be advantageous to use well-established and tested ML algorithms, such as the spectogram [18], [40], instead of a template-based recognizer. In this case, we would certainly need more templates per user and per gesture (only 5 in our case).

Finally, we mapped recognition performance (expressed by its recognition rate per gesture class derived from the confusion matrix) to gesture preference (expressed by measures from elicitation studies) using a performance-preference analysis. This analysis enabled us to divide the 9 gestures studied into four quadrants and to identify the gestures that preserve the balance between performance and preference, as well as to see which gestures should be worked on in the future. The three elicitation studies are just the beginning, and more are needed to support a true classification. Similarly, we could either try to extend the nine gesture classes to see the impact of this extension or consider other sets of up to 10 gestures, to preserve the user's ability to memorize them [83]. Other materials could also be considered, especially those offering good permittivity and those reported as promising in the catalog of Solids on Soli [33].

OPEN SCIENCE

Our companion website provides the reader with useful resources, including the Walabot novel dataset acquired and its description (Section IV), and the data from the testing in all conditions and materials (Section V-A), along with additional screenshots showing the configurations and some open-source code. Furthermore, we also suggest a starting point for the classification of radar gestures based on the three gesture elicitation studies to be further expanded.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose suggestions helped improve and clarify this manuscript and also would like to thank the participants of the gesture elicitation studies reported in the article for their participation and Emile Giot for acquiring the dataset and its testing.

REFERENCES

[1] L. Chen, V. Varadan, C. K. Ong, and C. P. Neo, *Microwave Electronics: Measurement and Materials Characterization* (Technology & Engineering). Hoboken, NJ, USA: Wiley, 2004.

[2] H.-S. Yeo and A. Quigley, "Radar sensing in human–computer interaction," *Interactions*, vol. 25, no. 1, pp. 70–73, Dec. 2017, doi: 10.1145/3159651.

[3] Q. Memon, B. Wodajo, S. Tekleab, and E. Alshehi, "Detection of static and moving objects behind walls and surfaces—An experimental investigation," in *Proc. 9th Int. Conf. Comput. Artif. Intell.*, New York, NY, USA, Mar. 2023, pp. 8–12, doi: 10.1145/3594315.3594317.

[4] T. Li, X. Xiong, Y. Xie, G. Hito, X.-D. Yang, and X. Zhou, "Reconstructing hand poses using visible light," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 3, pp. 1–20, Sep. 2017, doi: 10.1145/3130937.

[5] M. Alloulah, A. Isopoussu, and F. Kawsar, "On indoor human sensing using commodity radar," in *Proc. ACM Int. Joint Conf. Int. Symp. Pervasive Ubiquitous Comput. Wearable Comput.*, New York, NY, USA, Oct. 2018, pp. 1331–1336, doi: 10.1145/3267305.3274180.

[6] M. Hoffman, P. Varcholik, and J. J. LaViola, "Breaking the status quo: Improving 3D gesture recognition with spatially convenient input devices," in *Proc. IEEE Virtual Reality Conf. (VR)*, Piscataway, NJ, USA, Mar. 2010, pp. 59–66, doi: 10.1109/VR.2010.5444813.

[7] I. Chatterjee, R. Xiao, and C. Harrison, "Gaze+gesture: Expressive, precise and targeted free-space interactions," in *Proc. ACM Int. Conf. Multimodal Interact.*, New York, NY, USA, Nov. 2015, pp. 131–138, doi: 10.1145/2818346.2820752.

[8] D. Avrahami, M. Patel, Y. Yamaura, and S. Kratz, "Below the surface: Unobtrusive activity recognition for work surfaces using RF-radar sensing," in *Proc. 23rd Int. Conf. Intell. User Interfaces*, New York, NY, USA, Mar. 2018, pp. 439–451, doi: 10.1145/3172944.3172962.

[9] R. Tchantchane, H. Zhou, S. Zhang, and G. Alici, "A review of hand gesture recognition systems based on noninvasive wearable sensors," *Adv. Intell. Syst.*, vol. 5, no. 10, Oct. 2023, Art. no. 2300207, doi: 10.1002/aisy.202300207.

[10] A. D. Berenguer, M. C. Oveneke, H.-U.-R. Khalid, M. Alioscha-Perez, A. Bourdoux, and H. Sahli, "GestureVLAD: Combining unsupervised features representation and spatio-temporal aggregation for Doppler-radar gesture recognition," *IEEE Access*, vol. 7, pp. 137122–137135, 2019.

[11] U. Kunwar, S. Borar, M. Berghofer, J. Kylmälä, I. Aslan, L. A. Leiva, and A. Oulasvirta, "Robust and deployable gesture recognition for smartwatches," in *Proc. 27th Int. Conf. Intell. User Interfaces*, New York, NY, USA, Mar. 2022, pp. 277–291, doi: 10.1145/3490099.3511125.

[12] G. Mauro, M. Chmurski, L. Servadei, M. Pegalajar-Cuellar, and D. P. Morales-Santos, "Few-shot user-definable radar-based hand gesture recognition at the edge," *IEEE Access*, vol. 10, pp. 29741–29759, 2022, doi: 10.1109/ACCESS.2022.3155124.

[13] X. Lou, Z. Yu, Z. Wang, K. Zhang, and B. Guo, "Gesture-radar: Enabling natural human–computer interactions with radar-based adaptive and robust arm gesture recognition," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2018, pp. 4291–4297, doi: 10.1109/SMC.2018.00726.

[14] S. Palipana, D. Salami, L. A. Leiva, and S. Sigg, "Pantomime: Mid-air gesture recognition with sparse millimeter-wave radar point clouds," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 1, pp. 1–27, Mar. 2021, doi: 10.1145/3448110.

[15] A. Sluÿters, S. Lambot, and J. Vanderdonckt, "Hand gesture recognition for an off-the-shelf radar by electromagnetic modeling and inversion," in *Proc. 27th Int. Conf. Intell. User Interfaces*, New York, NY, USA, Mar. 2022, pp. 506–522, doi: 10.1145/3490099.3511107.

[16] B. Zhang, L. Zhang, M. Wu, and Y. Wang, "Dynamic gesture recognition based on RF sensor and AE-LSTM neural network," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Piscataway, NJ, USA, May 2021, pp. 1–5, doi: 10.1109/ISCAS51556.2021.9401065.

[17] A. Ali, P. Parida, V. Va, S. Ni, K. N. Nguyen, B. L. Ng, and J. C. Zhang, "End-to-end dynamic gesture recognition using mmWave radar," *IEEE Access*, vol. 10, pp. 88692–88706, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9858146

[18] N. T. Attygalle, L. A. Leiva, M. Kljun, C. Sandor, A. Plopski, H. Kato, and K. Čopič Pucihar, "No interface, no problem: Gesture recognition on physical objects using radar sensing," *Sensors*, vol. 21, no. 17, p. 5771, Aug. 2021, doi: 10.3390/s21175771.

[19] J.-W. Choi, S.-J. Ryu, and J.-H. Kim, "Short-range radar based real-time hand gesture recognition using LSTM encoder," *IEEE Access*, vol. 7, pp. 33610–33618, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8662554

[20] P. Goswami, S. Rao, S. Bharadwaj, and A. Nguyen, "Real-time multi-gesture recognition using 77 GHz FMCW MIMO single chip radar," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Piscataway, NJ, USA, Jan. 2019, pp. 1–4, doi: 10.1109/ICCE.2019.8662006.

[21] Y. Sun, T. Fei, X. Li, A. Warnecke, E. Warsitz, and N. Pohl, "Real-time radar-based gesture detection and recognition built in an edge-computing platform," *IEEE Sensors J.*, vol. 20, no. 18, pp. 10706–10716, Sep. 2020, doi: 10.1109/JSEN.2020.2994292.

[22] R. Wang, S. Xiang, C. Feng, P. Wang, S. Ergan, and Y. Fang, "Through-wall object recognition and pose estimation," in *Proc. 36th Int. Symp. Autom. Robot. Construct.*, M. Al-Hussein, Ed., May 2019, pp. 1176–1183, doi: 10.22260/ISARC2019/0157.

[23] F. Li, Y. Li, B. Du, H. Xu, H. Xiong, and M. Chen, "A gesture interaction system based on improved finger feature and WE-KNN," in *Proc. 4th Int. Conf. Math. Artif. Intell.*, New York, NY, USA, Apr. 2019, pp. 39–43, doi: 10.1145/3325730.3325759.

[24] S. Hazra and A. Santra, "Short-range radar-based gesture recognition system using 3D CNN with triplet loss," *IEEE Access*, vol. 7, pp. 125623–125633, 2019, doi: 10.1109/ACCESS.2019.2938725.

[25] S. Ahmed, K. D. Kallu, S. Ahmed, and S. H. Cho, "Hand gestures recognition using radar sensors for human–computer-interaction: A review," *Remote Sens.*, vol. 13, no. 3, p. 527, Feb. 2021, doi: 10.3390/rs13030527.

[26] A. Sluÿters, S. Lambot, J. Vanderdonckt, and R.-D. Vatavu, "RadarSense: Accurate recognition of mid-air hand gestures with radar sensing and few training examples," *ACM Trans. Interact. Intell. Syst.*, vol. 13, no. 3, pp. 1–45, Sep. 2023, doi: 10.1145/3589645.

[27] A.-I. Şiean, C. Pamparău, A. Sluÿters, R.-D. Vatavu, and J. Vanderdonckt, "Flexible gesture input with radars: Systematic literature review and taxonomy of radar sensing integration in ambient intelligence environments," *J. Ambient Intell. Humanized Comput.*, vol. 14, no. 6, pp. 7967–7981, Jun. 2023, doi: 10.1007/s12652-023-04606-9.

[28] A.-I. Siean, C. Pamparău, and R.-D. Vatavu, "Scenario-based exploration of integrating radar sensing into everyday objects for free-hand television control," in *Proc. ACM Int. Conf. Interact. Media Exper.*, New York, NY, USA, Jun. 2022, pp. 357–362, doi: 10.1145/3505284.3532982.

[29] J. O. Wobbrock, M. R. Morris, and A. D. Wilson, "User-defined gestures for surface computing," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, R. B. Arthur, K. Hinckley, M. R. Morris, S. E. Hudson, and S. Greenberg, Eds., Apr. 2009, pp. 1083–1092, doi: 10.1145/1518701.1518866.

[30] H. Xia, M. Glueck, M. Annett, M. Wang, and D. Wigdor, "Iteratively designing gesture vocabularies: A survey and analysis of best practices in the HCI literature," *ACM Trans. Comput.-Hum. Interact.*, vol. 29, no. 4, pp. 1–54, May 2022, doi: 10.1145/3503537.

[31] F. R. Ortega, A. Galvan, K. Tarre, A. Barreto, N. Rishe, J. Bernal, R. Balcazar, and J.-L. Thomas, "Gesture elicitation for 3D travel via multi-touch and mid-air systems for procedurally generated pseudo-universe," in *Proc. IEEE Symp. 3D User Interface (3DUI)*, Mar. 2017, pp. 144–153, doi: 10.1109/3DUI.2017.7893331.

[32] H. Dong, A. Danesh, N. Figueroa, and A. E. Saddik, "An elicitation study on gesture preferences and memorability toward a practical hand-gesture vocabulary for smart televisions," *IEEE Access*, vol. 3, pp. 543–555, 2015, doi: 10.1109/ACCESS.2015.2432679.

[33] K. Čopič Pucihar, N. T. Attygalle, M. Kljun, C. Sandor, and L. A. Leiva, "Solids on soli: Millimetre-wave radar sensing through materials," *Proc. ACM Hum.-Comput. Interact.*, vol. 6, pp. 1–19, Jun. 2022, doi: 10.1145/3532212.

[34] Y. Dong and W. Qu, "Review of research on gesture recognition based on radar technology," in *Artificial Intelligence for Communications and Networks*, S. Shi, L. Ye, and Y. Zhang, Eds. Cham, Switzerland: Springer, 2021, pp. 390–403.

[35] S. Ahmed, D. Wang, J. Park, and S. H. Cho, "UWB-gestures, a public dataset of dynamic hand gestures acquired using impulse radar sensors," *Sci. Data*, vol. 8, no. 102, pp. 1–9, Apr. 2021, doi: 10.1038/s41597-021-00876-0.

[36] H. Cheng, L. Yang, and Z. Liu, "Survey on 3D hand gesture recognition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1659–1673, Sep. 2016, doi: 10.1109/TCSVT.2015.2469551.

[37] Z. Wang, F. Liu, X. Li, M. Ma, X. Feng, and Y. Guo, "A survey of hand gesture recognition based on FMCW radar," in *Proc. 8th Int. Conf. Commun. Inf. Process.*, New York, NY, USA, Nov. 2022, pp. 73–79, doi: 10.1145/3571662.3571674.

[38] D. Avrahami, M. Patel, Y. Yamaura, S. Kratz, and M. Cooper, "Unobtrusive activity recognition and position estimation for work surfaces using RF-radar sensing," *ACM Trans. Interact. Intell. Syst.*, vol. 10, no. 1, pp. 1–28, Aug. 2019, doi: 10.1145/3241383.

[39] E. Hayashi, J. Lien, N. Gillian, L. Giusti, D. Weber, J. Yamanaka, L. Bedal, and I. Poupyrev, "RadarNet: Efficient gesture recognition technique utilizing a miniature radar sensor," in *Proc. CHI Conf. Hum. Factors Comput. Syst.* New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–14. [Online]. Available: https://doi.org/10.1145/3411764.3445367

[40] L. A. Leiva, M. Kljun, C. Sandor, and K. C. Pucihar, "The wearable radar: Sensing gestures through fabrics," in *Proc. 22nd Int. Conf. Hum.-Comput. Interact. Mobile Devices Services*, New York, NY, USA, Oct. 2020, pp. 1–4, doi: 10.1145/3406324.3410720.

[41] J. Lien, N. Gillian, M. E. Karagozler, P. Amihood, C. Schwesig, E. Olson, H. Raja, and I. Poupyrev, "Soli: Ubiquitous gesture sensing with millimeter wave radar," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–19, Jul. 2016, doi: 10.1145/2897824.2925953.

[42] D. Kim and C. Harrison, "EtherPose: Continuous hand pose tracking with wrist-worn antenna impedance characteristic sensing," in *Proc. 35th Annu. ACM Symp. User Interface Softw. Technol.*, New York, NY, USA, Oct. 2022, doi: 10.1145/3526113.3545665.

[43] M. Yu, N. Kim, Y. Jung, and S. Lee, "A frame detection method for real-time hand gesture recognition systems using CW-radar," *Sensors*, vol. 20, no. 8, p. 2321, Apr. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/8/2321

[44] J. Paradiso, C. Abler, K.-Y. Hsiao, and M. Reynolds, "The magic carpet: Physical sensing for immersive environments," in *Proc. Extended Abstr. Hum. Factors Comput. Syst.*, New York, NY, USA, 1997, pp. 277–278, doi: 10.1145/1120212.1120391.

[45] S. Chioccarello, A. Sluÿters, A. Testolin, J. Vanderdonckt, and S. Lambot, "FORTE: Few samples for recognizing hand gestures with a smartphone-attached radar," *Proc. ACM Hum.-Comput. Interact.*, vol. 7, pp. 1–25, Jun. 2023, doi: 10.1145/3593231.
[46] H.-S. Yeo, G. Flamich, P. Schrempf, D. Harris-Birtill, and A. Quigley, "RadarCat: Radar categorization for input & interaction," in *Proc. 29th Annu. Symp. User Interface Softw. Technol.*, New York, NY, USA, Oct. 2016, pp. 833–841, doi: 10.1145/2984511.2984515.
[47] H.-S. Yeo, R. Minami, K. Rodriguez, G. Shaker, and A. Quigley, "Exploring tangible interactions with radar sensing," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 2, no. 4, pp. 1–25, Dec. 2018, doi: 10.1145/3287078.
[48] R.-D. Vatavu, L. Anthony, and J. O. Wobbrock, "Gestures as point clouds: A $P recognizer for user interface prototypes," in *Proc. 14th ACM Int. Conf. Multimodal Interact.* New York, NY, USA: ACM, Oct. 2012, pp. 273–280, doi: 10.1145/2388676.2388732.
[49] P. Wang, J. Lin, F. Wang, J. Xiu, Y. Lin, N. Yan, and H. Xu, "A gesture air-writing tracking method that uses 24 GHz SIMO radar SoC," *IEEE Access*, vol. 8, pp. 152728–152741, 2020.
[50] R. Brunelli, *Template Matching Techniques in Computer Vision: Theory and Practice*. New York, NY, USA: Wiley, 2009.
[51] E. M. Taranta II, A. Samiei, M. Maghoumi, P. Khaloo, C. R. Pittman, and J. J. LaViola Jr., "Jackknife: A reliable recognizer with few samples and many modalities," in *Proc. CHI Conf. Hum. Factors Comput. Syst.* New York, NY, USA, May 2017, pp. 5850–5861, doi: 10.1145/3025453.3026002.
[52] R.-D. Vatavu, "The effect of sampling rate on the performance of template-based gesture recognizers," in *Proc. 13th Int. Conf. Multimodal Interfaces*, New York, NY, USA, Nov. 2011, pp. 271–278, doi: 10.1145/2070481.2070531.
[53] R.-D. Vatavu, "The impact of motion dimensionality and bit cardinality on the design of 3D gesture recognizers," *Int. J. Hum.-Comput. Stud.*, vol. 71, no. 4, pp. 387–409, Apr. 2013, doi: 10.1016/j.ijhcs.2012.11.005.
[54] Z. Chen, T. Zheng, and J. Luo, "Octopus: A practical and versatile wideband MIMO sensing platform," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, Oct. 2021, pp. 601–614, doi: 10.1145/3447993.3483267.
[55] S. Wang, J. Song, J. Lien, I. Poupyrev, and O. Hilliges, "Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum," in *Proc. 29th Annu. Symp. User Interface Softw. Technol.*, New York, NY, USA, Oct. 2016, pp. 851–860.
[56] N. Magrofuoco, J.-L. Pérez-Medina, P. Roselli, J. Vanderdonckt, and S. Villarreal, "Eliciting contact-based and contactless gestures with radar-based sensors," *IEEE Access*, vol. 7, pp. 176982–176997, 2019, doi: 10.1109/ACCESS.2019.2951349.
[57] A. Sluÿters, Q. Sellier, J. Vanderdonckt, V. Parthiban, and P. Maes, "Consistent, continuous, and customizable mid-air gesture interaction for browsing multimedia objects on large displays," *Int. J. Hum.–Comput. Interact.*, vol. 39, no. 12, pp. 2492–2523, Jul. 2023, doi: 10.1080/10447318.2022.2078464.
[58] C. Huesser, S. Schubiger, and A. Çöltekin, "Gesture interaction in virtual reality," in *Proc. IFIP TC 13 Conf. Hum.–Comput. Interact.*, C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen, Eds. Cham, Switzerland: Springer, 2021, pp. 151–160, doi: 10.1007/978-3-030-85613-7_11.
[59] F. Soldovieri, O. Lopera, and S. Lambot, "Combination of advanced inversion techniques for an accurate target localization via GPR for demining applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 451–461, Jan. 2011, doi: 10.1109/TGRS.2010.2051675.
[60] Y. Huang, Z. Tian, and Q. Jiang, "A radar and monocular camera-based fusion approach for pedestrian detection," in *Proc. 2nd Int. Conf. Comput. Data Sci. (CONF-CDS)*. New York, NY, USA: Association for Computing Machinery, 2021, pp. 1–7. [Online]. Available: https://doi.org/10.1145/3448734.3450461
[61] G. Agresti and S. Milani, "Material identification using RF sensors and convolutional neural networks," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 3662–3666.
[62] R. N. Khushaba and A. J. Hill, "Radar-based materials classification using deep wavelet scattering transform: A comparison of centimeter vs. millimeter wave units," *IEEE Robot. Autom. Lett.*, vol. 7, no. 2, pp. 2016–2022, Apr. 2022, doi: 10.1109/LRA.2022.3143200.
[63] K. Čopič Pucihar, C. Sandor, M. Kljun, W. Huerst, A. Plopski, T. Taketomi, H. Kato, and L. A. Leiva, "The missing interface: Micro-gestures on augmented objects," in *Proc. Extended Abstr. CHI Conf. Human Factors Comput. Syst.*, New York, NY, USA, May 2019, pp. 1–6, doi: 10.1145/3290607.3312986.
[64] M. Sato, I. Poupyrev, and C. Harrison, "Touché: Enhancing touch interaction on humans, screens, liquids, and everyday objects," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, May 2012, pp. 483–492, doi: 10.1145/2207676.2207743.
[65] W. Han, S. Dai, and M. R. Yuce, "Real-time contactless respiration monitoring from a radar sensor using image processing method," *IEEE Sensors J.*, vol. 22, no. 19, pp. 19020–19029, Oct. 2022.
[66] S. Lambot, E. C. Slob, I. van den Bosch, B. Stockbroeckx, and M. Vanclooster, "Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 11, pp. 2555–2568, Nov. 2004, doi: 10.1109/TGRS.2004.834800.
[67] A. De Coster and S. Lambot, "Fusion of multifrequency GPR data freed from antenna effects," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 2, pp. 664–674, Feb. 2018, doi: 10.1109/JSTARS.2018.2790419.
[68] L. Spassova, R. Wasinger, J. Baus, and A. Kruger, "Product associated displays in a shopping scenario," in *Proc. 4th IEEE ACM Int. Symp. Mixed Augmented Reality (ISMAR)*, 2005, pp. 210–211, doi: 10.1109/ISMAR.2005.46.
[69] S. Longo, E. Kovacs, J. Franke, and M. Martin, "Enriching shopping experiences with pervasive displays and smart things," in *Proc. ACM Conf. Pervasive Ubiquitous Comput.*, New York, NY, USA, Sep. 2013, pp. 991–998, doi: 10.1145/2494091.2496013.
[70] M. Cai, S. Masuko, and J. Tanaka, "Gesture-based mobile communication system providing side-by-side shopping feeling," in *Proc. 23rd Int. Conf. Intell. User Interfaces Companion (IUI)*. New York, NY, USA: Association for Computing Machinery, 2018, pp. 1–2. [Online]. Available: https://doi.org/ 10.1145/3180308.3180310
[71] K. Cheverst, A. Dix, D. Fitton, A. Friday, and M. Rouncefield, "Exploring the utility of remote messaging and situated office door displays," in *Human-Computer Interaction With Mobile Devices and Services*, L. Chittaro, Ed. Berlin, Germany: Springer, 2003, pp. 336–341, doi: 10.1007/978-3-540-45233-1_24.
[72] Md. Z. Uddin, F. M. Noori, and J. Torresen, "In-home emergency detection using an ambient ultra-wideband radar sensor and deep learning," in *Proc. IEEE Ukrainian Microw. Week (UkrMW)*, Sep. 2020, pp. 1089–1093, doi: 10.1109/UkrMW49653.2020.9252708.
[73] M. Mercuri, P. Karsmakers, B. Vanrumste, P. Leroux, and D. Schreurs, "Biomedical wireless radar sensor network for indoor emergency situations detection and vital signs monitoring," in *Proc. IEEE Topical Conf. Biomed. Wireless Technol., Netw., Sens. Syst. (BioWireleSS)*, Jan. 2016, pp. 32–35, doi: 10.1109/BIOWIRELESS.2016.7445554.
[74] L. Santana, A. P. Rocha, A. Guimarães, I. C. Oliveira, J. M. Fernandes, S. Silva, and A. Teixeira, "Radar-based gesture recognition towards supporting communication in aphasia: The bedroom scenario," in *Mobile and Ubiquitous Systems: Computing, Networking and Services*, T. Hara and H. Yamaguchi, Eds. Cham, Switzerland: Springer, 2022, pp. 500–506, doi: 10.1007/978-3-030-94822-1_30.
[75] M. Korkman, U. Kirk, and S. Kemp, *NEPSY: A Developmental Neuropsychological Assessment*. San Antonio, TX, USA: Psychol. Corp., 1998.
[76] R.-D. Vatavu and J. O. Wobbrock, "Formalizing agreement analysis for elicitation studies: New measures, significance test, and toolkit," in *Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst.* New York, NY, USA: ACM, Apr. 2015, pp. 1325–1334, doi: 10.1145/2702123.2702223.
[77] M. R. Morris, A. Danielescu, S. Drucker, D. Fisher, B. Lee, M. C. Schraefel, and J. O. Wobbrock, "Reducing legacy bias in gesture elicitation studies," *Interactions*, vol. 21, no. 3, pp. 40–45, May 2014, doi: 10.1145/2591689.
[78] S. Villarreal-Narvaez, A.-I. Şiean, A. Sluÿters, R.-D. Vatavu, and J. Vanderdonckt, "Informing future gesture elicitation studies for interactive applications that use radar sensing," in *Proc. Int. Conf. Adv. Vis. Interfaces (AVI)*. New York, NY, USA: Association for Computing Machinery, 2022, pp. 1–3. [Online]. Available: https://doi.org/10.1145/3531073.3534475
[79] I.-A. Zaiţi, Ş.-G. Pentiuc, and R.-D. Vatavu, "On free-hand TV control: Experimental results on user-elicited gestures with leap motion," *Pers. Ubiquitous Comput.*, vol. 19, nos. 5–6, pp. 821–838, Aug. 2015, doi: 10.1007/s00779-015-0863-y.
[80] R. Aigner, D. Wigdor, H. Benko, M. Haller, D. Lindbauer, A. Ion, S. Zhao, and J. T. K. V. Koh, "Understanding mid-air hand gestures: A study of human preferences in usage of gesture types for HCI," Microsoft Res., Tech. Rep. MSR-TR-2012-111, Nov. 2012. [Online]. Available: https://www.microsoft.com/en-us/research/publication/understanding-mid-air-hand-gestures-a-study-of-human-preferences-in-usage-of-gesture-

[81] M. G. Amin, Z. Zeng, and T. Shan, "Hand gesture recognition based on radar micro-Doppler signature envelopes," in *Proc. IEEE Radar Conf. (RadarConf)*, Mar. 2019, pp. 1–6, doi: 10.1109/RADAR.2019.8835661.
[82] T. Zheng, Z. Chen, J. Luo, L. Ke, C. Zhao, and Y. Yang, "SiWa: See into walls via deep UWB radar," in *Proc. 27th Annu. Int. Conf. Mobile Comput. Netw.*, New York, NY, USA, Oct. 2021, pp. 323–336, doi: 10.1145/3447993.3483258.
[83] M. A. Nacenta, Y. Kamber, Y. Qiang, and P. O. Kristensson, "Memorability of pre-designed and user-defined gesture sets," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, New York, NY, USA, Apr. 2013, pp. 1099–1108, doi: 10.1145/2470654.2466142.
[84] J. Nielsen and J. Levy, "Measuring usability: Preference vs. performance," *Commun. ACM*, vol. 37, no. 4, pp. 66–75, Apr. 1994, doi: 10.1145/175276.175282.
[85] S. Zhu, J. Xu, H. Guo, Q. Liu, S. Wu, and H. Wang, "Indoor human activity recognition based on ambient radar with signal processing and machine learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Oct. 2018, pp. 1–6, doi: 10.1109/ICC.2018.8422107.
[86] S. Lambot and F. André, "Full-wave modeling of near-field radar data for planar layered media reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2295–2303, May 2014, doi: 10.1109/TGRS.2013.2259243.
[87] M. Heideman, D. Johnson, and C. Burrus, "Gauss and the history of the fast Fourier transform," *IEEE ASSP Mag.*, vol. M-1, no. 4, pp. 14–21, Oct. 1984. [Online]. Available: https://ieeexplore.ieee.org/document/1162257
[88] P. Sharma, B. Kumar, D. Singh, and S. P. Gaba, "Critical analysis of background subtraction techniques on real GPR data," *Defence Sci. J.*, vol. 67, no. 5, p. 559, Sep. 2017. [Online]. Available: https://core.ac.uk/download/pdf/333722781.pdf
[89] A. Dickman, *Inverse Fast Fourier Transform (IFFT)*. Cham, Switzerland: Springer, 2022, pp. 119–124, doi: 10.1007/978-3-030-93363-0_9.
[90] R. Phumvijit, P. Supanakoon, and S. Promwong, "Measurement scheme of radar cross section with time gating," in *Proc. 14th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTI-CON)*, Jun. 2017, pp. 822–825, doi: 10.1109/ECTICon.2017.8096365.
[91] F. Zhou, X. Li, and Z. Wang, "Efficiently user-independent ultrasonic-based gesture recognition algorithm," in *Proc. IEEE SENSORS*, Oct. 2019, pp. 1–4.
[92] I. Cui nas and M. G. Sánchez, "Permittivity and conductivity measurements of building materials at 5.8 GHz and 41.5 GHz," *Wireless Pers. Commun.*, vol. 20, no. 1, pp. 93–100, Jan. 2002, doi: 10.1023/A:1013886209664.
[93] J. A. Martilla and J. C. James, "Importance-performance analysis," *J. Mark*, vol. 41, no. 1, pp. 77–79, Jan. 1977, doi: 10.1177/002224297704100112.
[94] F. M. Caputo, P. Prebianca, A. Carcangiu, L. D. Spano, and A. Giachetti, "Comparing 3D trajectories for simple mid-air gesture recognition," *Comput. Graph.*, vol. 73, pp. 17–25, Jun. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0097849318300335
[95] F. M. Caputo, S. Burato, G. Pavan, T. Voillemin, H. Wannous, J. P. Vandeborre, M. Maghoumi, E. M. Taranta II, A. Razmjoo, J. J. LaViola Jr., F. Manganaro, S. Pini, G. Borghi, R. Vezzani, R. Cucchiara, H. Nguyen, M. T. Tran, and A. Giachetti, "Online gesture recognition," in *Proc. Eurograp. Workshop 3D Object Retr.*, S. Biasotti, G. Lavoué, and R. Veltkamp, Eds., 2019, pp. 93–102.
[96] A. Carcangiu and L. D. Spano, "G-gene: A gene alignment method for online partial stroke gestures recognition," *Proc. ACM Hum.-Comput. Interact.*, vol. 2, pp. 1–17, Jun. 2018, doi: 10.1145/3229095.

ARTHUR SLUŸTERS (Senior Member, IEEE) received the master's degree (magna cum laude) in computer science and engineering from Université Catholique de Louvain (UCLouvain), Belgium, in 2020, where he is currently pursuing the Ph.D. degree in computer science with Louvain Research Institute in Management and Organizations (LouRIM). His research interests include gesture interaction, radar-based gesture recognition, and software engineering.

SÉBASTIEN LAMBOT (Member, IEEE) received the M.Sc. and Ph.D. degrees in agricultural and environmental engineering from Université Catholique de Louvain (UCLouvain), Belgium, in 1999 and 2003, respectively. Following the Ph.D. degree, he was awarded a European Marie-Curie Post-Doctoral Fellowship and conducted research with Delft University of Technology, The Netherlands, from 2004 to 2005. From 2006 to 2012, he held joint appointments with UCLouvain, where he was appointed as a Professor and Forschungszentrum Jülich, Germany, where he was the Research Group Leader. He is currently a Full Professor (part-time) and the FNRS Research Director of UCLouvain. He has dedicated his research to advancing ground-penetrating radar (GPR) full-wave modeling and inversion for digital soil mapping and non-destructive testing. His research interests include remote sensing, hydrogeophysics, road inspection, and gesture recognition. His main fundamental contributions in this field include the development of full-wave far-field and near-field radar equations. He served as the General Chair for the Third International Workshop on Advanced Ground Penetrating Radar in 2005 and the 15th International Conference on Ground Penetrating Radar in 2014. More recently, he co-chaired IGARSS 2021. He has also edited several special issues in *Near Surface Geophysics*, *Vadose Zone Journal*, and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

JEAN VANDERDONCKT (Senior Member, IEEE) received the M.Sc. degree in mathematics, the M.Sc. degree in computer science, and the Ph.D. degree from the University of Namur, Belgium, in 1987, 1989, and 1997, respectively. He was a Visiting Associate Professor with Stanford University, in 2000. He is currently a Full Professor with Louvain School of Management, Université Catholique de Louvain, Belgium, where he has been leading the Louvain Interaction Laboratory, since 1998. His research interests include human–computer interaction (HCI), engineering interactive computing systems (EICS), intelligent user interfaces (IUI), usability engineering, and software engineering. He is an Associate Editor of *ACM Transactions on Interactive Intelligent Systems* and the Co-Editor-in-Chief of the Springer Series of HCI and the SpringerBriefs in HCI.

SANTIAGO VILLARREAL-NARVAEZ received the B.Sc. degree in computer science and the M.Sc. degree in web intelligence from Jean Monnet University, Saint-Étienne, France, in 2007 and 2010, respectively, and the Ph.D. degree from Université Catholique de Louain, Belgium, in 2023. He is currently a Postdoctoral Researcher with Namur Digital Institute (NADI), Université de Namur, Belgium. He has been a Professor with Universidad de Las Américas, teaching software development and has also collaborated with the Intelligent and Interactive Systems Laboratory (SI2 Lab-UDLA) on the ePHoRt Project. His research interests include human–computer interaction, gesture interaction, multimodal interfaces, and augmented reality.