An Electromagnetic Model and Inversion-based Approach for Radar Gesture Recognition

Arthur Sluÿters , *Student Member*, Sébastien Lambot, *Member, IEEE*, Nuwan T. Attygalle, and Jean Vanderdonckt, *Senior Member, IEEE*

Accepted Author Manuscript
Accepted for publication in *IEEE Transactions on Instrumentation and Measurement.*
DOI: 10.1109/TIM.2026.3729019

***Abstract*—Radars offer robust gesture recognition capabilities over traditional vision, capacitance, and ultrasound sensors, with flexibility in preserving user privacy, sensing in different lighting conditions, higher temporal resolution, and sensing through materials. However, raw radar signals are cluttered, difficult to interpret, and strongly dependent on the radar hardware and scene configuration, which often forces gesture recognizers to be retrained for each new radar and environment. To address these limitations, we build on a full-wave electromagnetic radar model and inversion-based approach to calibrate and normalize radar measurements into a physically meaningful Green's function that is independent of the radar hardware and background scene. In this paper, we (1) experimentally validate the radar system invariance of the model by comparing signals from two distinct radars observing the same scene and frequency range, demonstrating that their diverse raw signals align to each other after normalization; (2) analyze the sensitivity of the processing pipeline to the distance and angle of the target, revealing the physical limitations of the infinite-plane assumption for finite hands; and (3) empirically evaluate the method on a comprehensive dataset of 20 gestures in user-dependent, user-independent, and mixed scenarios. Our results achieve classification accuracies of 90.58%, 97.78%, and 98.54% for the full set, and subsets of 12 and 8 gestures, respectively, confirming the approach's robustness and flexibility.**

***Index Terms*—Gesture input, Gesture recognition, Hand gesture, Radar sensing.**

I. Introduction

GESTURE input enables end users to use body pose and movement to interact effectively and efficiently with a computer-based interactive system. Gestures range from symbolic stroke gestures performed on touch-enabled mobile devices [1], [2] to finger, hand, and arm movements sensed by wearables [3] to mid-air gestures to interact with large displays [4] to foot-based interaction [5] and whole-body gestures [6]. A gesture recognition system typically includes an *input sensor* (e.g., the accelerometer integrated into a smart ring [3]), a *gesture recognizer* (e.g., Nearest-Neighbor classification of whole-body movement with the Dynamic Time Warping dissimilarity [6]), and a *gesture set* containing classes of gesture templates used for recognition. Each gesture class is assigned to a system function, such as opening the hand as a blossom to open a file [7] or pinching in/out to zoom in/out a picture.

Input sensors typically fall into six categories: touch surfaces for contact-based or close-to-surface touch gestures [8], computer vision techniques for image-based gesture recognition [9], [10], wearables [11], radars [12], Lidars [13], and other types of signals (e.g., Wi-Fi signals [14] and radio frequency identification (RFID) [15]). Among sensors detecting contactless gestures [15], [16], radars offer unique opportunities [17], including vital sign detection [18]–[21], finger-level gesture recognition [22], [23], on-body input [24], tangible interaction involving physical objects [25], [26], through-the-wall sensing [27]–[30], automotive perception [31], and activity recognition [30], [32].

Machine Learning (ML) and Deep Learning (DL) techniques are particularly suited for representing radar signals as images on multiple layers of abstraction, on which processing techniques are applied to accurately recognize the gestures from the radar signals [33]–[36].

Many radar-based gesture recognition systems use image-based representations, such as range-Doppler [37], [38], micro-Doppler [39], and range-Angle [37], [39], as inputs to deep neural networks (DNNs). For example, *Li et al.* [40] utilized micro-Doppler spectrograms for sign language classification, feeding them directly into a DNN for gesture classification. Similarly, *Wang et al.* [41] fused range-time, angle-time, and velocity-time spectrograms into a composite representation, which was input into a DNN for robust gesture classification. The *u-DeepHand* [42] explored denoised range spectrograms, using an auto-encoder to remove micro-motions from non-target body parts before feeding the spectrograms into a DNN. These image-based approaches achieve high accuracy but are computationally intensive and sensitive to radar configurations.

Instead of using dense image-based signal representations, some methods extract fine-grained features to reduce computational complexity. *Goswami et al.* [43] computed range-Doppler and range-Angle images but ex-

Manuscript received July 10, 2025; accepted September XX, YYYY. Arthur Sluÿters is funded by the Fonds de la Recherche Scientifique – FNRS under Grants n°40001931 and n°4001162. (Corresponding authors: Nuwan T. Attygalle; Jean Vanderdonckt.)

Arthur Sluÿters, Nuwan T. Attygalle (e-mail: nuwan.attygalle@uclouvain.be) and Jean Vanderdonckt (e-mail: jean.vanderdonckt@uclouvain.be) are with the Louvain Research Institute in Management and Organizations (LouRIM), Université catholique de Louvain (UCLouvain), 1348 Louvain-la-Neuve, Belgium.

Sébastien Lambot is with the Université catholique de Louvain (UCLouvain), 1348 Louvain-la-Neuve, Belgium.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TIM.2026.3729019

tracted features like reflected energy, range, velocity, and angle, inputting these concatenated features into a feed-forward neural network for gesture classification. Similarly, the *Pantomime* [44] used point-cloud data as input for a DNN, leveraging spatial feature representations for hand gesture recognition (HGR). *Wang et al.* [23] developed an interferometric Multiple-Input-Multiple-Output (MIMO) radar system using Texas Instruments' AWR1843 [1] radar to accurately sense, track, and recognize macro hand gestures (e.g., numbers 0–9, letters A–Z) and micro finger gestures (e.g., finger tap, thumb slide). Their method uses temporal and spatial interferometry to measure subtle range and angular displacements, creating 3-D trajectories (2D spatial coordinates plus time) for robust gesture recognition. Unlike traditional approaches that use micro-Doppler spectrograms, it feeds time-varying 3D coordinates into a ResNet50 [45] CNN for classification, reducing computational complexity and real-time processing.

However, the feature extraction methods described by *Goswami et al.* [43], *Pantomime* [44], and *Wang et al.* [23] are highly dependent on the radar configuration, antenna layout, and MIMO antenna configuration. For instance, *Wang et al.*'s digital signal processing (DSP) for extracting 3-D coordinates is sensitive to the radar's antenna array and millimeter-wave (mmWave) wave front configuration, as Angle of Arrival (AoA) estimation relies on the MIMO virtual array. Similarly, image-based methods [40]–[42] face challenges in generalizing across different radar systems or environmental conditions. These approaches are often tailored to specific radar types operating within particular frequency ranges, requiring large amounts of labeled data for a given context [23], [26], [46], [47]. As a result, gesture datasets recorded with one radar system are often incompatible with other systems, necessitating the creation of new datasets for each unique radar and environment rather than leveraging existing ones. This lack of adaptability hinders the development of robust radar-based gestural interfaces, as accuracy may degrade when using a different radar or operating in environments with varying backgrounds.

To address these limitations, this paper applies a prior electromagnetic model-based approach with inversion and a gesture recognizer [48], [49] to satisfy three key properties, which offer a unique combination:

- P_1. *Radar system invariance*: Once trained with a particular radar type that covers a given frequency range, our model-based approach remains applicable to any other radar type without retraining, provided that the same frequency range is preserved. For example, replacing the Walabot radar originally used for training a recognizer with another Walabot or radar operating in the same frequency range does not affect recognition. The approach does not depend on whether the radar is a time- or frequency-domain radar. However, differences in the radiation pattern may affect performance, as the model assumes a dipole-like antenna pattern.
- P_2. *Background scene invariance*: Through full-wave inversion and signal normalization, our gesture recognition approach relies solely on the estimated effective permittivity and distance of the moving hand. As a result, it is inherently insensitive to variations in the stationary background environment (e.g.,, walls, furniture), which are removed from the signal prior to recognition.
- P_3. *One-shot calibration*: Our approach relies on two distinct calibration steps. First, the radar is calibrated once, typically in factory settings, using our full-wave electromagnetic model based on far-field theory. This yields a normalized Green's function that effectively removes radar- and antenna-specific effects from the measured signal. Second, a gesture recognition model is trained on these normalized signals. As a result, the recognizer becomes independent of the specific radar unit: using a different radar of the same type (or operating in the same frequency range) does not require retraining or recalibration of the recognizer. A new radar would require a new electromagnetic calibration, but the resulting Green's function and signal characteristics would remain consistent. Recalibration of the radar may only be necessary in case of significant environmental changes such as temperature, humidity, or pressure variations.

The remainder of this paper is organized as follows. Section II explains the mathematical theory for full-wave electromagnetic modeling and inversion. Section III then describes the six stages of our model-based approach to radar signal processing. Two experiments are conducted in Section IV to evaluate the efficacy of this approach on a set of twenty gesture classes. Finally, Section V summarizes this paper and discusses avenues for future research.

II. Mathematical Theory

A. Full-wave Radar Model

The far-field full-wave radar equation [48] assumes a homogeneous field distribution for the back-scattered field over the antenna aperture, which holds when the target is far enough from the antenna. The linear block diagram in Fig. 1 describes the antenna using four characteristic global reflection (R) and transmission (T) coefficients: $R_i(\omega)$, $T_i(\omega)$, $T_s(\omega)$, and $R_s(\omega)$, where ω is the angular frequency and $a(\omega)$ and $b(\omega)$ are the source and returned signals in the frequency domain, respectively. Wave propagation from the radar antenna to the target is described using the 3D multilayered media Green function, defined as the x-directed backscattered electric field for a unit-strength electric source, both located at the antenna phase center [48], [50]–[52]. This equivalent medium is used for our hand gesture application, as an analytical solution of Maxwell's equations exists, while for real hands a 3D numerical solution would be needed, but would be impractical in terms of computing resources and too complex to

[1] https://www.ti.com/product/AWR1843

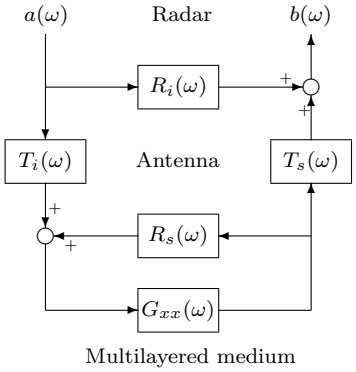


Fig. 1: Equivalent linear block diagram of the radar-antenna-medium corresponding to the far-field radar equation [48].

use. The equivalent linear block diagram shown in Fig. 1 is solved as follows:

$$b(\omega) = a(\omega)R_i(\omega) + \frac{a(\omega)T_i(\omega)G_{xx}(\omega)}{1 - R_s(\omega)G_{xx}(\omega)}T_s(\omega) \tag{1}$$

Defining the ratio $S(\omega) = b(\omega)/a(\omega)$, like the so-called S_{11} quantity measured by a Vector Network Analyzer (VNA), and the product $T(\omega) = T_i(\omega)T_s(\omega)$, Eq. (1) becomes [48]:

$$S(\omega) = R_i(\omega) + \frac{T(\omega)G_{xx}(\omega)}{1 - R_s(\omega)G_{xx}(\omega)} \tag{2}$$

For a pulse radar, the quantity $s(t)$ is measured with t being time and $b(\omega)$ is obtained using the Fourier transform:

$$b(\omega) = \mathcal{F}\{s(t)\} = \int_{-\infty}^{\infty} s(t)e^{-2\pi jt\omega}dt \tag{3}$$

Then, defining $H_i(\omega) = a(\omega)R_i(\omega)$, $H(\omega) = a(\omega)T(\omega)$ and $H_s(\omega) = R_s(\omega)$, Eq. (1) becomes:

$$\mathcal{F}\{s(t)\} = H_i(\omega) + \frac{H(\omega)G_{xx}(\omega)}{1 - H_s(\omega)G_{xx}(\omega)} \tag{4}$$

Whether the radar provides the ratio $b(\omega)/a(\omega)$ or the signal $b(\omega)$, the radar-antenna system is fully described by only three independent functions. $R_i(\omega)$ or $H_i(\omega)$ corresponds to a free-space measurement, for which $G_{xx}(\omega) = 0$, while $R_s(\omega)$ is responsible for the multiple reflections between the antenna and the target medium. Eq. (2) is analytically inverted to derive Green's function from the measurements, thereby filtering out the effects of the radar-antenna system and resulting in a normalized quantity that is independent of the radar in a given frequency range, i.e., the Green's function:

$$G_{xx}(\omega) = \frac{S(\omega) - R_i(\omega)}{T(\omega) + S(\omega)R_s(\omega) - R_i(\omega)R_s(\omega)} \tag{5}$$

A practical application of the far-field radar equation is to improve radar imaging, e.g., for landmine detection [53].

B. Determination of the Radar-antenna Characteristic Functions

The characteristic radar functions $R_i(\omega)$, $T(\omega)$ and $R_s(\omega)$ (or $H_i(\omega)$, $H(\omega)$ and $H_s(\omega)$) can be determined by solving a system of equations (2) containing at least three different model configurations (denoted k). In our practice, the antenna is placed at varying distances from a metal sheet, representing an infinite Perfect Electric Conductor (PEC). Knowing the properties of the medium, Green's functions are computed, and the corresponding $S_k(\omega)$ are measured. The system of equations can be expressed in a linear form by reformulating Eq. (2) assuming frequency dependence ω:

$$S_k = R_i + S_k G_{xx,k} R_s + G_{xx,k}\left(T - R_i R_s\right) \tag{6}$$

Hence, the resulting linear system of equations can be written in matrix form as:

$$\mathbf{b} = \mathbf{A}\mathbf{x} \tag{7}$$

where

$$\mathbf{b} = \begin{bmatrix} S_1 \\ \vdots \\ S_k \\ \vdots \\ S_n \end{bmatrix}, \mathbf{A} = \begin{bmatrix} 1 & S_1 G_{xx,1} & G_{xx,1} \\ \vdots & \vdots & \vdots \\ 1 & S_k G_{xx,k} & G_{xx,k} \\ \vdots & \vdots & \vdots \\ 1 & S_n G_{xx,n} & G_{xx,n} \end{bmatrix}, \mathbf{x} = \begin{bmatrix} R_i \\ R_s \\ T - R_i R_s \end{bmatrix} \tag{8}$$

The vector of unknowns is then determined using the least squares approach:

$$\mathbf{x} = \left(\mathbf{A}^{\mathrm{H}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{H}}\mathbf{b} \tag{9}$$

where symbol H denotes the Hermitian or conjugate transpose ($\mathbf{A}^{\mathrm{H}} \equiv \overline{\mathbf{A}^{T}}$). To apply the Green's function model, which assumes infinite layers, the metal sheet should be larger than five times the maximum wavelength, depending on the antenna radiation pattern. For smaller sizes, edge effects can affect the results.

C. Full-Wave Signal Inversion

Full-wave inversion can be used to determine the properties of the layered medium by minimizing an objective function formulated in the least squares sense as:

$$\phi(\mathbf{b}) = \left|\mathbf{G}^{*}_{\mathbf{xx}} - \mathbf{G}_{\mathbf{xx}}\right|^{T}\mathbf{C}^{-1}\left|\mathbf{G}^{*}_{\mathbf{xx}} - \mathbf{G}_{\mathbf{xx}}\right| \tag{10}$$

where $\mathbf{G}^{*}_{\mathbf{xx}} = G_{xx}(\omega)$ and $\mathbf{G}_{\mathbf{xx}} = G_{xx}(\omega, \mathbf{b})$ are the vectors containing the measured and simulated radar signals, respectively. The vector $\mathbf{b} = [\varepsilon_n, \sigma_n, h_n]$ with $n = 1, ..., N$ contains the unknowns, while $\mathbf{C}$ is the measurement error covariance matrix. The discrepancy between observed and modeled data is expressed by the amplitude of the errors in the complex plane, where the radar data vectors are arranged by frequency since the radar data are complex-valued. To focus on a particular range window, containing, for example the medium surface only [54], inversion is performed in the time domain. Reflections occurring after the reflection of interest are thereby ignored and the inverse model configuration can be simplified, e.g., reducing it to a half-space medium. To convert the frequency domain data into the time domain, the Inverse Fast Fourier Transform (IFFT) is applied:

$$g_{xx}(t) = \mathcal{F}^{-1}\{G_{xx}(\omega)\} = \int_{-\infty}^{\infty} G_{xx}(\omega)e^{2\pi j\omega t}d\omega \tag{11}$$

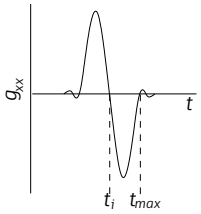


Fig. 2: Sketch of the radar signal in the time domain, $g_{xx}(t)$, with antenna effects removed. The time window for the inversion, denoted as t_{max}, is chosen to focus, as an example, on the surface reflection. t_i denotes the approximate time corresponding to the surface interface for Green's function.

Then, a maximum time t_{max} can be set to focus on the initial reflectors of interest, such as the surface reflection (Fig. 2). The objective function becomes:

$$\phi(\mathbf{b}) = \left(\mathbf{g}^*_{\mathbf{xx}} - \mathbf{g}_{\mathbf{xx}}\right)^T \mathbf{C}^{-1} \left(\mathbf{g}^*_{\mathbf{xx}} - \mathbf{g}_{\mathbf{xx}}\right) \tag{12}$$

where

$$\mathbf{g}^*_{\mathbf{xx}} = g^{\uparrow *}_{xx}(t)\big|_0^{t_{max}} \quad \text{and} \quad \mathbf{g}_{\mathbf{xx}} = g_{xx}(t, \mathbf{b})\big|_0^{t_{max}} \tag{13}$$

are the vectors containing, respectively, the observed and simulated time-domain windowed Green functions. Depending on the number of unknowns to estimate and the information available from the radar data, the topography of the objective function to be optimized can be quite complex, and, in some cases, the inverse problem can be ill-posed. When dealing with one or two unknowns, local optimization using a method such as the Levenberg-Marquardt Algorithm (LMA) [55] or a Look-Up Table (LUT) approach, which calculates the objective function throughout the parameter space, is reliable and efficient. However, for more complex problems, global optimization is usually required.

III. Radar Data Processing Strategy

Based on full-wave electromagnetic modeling and inversion (Section II), a six-stage pipeline processes the radar signals for gesture recognition (Fig. 3). Data from each stage can serve as input to gesture recognition: data from stages (b), (c), and (d) can be better suited for image processing algorithms, such as DL [33], while data from stages (e) and (f) could be used with template matching algorithms [56].

A. Raw Data Capture

Data is retrieved from one or more receiving radar antennas, either in the time or frequency domain, depending on the radar used (Fig. 3a). Raw data consist of the relevant signal with noise and parasitic reflections, which vary depending on the environment and on the radar system. To perform subsequent processing stages, the time-domain signal should first be converted to the frequency domain using a Fast Fourier Transform (FFT) Eq. 3) in order to apply the normalizing radar Eq. 5. The FFT yields a spectrum spanning from 0 to the Nyquist frequency. Within this spectrum, a specific sub-range is strategically chosen to strike an ideal balance between bandwidth and signal-to-noise ratio. In the context of pulse radar applications, it is common to narrow down this sub-range to a range approximately one fourth below the center frequency up to twice the center frequency.

B. Radar Source and Antenna Effects Removal

This second stage removes the effects of the radar source and antenna (except for radiation pattern) by applying the radar equation (Eq. 5) [48], [57] to the raw signal (Fig. 3b). These effects include internal reflections and transmissions between antennas, as well as antenna-target interactions. In the case of antenna arrays, this stage requires the generation of a far-field model of each emitter/receiver pair of antennas by computing their respective characteristic functions. The generation of the model is a one-shot process (property P_3) that consists of performing several measurements at different distances from a metal sheet (Section II-B). The far-field version of the radar model is chosen because it is much less resource-intensive to compute than the near-field version and is valid as long as users stay sufficiently far away from the radar, i.e., the distance to the target should be longer than the radar aperture size [58]. This stage produces a Green's function that is independent of the radar and thereby normalized in the operating frequency range of interest (property P_1) [59]. We experimentally evaluate the radar source and antenna effect removal. We provide the computational complexity of each stage for processing one radar frame based on the variable n, which is the size of the radar signal from a single receiving antenna. The size of the signal is affected by the sampling rate of the radar. The computational complexity of this stage is $\mathcal{O}(n)$.

C. Background Scene Removal

Based on the superposition principle, the reflections from static objects in the surrounding environment (e.g., walls, plants, furniture) are subtracted from the normalized signal (Fig. 3). This leaves only the reflections of moving objects and people, including the end user, in the resulting signal (property P_2), resulting in greater accuracy and efficiency of full-wave inversion. It is important to acknowledge that while the interactions between the static objects and the hand are not eliminated with the application of the superposition principle, they are anticipated to be minimal or negligible when the static objects are positioned farther away from the endpoint, as detailed in the following section. The computational complexity of this stage for processing one radar frame is $\mathcal{O}(n)$, where n is the size of the radar signal. Based on the superposition principle, the signal from the surrounding environment (e.g., walls, plants, furniture) is subtracted from the normalized signal. The resulting signal only contains the reflections of moving objects and persons, including the user, thus improving the accuracy and efficiency of full-wave inversion.

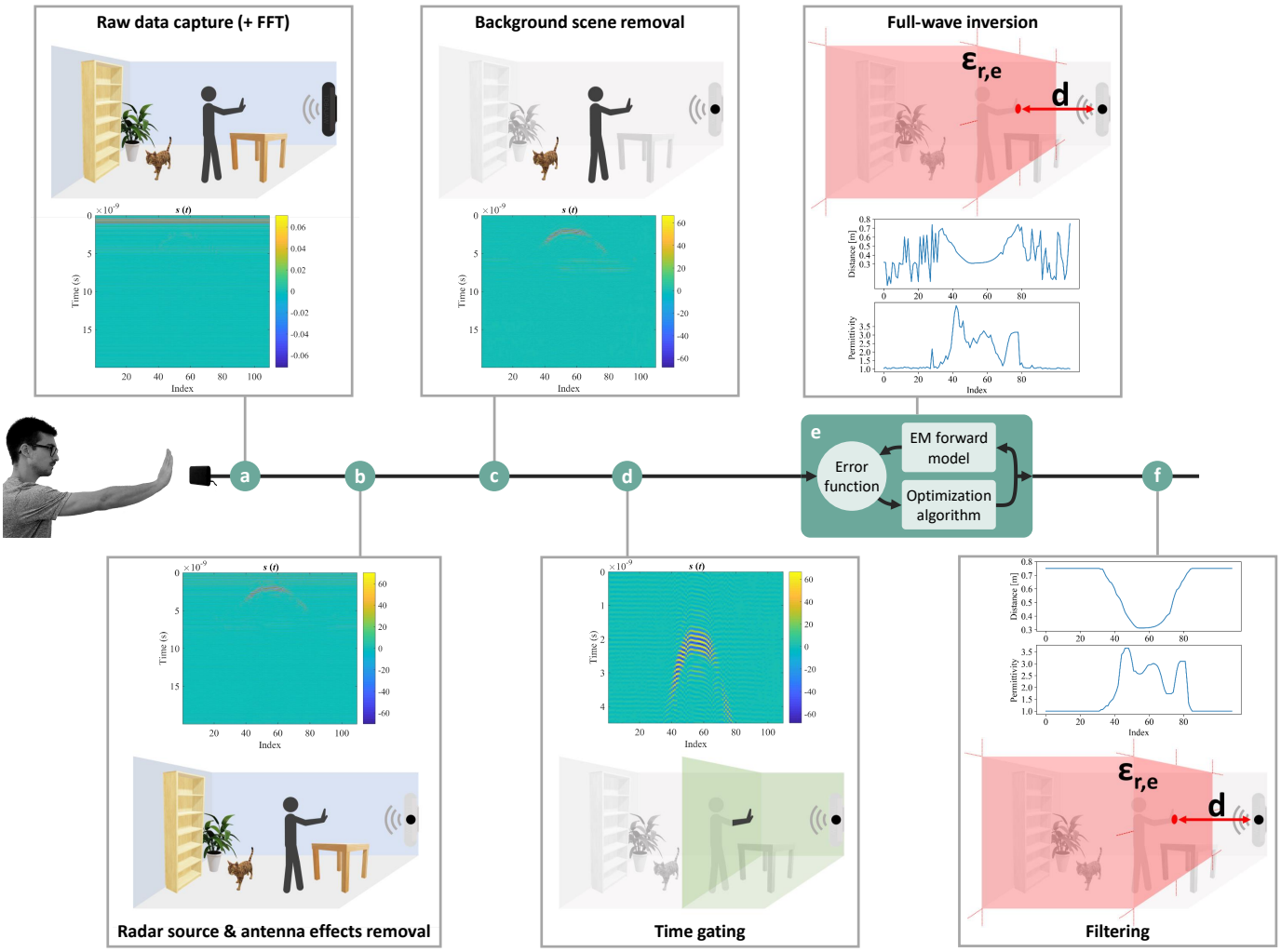


Fig. 3: Overview of the radar data processing strategy. For each stage, a visualization of its effect on the radar data and an example of the corresponding radar signal are provided. The radar signal is represented in the time domain for stages (a) to (d) and as time-distance and time-permittivity plots for stages (e) and (f).

D. Time Gating

An IFFT (Eq. 11) is first applied to the signal to convert it to the time domain (Fig. 3d), which simplifies subsequent computations. The time domain signal is then truncated (Fig. 3e) to exclude all reflections occurring outside some distance range of the radar (Fig. 2 and Eq. 13). The distance range is carefully selected to maximize the signal-to-noise ratio while maintaining enough distance between the user and the radar to ensure the validity of our far-field antenna model (Eq. 5). Consequently, movements occurring behind the user are ignored, thus improving the efficacy of the next stages and reducing the amount of data to be processed in the full-wave inversion, which increases its efficiency and enables faster real-time processing without compromising accuracy. It should be noted that this stage does not eliminate reflections from moving objects situated between the user and the radar, which could affect subsequent stages. However, such reflections are typically avoidable, and within the user's control. Because the truncated radar signal is copied into a new frame, the computational complexity of this stage for processing one radar frame is $\mathcal{O}(n)$, where n is the size of the truncated radar signal. The computational complexity could be reduced to a constant value ($\mathcal{O}(1)$) if the signal was truncated without copying it into a new frame.

E. Full-wave Inversion

The process of full-wave inversion reduces the radar signal to two physically-meaningful dimensions (Fig. 3e): the hand-radar distance d and the relative permittivity $\epsilon_{r,e}$ of an infinite half-space medium equivalent to the hand. The interface between the air and this medium is situated at a distance d from the radar and is orthogonal to the radar-hand direction. The obtained permittivity is an apparent value that remains unaffected by variations in the hand-radar distance. This apparent permittivity is determined by the quantity of signal reflected by the hand and should not be confused with its true permittivity. As such, it provides information on the pose of the hand, whereas the distance gives the dynamic position of the hand. The inversion algorithm iteratively adjusts the distance and apparent relative permittivity parameters of

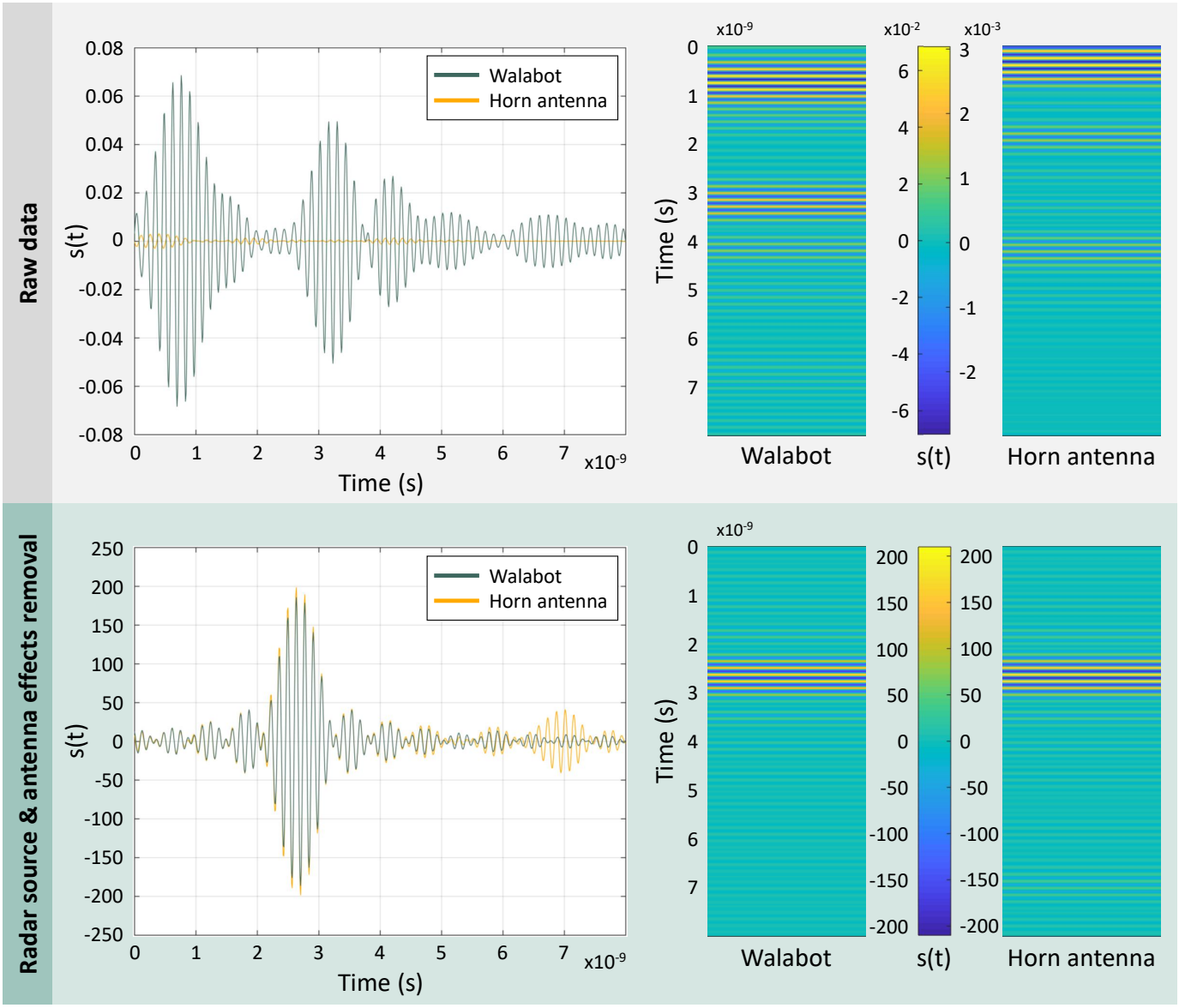


Fig. 4: Radar signal from two radars (Walabot and vector network analyzer + horn antenna) at 40 cm from a large copper plate. The top plots show two representations of the raw time-domain signal. The signal from both radars have different orders of magnitude and feature radar source and antenna effects, including internal reflections, making them difficult to compare. The bottom plots show how the signal from both radars is normalized after removing the radar source and antenna effects.

the model to find the best fit between the modeled and measured signals (Eq. 12). The computations are sped up by the use of a pre-calculated LUT for real-time objective function minimization. This drastic data reduction keeps essential information from the radar signal in the form of distance and permittivity values, thus enabling to use of less sophisticated algorithms, such as template-matching algorithms for gesture recognition. The computational complexity of this step is $\mathcal{O}(n \times k)$, where n is the size of the truncated radar signal and k is the size of the LUT. Despite the use of a LUT, this is the most expensive stage of the pipeline.

We conducted a small experiment to evaluate the sensitivity of the full-wave inversion process to three main variables: (1) the presence of strong reflectors in the

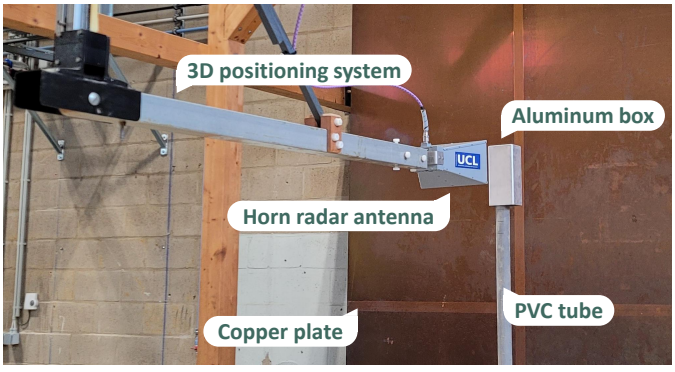


Fig. 5: Experimental setup for evaluating the sensitivity of the full-wave inversion process with respect to distance, angle, and background scene.

background, (2) the distance between the hand and the radar, from 30 cm to 110 cm, and (3) the angle between the hand and the radar, from 0° to 60° at a distance of 30 cm and from 0° to 45° at a distance of 50 cm. A custom radar system, comprising a vector network analyzer and a BBHA 9120 A horn antenna (0.8–5 GHz), was mounted on a 3D positioning system. The antenna exhibits high directivity, with a –3 dB beamwidth of approximately 45° in the E-plane and 30° in the H-plane at 1 GHz. The radar was directed toward a hollow aluminum box (21.6 cm height × 10 cm width × 8 cm depth) mounted on a PVC pole. A 3 m × 3 m copper plate placed on the wall 127.5 cm behind the aluminum box acted as a strong background reflector. Fig. 5 and 6 illustrate the experimental setup and the results of this experiment. The copper plate in the background did not have an impact on the inversion process as its reflections were removed in the time gating stage. Similarly, the inversion process correctly estimated the distance between the radar and the aluminum box up to 110 cm. Worse results should be expected when using radars with a lower signal-to-noise ratio and/or estimating the distance of less reflective objects. The maximum interaction distance should thus be adjusted accordingly.

The results of the inversion process when increasing the angle between the aluminum box and the radar antenna were consistent with its high directivity of 45°: the inversion process failed to correctly estimate the distance from 40° at 30 cm and 46° at 50 cm. The permittivity estimations changed drastically depending on the distance and angle between the aluminum box and the radar antenna, suggesting that this measure should not be used in scenarios where users are expected to move a lot around the radar. Although the EM model accounts for the decrease of signal strength with respect to distance (3D spherical divergence), these results make sense as the inversion process considers the hand to be an infinite medium to simplify computations. Because the hand is a finite medium, some of the radar signal does not interact with it, resulting in less signal being reflected than anticipated, which corresponds to a lower estimated relative permittivity. This phenomenon is emphasized when the hand is moved further away and/or more offset from the radar antenna. We believe that the model could be improved to account for the finite size of the hand, thus providing an estimation of relative permittivity that would be independent of the hand-radar distance.

F. Filtering

The result of the full-wave inversion depends on the signal-to-noise ratio, i.e., on the quality of the radar data. A low signal-to-noise ratio and moving objects behind or in front of the user may produce an incorrect estimation of the distance and apparent permittivity. Two strategies are employed to clean up the signal produced by the inversion stage (Fig. 3f). First, the distance and permittivity values are discarded and replaced with default values (d=75cm, $\epsilon_{r,e}$=1.0) when the estimated permittivity is close to the relative permittivity of the air, i.e., when no hand is detected in front of the radar. In addition, a moving median filter is applied to smooth out too-fast variations in the estimated distance and permittivity. The moving median filter was selected for its ability to filter out large and sudden variations in the data that would not be physically possible. Employing radars with a large dynamic range and leading to a higher signal-to-noise ratio and a wider field of view could decrease the need for filtering, as they would better dissociate the hand signal from the noise and hand reflections, even when the hand is not perfectly in front of the radar.

G. Interfacing with External Components and Comparison with Recent Radar Representations

The stages of our pipeline can interface with a variety of external components, including signal processing techniques and gesture recognition algorithms (see Table I).

Recent radar-based gesture-recognition pipelines can be broadly divided into dense map-based and sparse point- [44] or trajectory-based [49] approaches. Dense representations, including micro-Doppler spectrograms, range–Doppler maps, range–angle maps, and fused range–time, velocity–time, and angle–time representations, preserve rich spatiotemporal information and are well suited to convolutional and other deep neural networks [40]–[42], [60]–[64]. However, their dimensions and statistical distributions are determined by the radar waveform, bandwidth, sampling configuration, range and Doppler resolution, antenna array, and preprocessing chain [64] Consequently, a network trained on these representations can be sensitive to changes in radar hardware or acquisition parameters. Sparse point-cloud and trajectory approaches [23], [44] reduce the input volume and can support real-time or resource-constrained inference. Nevertheless, their outputs depend on target-detection thresholds, beamforming or angle-of-arrival estimation, MIMO array geometry, and the density and stability of valid target detections.

The proposed pipeline is complementary to both families of methods. Signals from stages (b)–(d) can be used to generate range–Doppler maps, range–angle maps, or point clouds after radar-source, antenna, stationary-background, and out-of-range components have been attenuated, thereby reducing the sensor- and scene-specific variability presented to a downstream recognizer. Stages (e)–(f) provide a different trade-off by compressing each radar frame to the hand–radar distance and apparent relative permittivity, enabling physically interpretable, low-dimensional template-based recognition. Unlike recent data-driven generalization strategies that enlarge the training distribution through synthetic radar-data generation [65], the present approach seeks generalization through physics-based calibration and normalization. This compression necessarily discards part of the spatiotemporal information exploited by dense deep-learning representations and remains subject to the far-field, finite-target,

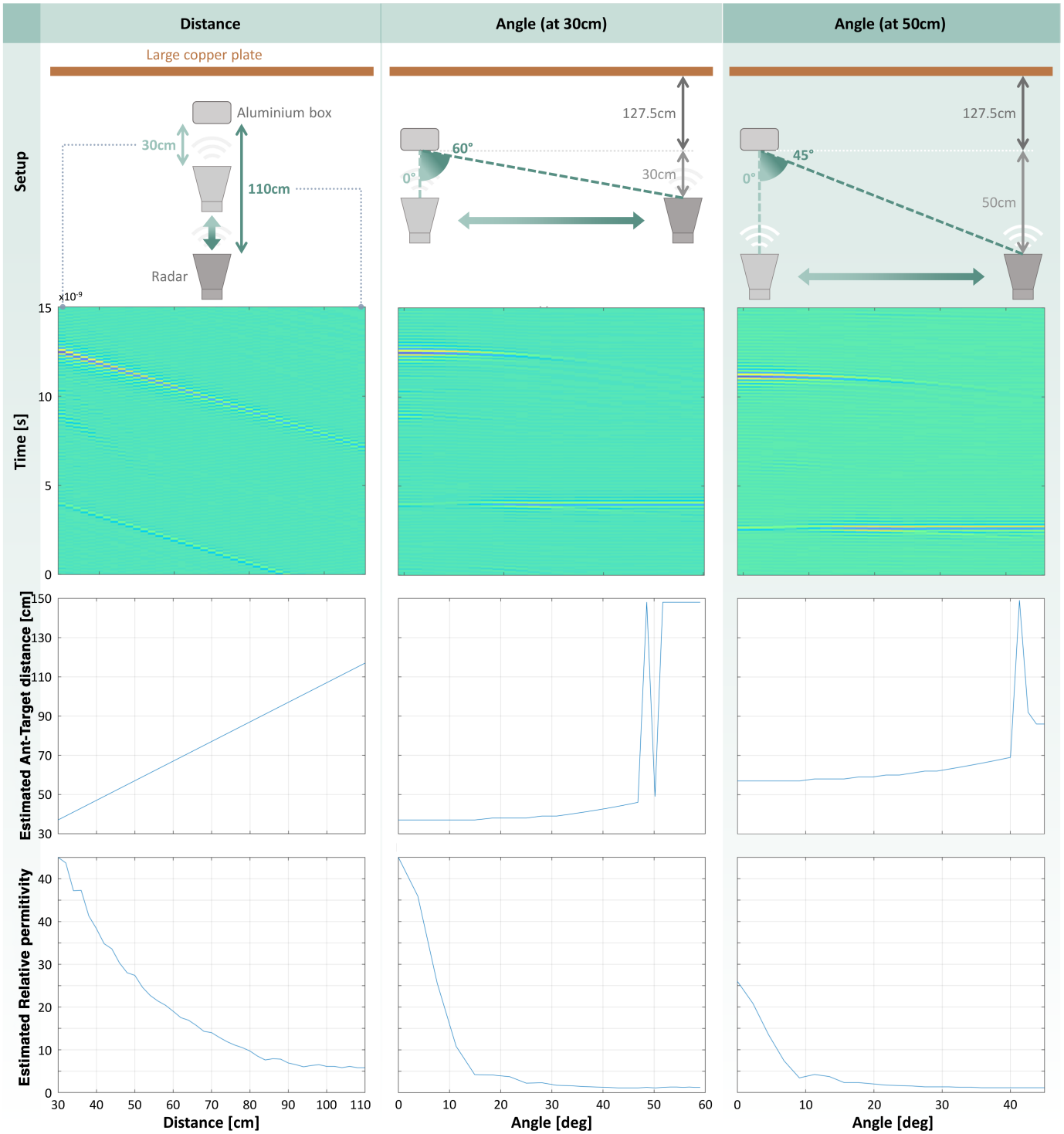


Fig. 6: Effect of radar-target angle and distance on the signal after radar source and antenna effects removal and inversion.

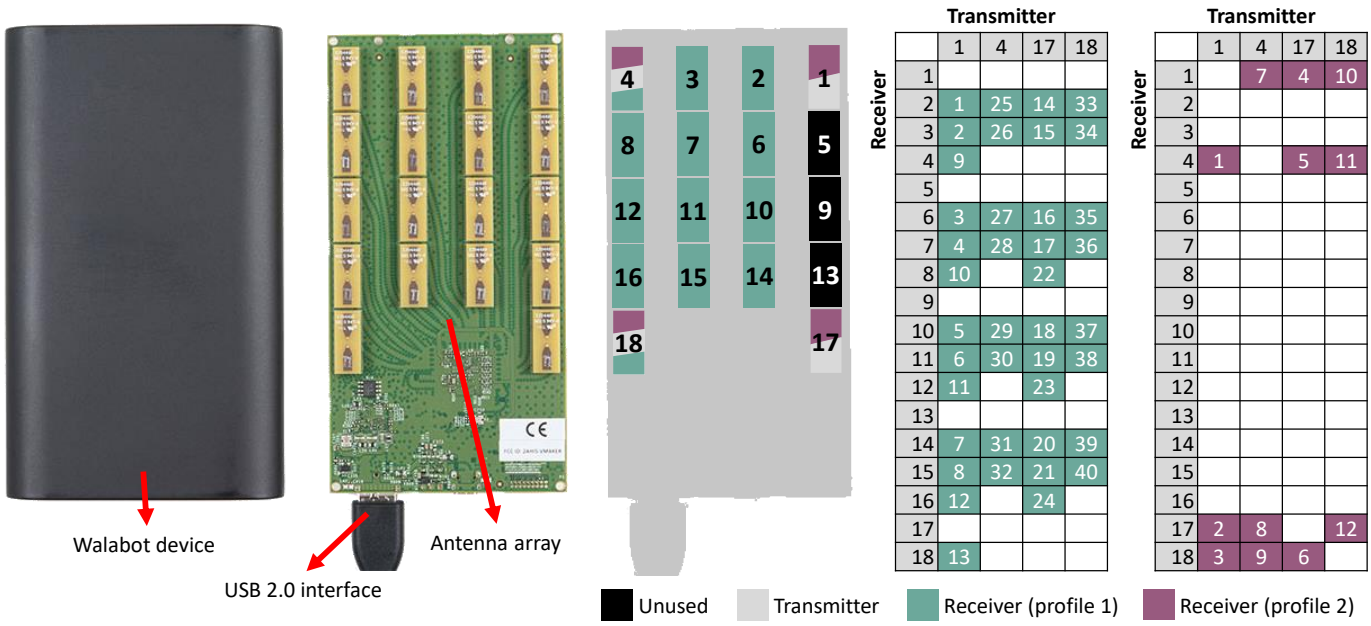


Fig. 7: The Walabot device: plastic enclosure and PCB, antennas IDs, and antenna pairs IDs (profiles 1 and 2).

and angular limitations analyzed in subsection IV-E. Because existing studies use different radar systems, gesture vocabularies, participant populations, and evaluation protocols, their reported classification accuracies are not directly comparable; evaluation on a common multi-radar dataset remains necessary for a controlled performance comparison. Finally, data from the inversion and filtering stages (Sections III-E and III-F) could facilitate breathing patterns recognition, by feeding the distance metric to K-Nearest Neighbor (KNN) algorithms, and enable robust tracking and recognition of hand trajectories with trilateration, by using the distances computed by the inversion stage from multiple spatially separated radar antennas to determine the 2D or 3D coordinates of the hand.

IV. Experiments

Two experiments were conducted to evaluate the efficacy of the radar data processing pipeline for hand gesture recognition. This section first describes the sensor and gesture dataset used in this testing and then details both experiments.

A. Walabot

The Walabot represents a commercially available, ultra-wideband (UWB) radar sensor employing frequency-modulated continuous-wave (FMCW) technology, and it incorporates an array of bowtie antennas. The advantage of using an antenna array, as opposed to a monostatic or bistatic radar configuration, is its capability to furnish not only distance information but also direction information. This work focuses on the higher end "Developer" version, which operates in the 6.3-8 GHz range and features 18 radar antennas, including four used as transmitters and 15 as receivers (Fig. 7). Its compact form factor (72×140 mm board) and USB connectivity make it suitable for use in (semi-)mobile and stationary contexts of use. The Walabot Developer SDK enables the retrieval of raw time-domain radar data for use in custom applications.

B. Custom radar with Horn antenna and VNA

In addition to the Walabot sensor, we used a custom stepped-frequency continuous-wave (SFCW) radar built around a lightweight vector network analyzer (R60, Copper Mountain Technologies, Indianapolis, IN, USA) connected to a broadband double-ridged horn antenna operating in a monostatic configuration (BBHA 9120 A, Schwarzbeck Mess-Elektronik, Schönau, Germany). The system operates over the 0.8–5 GHz frequency band and measures the complex reflection coefficient $S(\omega)$ over a dense set of frequency samples, which are directly used in the full-wave processing pipeline described in Section III. The antenna exhibits high directivity, with a -3 dB beamwidth of approximately 45° in the E-plane and 30° in the H-plane at 1 GHz, making the system well suited for controlled laboratory experiments and for applying the full-wave inversion model over a well-defined field of view [48]. The radar was mounted on a motorized three-axis positioning system, enabling precise control of its position and distance relative to the scene. This SFCW radar serves as a flexible, high-signal-to-noise reference system for the radar-system invariance and sensitivity analyses presented in Experiments 1 and 2.

C. Gesture Set

A dataset of 20 frequent gestures (Fig. 11 and 12) was designed based on the results of a systematic literature review [17] and gesture elicitation studies [1]: (a) open hand, (b) close hand, (c) open then close hand, (d) extend

TABLE I: Summary of the pipeline stages, with their name and description, the benefits and limitations of using the signal they produce, and potential external components that they could interface with. External components deemed most suited to a particular stage are highlighted in a **bolder font**.

ID	Name	Description	Benefits (+) and limitations (−)	External components
a	Raw data capture (+ FFT)	Retrieve raw radar signal in the time or frequency domain.	+ No processing required − Internal reflections − Transmissions between antennas − Antenna-target interactions − Reflections of the surrounding environment − High complexity of the signal	DNN, KNN, RDM, RAM, point cloud generation
b	Radar source & antenna effects removal	Apply the radar equation to the raw radar signal to remove internal reflections, direct transmissions between the antennas, and antenna-target interactions.	+ Normalization wrt. the radar system − Reflections of the surrounding environment − High complexity of the signal	DNN, KNN, **RDM**, **RAM**, point cloud generation
c	Background scene removal	Apply the superposition principle to the radar signal to remove the reflections of static objects in the surrounding environment of the radar.	+ Normalization wrt. the radar system + Normalization wrt. the surrounding environment − High complexity of the signal	DNN, KNN, RDM, RAM, **point cloud generation**
d	Time gating	Temporally truncate the radar signal to exclude reflections occurring outside the "useful" range of the radar.	+ Normalization wrt. the radar system + Normalization wrt. the surrounding environment − High complexity of the signal	**DNN**, KNN, RDM, RAM, point cloud generation
e	Full-wave inversion	Apply full-wave inversion to the radar signal to extract the hand-radar distance d and relative permittivity $\epsilon_{r,e}$ metrics.	+ Normalization wrt. the radar system + Normalization wrt. the surrounding environment + Lower complexity (only two metrics) − Loss of some information − Incorrect estimations of distance and relative permittivity	Trilateration, KNN
f	Filtering	Apply two filtering strategies to clean up the signal produced by the full-wave inversion stage.	+ Normalization wrt. the radar system + Normalization wrt. the surrounding environment + Lower complexity (only two metrics) − Loss of some information	**Trilateration**, **KNN**

one finger, (e) extend two fingers, (f) extend three fingers, (g) extend four fingers, (h) swipe right, (i) swipe left, (j) swipe up, (k) swipe down, (l) draw infinity, (m) draw circle, (n) draw Z, (o) push fist, (p) push palm, (q) hello, (r) knock twice, (s) barrier gesture, (t) touch nose. The gestures are classified according to Aigner et al.'s taxonomy [66] and our taxonomy, inspired by Vatavu's iFad gestures [67] (Fig. 8). Each gesture was performed ten times by six participants in a controlled environment, resulting in a total of 20 gestures × 6 participants × 10 repetitions = 1200 gesture recordings. All participants gave their informed consent for inclusion before participating in the study, and the protocol was approved by the Ethics Commission of the Institute for Research in the Psychological Sciences (Université catholique de Louvain).

The recording setup is shown in Fig. 9. The Walabot was placed vertically on a cardboard box, a material relatively transparent to RF waves [36]. The participants were sitting in front of the Walabot at a distance of approximately 70 cm, enough to ensure that the far-field hypothesis of the antenna model is valid at all times. The Walabot was configured to record data from 12 antenna pairs at around

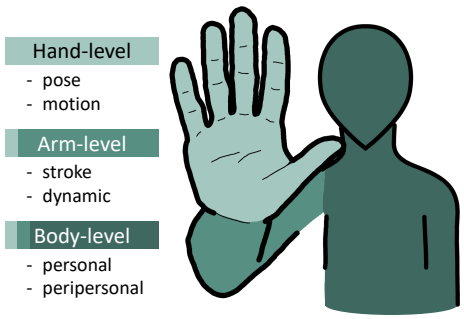


Fig. 8: Taxonomy inspired by Vatavu's iFAD gestures [67].

41 frames/sec and 1024 time samples/frame. A custom script was developed to streamline the recording process.

The recording procedure for one gesture is as follows (Fig. 10): show the gesture to the participant, then, for each template (1) the participant starts with their hands on their laps, (2) the recording is started, (3) the participant produces the gesture, (4) the participant places their hands back on their laps, (5) the recording is stopped and the data are saved to a text file. This sequence of actions is repeated for each repetition of each gesture. On average,



Fig. 9: Setup used for acquiring the hand gesture dataset: (1) acquisition station, (2) Walabot device, (3) participant producing a gesture.

a recording session (i.e., records 10 templates for each of the 20 gestures) takes around 30 minutes.

D. Experiment 1: Radar-system Invariance

This experiment evaluates the radar-system invariance property (P1) of our full-wave normalization. Two different radar systems observe exactly the same scene (air above a large copper plate), and we verify that, after removing radar-source and antenna effects using the far-field radar equation, both systems yield the same normalized Green's function $G_{xx}(\omega)$.

1) Experimental Setup and Data Acquisition: We consider the configuration shown in Fig. 5: a large copper plate, modeled as an infinite Perfect Electric Conductor (PEC), is placed in front of the radar in air. The medium therefore consists of a semi-infinite air layer over an infinite PEC, for which an analytical expression of the multi-layered Green's function $G_{xx}(\omega)$ is available in previous work [48], [69]. We used 2 ultra-wideband radar systems:

- A stepped-frequency continuous-wave (SFCW) radar implemented with a vector network analyzer (VNA) and a BBHA 9120 A horn antenna (as described in Section II), and
- A frequency-modulated continuous-wave (FMCW) Walabot radar (described in Section IV-A).

Both systems operate over the same frequency range used in the rest of this paper and are positioned in the far field at a distance of $d = 40$ cm from the copper plate, as illustrated in Fig. 4. For each radar, raw time-domain signals $s(t)$ are considered for this configuration. The top row of Fig. 4 shows that these raw signals have different amplitudes and waveforms due to the different radar electronics, modulation schemes, and antenna patterns, as well as relatively strong internal reflections and direct coupling.

Prior to this experiment, each radar is externally calibrated following the procedure of Section II-B, which is based on the far-field full-wave model of Lambot et al. [48] and its experimental validation for pulsed radars by Mahmoudzadeh and Lambot [69]. For each system we acquired measurements $S_k(\omega)$ at several antenna distances from a PEC, compute the corresponding Green's functions $G_{xx,k}(\omega)$ for the air–PEC configuration, and solve the linear system in Eqs. 6–9 to estimate the characteristic functions $R_i(\omega)$, $R_s(\omega)$ and $T(\omega)$ of the radar–antenna system. It is worth noting that the antenna radiation pattern is not explicitly accounted for in the far-field model and its calibration used in this study. While the point source–receiver formulation provides a highly accurate representation of the angular spectrum for infinite planar layered media, making the model well suited for such configurations, deviations may arise for finite, localized targets observed at nonzero incidence angles. In these cases, differences between the actual antenna radiation pattern and the idealized model may introduce angle-dependent bias, particularly at large offsets from the antenna boresight.

2) Data Analysis: For the Walabot (FMCW) radar, the raw time-domain signal $s(t)$ is first transformed into the frequency domain $b(\omega)$ using the Fourier transform in Eq. 3. For the SFCW radar, the VNA directly provides the complex ratio

$$S(\omega) = \frac{b(\omega)}{a(\omega)},$$

as in the far-field radar equation (Eq. 2). In both cases, the measured quantity $S(\omega)$ is related to the scene Green's function $G_{xx}(\omega)$ through the far-field radar equation (Eq. 2:

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

Using the calibrated characteristic functions $R_i(\omega)$, $R_s(\omega)$ and $T(\omega)$ for each radar, we analytically invert the radar equation with Eq. (5) to obtain the normalized Green's function:

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

This operation removes all radar-source and antenna effects, so that $G_{xx}(\omega)$ depends only on the air–PEC configuration and not on the specific radar hardware. For visualization and comparison, the normalized Green's function is then transformed back to the time domain using the inverse Fourier transform (Eq. 13) to obtain $g_{xx}(t)$:

$$g_{xx}(t) = \mathcal{F}^{-1}\{G_{xx}(\omega)\}.$$

3) Results: The bottom row of Fig. 4 presents the normalized time-domain signals $g_{xx}(t)$ obtained after applying the full-wave normalization to both radars. In contrast to the raw data, the normalized signals now have comparable amplitudes and nearly identical waveforms. The two curves almost overlap over the entire time window, indicating that both radar systems produce the same Green's function $G_{xx}(\omega)$ for the same scene.

This result experimentally confirms the radar-system invariance property (P1): once radar-source and antenna effects are removed using the calibrated full-wave model (Eqs. 2, 5–9, the remaining signal is a normalized Green's

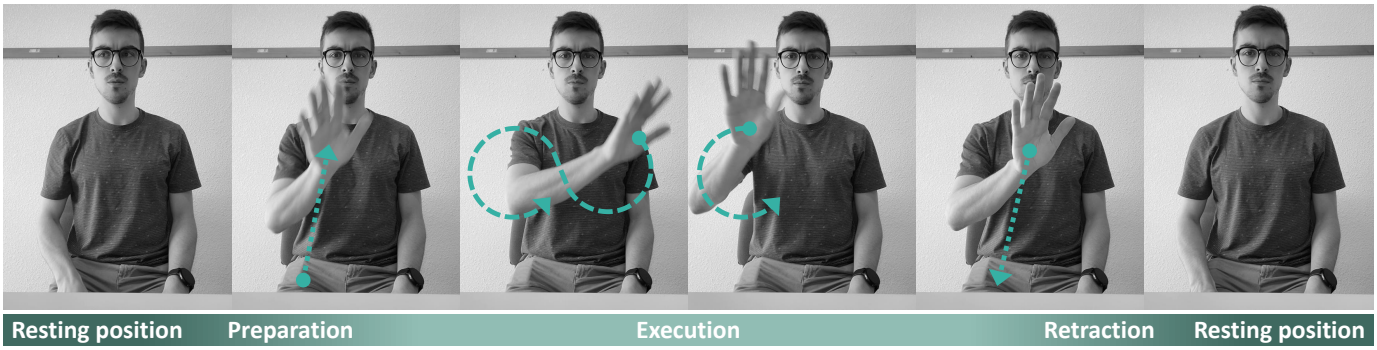


Fig. 10: The five phases of a gesture, based on [68].

function that depends only on the observed medium and not on the particular radar implementation.

In particular, previous work by Mahmoudzadeh et al. [69] already demonstrated the compatibility of the full-wave model between a frequency-domain SFCW GPR (800–4000 MHz) and a time-domain pulsed GPR with a 225–2500 MHz bandwidth, i.e., over only a *partially overlapping* frequency range. In contrast, our experiment compares an SFCW horn-antenna radar and an FMCW Walabot sensor that are both constrained to exactly the same 0.8–5 GHz operating band, and for which we directly normalize and overlay the measured signals using the same full-wave calibration. This provides a more stringent and practically relevant validation of radar-system invariance, reinforcing the hardware-agnostic nature of the proposed processing pipeline.

E. Experiment 2: Sensitivity of Full-Wave Inversion

The goal of this experiment is to **assess the sensitivity** of the full-wave inversion stage (Subsection III-E) to three key factors: (1) The presence of strong static reflectors in the background, (2) The distance between the target and the radar (30–110 cm), and (3) The angle between the target and the radar boresight (0–60° at 30 cm and 0–45° at 50 cm). Rather than using a human hand, we use a simple, highly reflective target to obtain controlled measurements and to better understand how distance and aspect angle affect the estimated distance and apparent permittivity after full-wave inversion (Fig. 5).

1) Experimental Setup and Data Acquisition: The SFCW radar was aimed at a hollow aluminium box (height 21.6 cm, width 10 cm, depth 8 cm) mounted on a PVC pole. A 3×3 m^2 copper plate was fixed to the wall 127.5 cm behind the box to act as a strong background reflector (Fig. 5). Two sets of measurements were performed:

- **Distance sweep:** the box was placed on the radar boresight at distances from 30 to 110 cm, in steps of 10 cm.
- **Angle sweep:** at two reference distances (30 and 50 cm), the box was rotated laterally so that the radar–target angle varied from 0° to 60° (at 30 cm) and from 0° to 45° (at 50 cm).

For each configuration, the raw frequency-domain signal $S(f)$ was acquired and converted into the time domain ($s(t)$), and the corresponding normalized Green's function was obtained using the processing pipeline (Fig. 3). Figs. 6 and 5 summarize the configurations and the resulting inversion outputs.

2) Data Analysis: The acquired radar data were processed using the full pipeline of Fig. III:

1) **Radar source and antenna effects removal** (Subsection III-B): the far-field radar equation (Eq. 2 and its inversion (Eq. 5) were used together with the previously calibrated radar-antenna characteristic functions $R_i(\omega)$, $R_s(\omega)$, and $T(\omega)$ to obtain the normalized Green's function $G_{xx}(\omega)$ of the scene (air + aluminium box + copper plate).
2) **Background scene removal** (Subsection III-C): reflections from the static background, including the large copper plate on the wall, were subtracted from the normalized signal using the superposition principle, leaving the contribution of the aluminium box.
3) **Time gating** (Subsection III-D): an inverse FFT (Eq. (11)) was applied to obtain the time-domain Green's function $g_{xx}(t)$, and a time window was selected so as to isolate the reflection from the box and discard later reflections from the plate and other distant objects (Fig. 3d,e).
4) **Full-wave inversion** (subsection III-E): for each configuration, the windowed time-domain signal was fed to the full-wave inversion module, which estimates the antenna–target distance d and the apparent relative permittivity $\varepsilon_{r,e}$ of an equivalent infinite half-space (Eq. 12. The inversion is implemented through a look-up table (LUT) over $(d, \varepsilon_{r,e})$, minimizing the objective function defined in Eq. 12.

3) Results and Interpretation: The effect of the background, distance, and angle on the inversion results is summarized in Fig. 6.

Effect of the background copper plate: Despite the presence of a large copper plate 127.5 cm behind the box, the inversion results are essentially unaffected by the background reflector: after background subtraction and time gating, the inversion effectively isolates the response of the aluminium box, such that the measurements can be interpreted as if only this target were present. This confirms that the time-gating stage successfully removes late reflections from strong static reflectors located behind

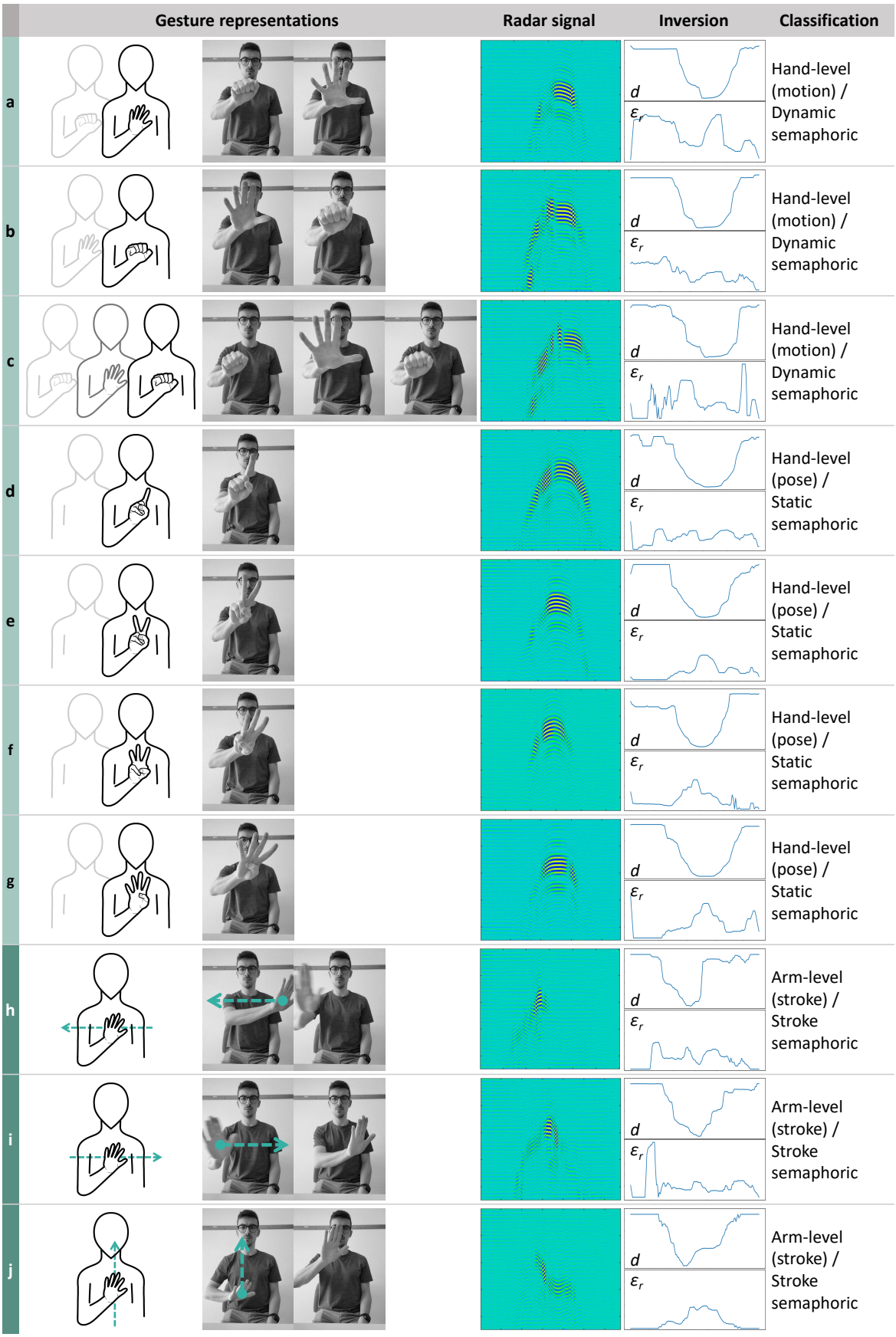


Fig. 11: The 20 gestures of our dataset and their classification according to our gesture taxonomy (Fig. 8) and the taxonomy of Aigner et al. [66].

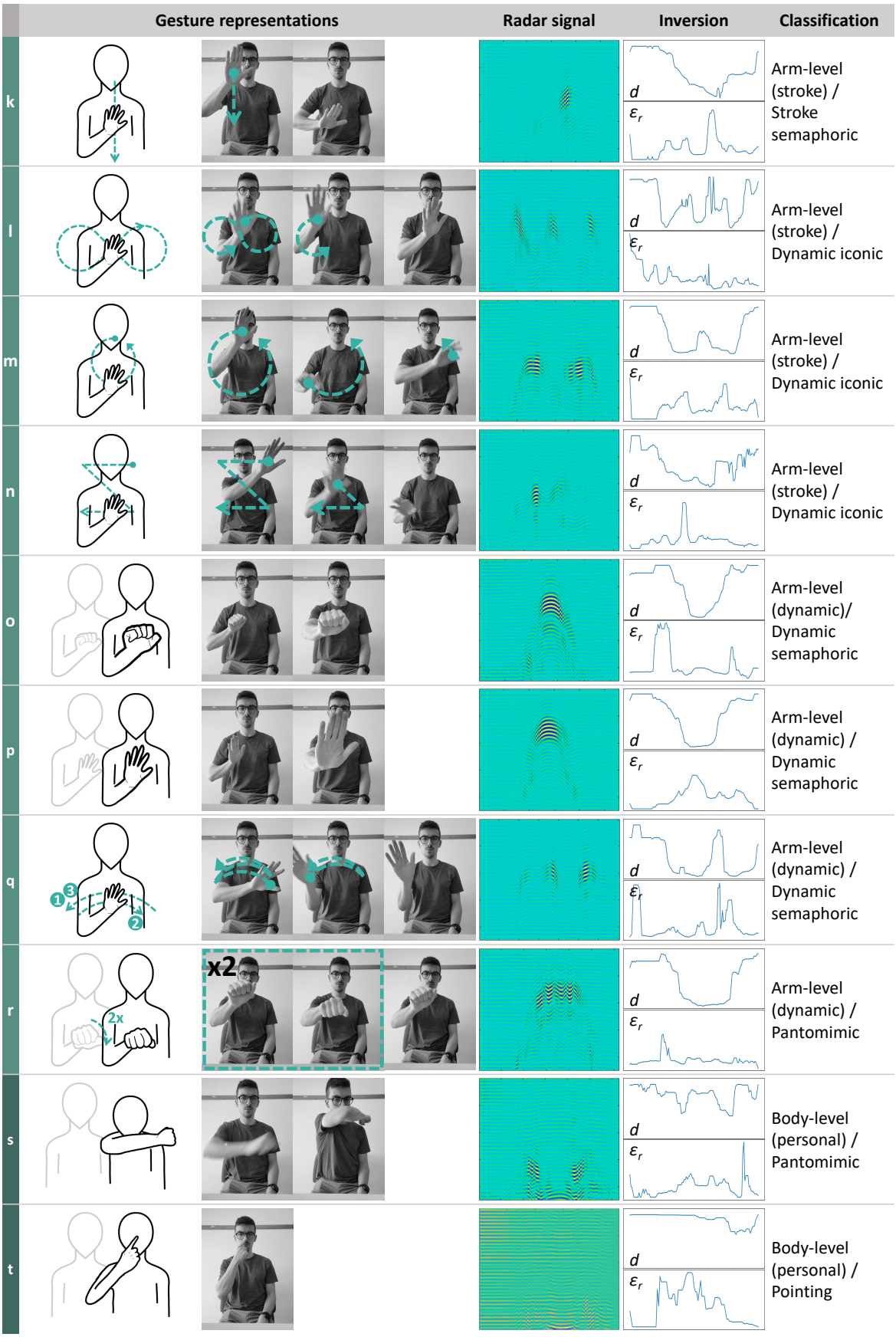


Fig. 12: The 20 gestures of our dataset and their classification according to our gesture taxonomy (Fig. 8) and the taxonomy of Aigner et al. [66] (cont.).

the user.

Effect of distance: As the radar–target distance increases from 30 to 110 cm, the estimated distance d closely follows the true distance, indicating that the inversion remains accurate over this range for a strong, well-aligned reflector. In practice, the maximum reliable distance will depend on the radar's dynamic range and the reflectivity of the target: systems with lower signal-to-noise ratio, or weaker targets, are expected to require a shorter maximum interaction distance.

Effect of angle: When the target is rotated away from the radar boresight, the behaviour of the inversion is consistent with the high directivity of the horn antenna. For angles up to roughly the -3 dB beamwidth, the estimated distance remains close to the true value. However, the inversion fails to provide accurate distance estimates beyond about 40° at 30 cm and 46° at 50 cm (Fig. 6), where the target lies outside the main lobe of the antenna. This illustrates the practical limitation imposed by the antenna radiation pattern on the usable field of view.

Effect on apparent permittivity: In contrast with the distance estimate, the apparent permittivity $\varepsilon_{r,e}$ varies strongly with both distance and angle. As the box is moved farther away or increasingly off-axis, the estimated permittivity decreases markedly. This behavior is consistent with the assumptions of the EM model: the inversion assumes an infinite half-space, whereas the box—and, in our application, the hand—is a finite, relatively small object. At larger distances and angles, a smaller fraction of the emitted energy interacts with the target (due to 3D spherical divergence and the antenna radiation pattern), leading to weaker reflections than predicted by the infinite-medium model and thus to a lower apparent permittivity. The sensitivity of $\varepsilon_{r,e}$ to position and orientation suggests that, in dynamic scenarios where users move around the radar, distance is a more reliable feature than permittivity, while the latter may primarily encode aspects of hand pose or target orientation.

Overall, this experiment shows that the full-wave inversion is robust to strong static reflectors in the background and can accurately estimate distance over a wide range of configurations, but that the apparent permittivity is highly dependent on distance and angle. These observations motivate future extensions of the forward model to better account for finite-size targets and the antenna radiation pattern, with the goal of making $\varepsilon_{r,e}$ less dependent on the hand–radar geometry.

F. Experiment 3: All Gestures

This first experiment evaluates the efficacy of two stages of our pipeline by training Jackknife [70], a 1-NN gesture recognizer, on the twenty classes of the dataset to determine whether their output data contain enough information to accurately classify a large and diverse set of gestures. Jackknife [**?**] was set up with its default settings for the inner product distance measure. This gesture recognizer was selected because it can accommodate different data types, works with as few as one training template, and is fast to train and execute. We chose to work with an existing recognizer, as implementing the best-performing algorithm for radar gesture recognition was outside of the scope of this paper. However, other algorithms (e.g.„ deep-learning-based) and signal processing techniques could be used in conjunction with our pipeline to achieve better performance (Section III-G).

1) Procedure: A leave-one-out cross-validation (LOOCV) is conducted on data from the time gating (Fig. 3d) and filtering (Fig. 3f) stages of the pipeline in three different scenarios: (1) user-dependent, i.e., for each gesture sample in the dataset, the sample is selected for testing and all other samples performed by the same user are used for training, (2) user-independent, i.e., for each gesture sample in the dataset, the sample is selected for testing and all other samples performed by the other users are used for training, and (3) mixed, i.e., for each gesture sample in the dataset, the sample is selected for testing and all other samples are used for training. The execution time and recognition rate are measured for each trial, and all trials are run on a Windows 10 laptop with 32GB of RAM and an Intel i7-10875H CPU. Taranta et al.'s Jackknife [70] with inner product distance measure is used for gesture recognition and trained with $N{=}\{x \in \mathbb{N} \mid 4 \leq x \leq 40\}$ sampling points and antenna pairs $AP = \{(1,2,3,6,8,9),(4,5,7,10,11,12),(1,2,3,4,5,6,7,8,9,10,11,12)\}$. The total number of training templates per trial is 199 (9 to 10 per gesture class) for user-dependent, 1000 (50 per gesture class) for user-independent, and 1199 (59 to 60 per gesture class) for mixed.

2) Results and discussion: The system reaches 90.58% accuracy over the whole dataset in the user-dependent scenario using data from the time gating stage, with an average execution time of 34 ms ($N{=}38$, $AP{=}(4,5,7,10,11,12)$). We observe more confusion between gesture classes d, e, f, g, and p, most of which fall under 90% accuracy. These gestures feature small changes in hand pose (and, thus, slight variations in the amplitude of the reflected signal) and similar trajectories, making them tricky to differentiate. Some confusion is also observed between the gestures j and k (swipe up and down), whose opposite trajectories result in very similar radar signatures. Conversely, seven gesture classes reach over 95% accuracy, including one hand-level gesture (j), five arm-level gestures (h, i, l, m, and n) and one body-level gesture (t). The accuracy drops to 45.17% with data from the filtering stage ($N{=}26$, $AP{=}(1,2,3,4,5,6,7,8,9,10,11,12)$), with a much shorter execution time at 2 ms. Only seven gesture classes are recognized with more than 50% accuracy, namely b, k, l, m, n, p, and r. All but two of these classes (b and r) feature different motions in space with similar hand poses. The distance measurement is thus crucial for distinguishing between these gesture classes, while the relative permittivity is less relevant. We also observe some confusion between gesture classes d, e, f, g, o, and p, which feature different hand poses but similar trajectories. In

particular, gestures d and o are recognized with less than 30% accuracy. The relative permittivity is more relevant than the distance measure to identify hand poses. As such, the permittivity measures may lack the required precision to accurately distinguish between these slight variations in hand poses, thus explaining the lower accuracy on these gesture classes. In this context, it is noteworthy to mention that the obtained permittivity values exhibit a high sensitivity to minor shifts in hand orientation, rendering them susceptible to recognition inaccuracies. Fig. 13 shows the confusion matrices corresponding to these two configurations.

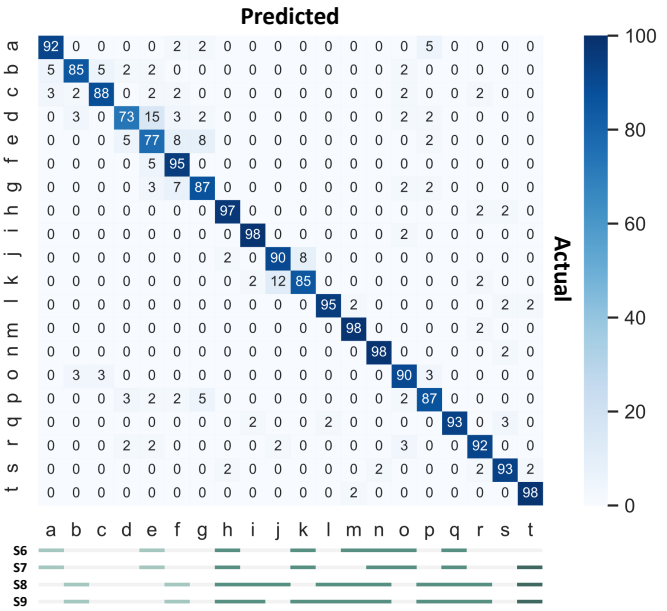
Accuracy decreases sharply in the user-independent scenario, dropping to 24.58% using data from the time gating stage (N=38, AP=(1, 2, 3, 6, 8, 9)). Looking at the performance for each gesture class shows similar patterns to the user-dependent scenario, albeit with much lower accuracy. There is a lot of confusion between gesture classes a, b, c, d, e, f, g, and p. Only five gesture classes reach more than 30% accuracy, namely k, m, p, s, and t. In particular, class k, which has a very different radar signature than most other gestures, is accurately recognized 77% of the time. Accuracy is only 15.75% with data from the filtering stage (N=29, AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)). Patterns are very hard to distinguish in this scenario because of the very low accuracy. Only two gesture classes reach 30% accuracy, namely, classes m and n. Some form of normalization of the signal with respect to the users could be applied to improve the performance of the system.

In the mixed scenario, accuracy is only slightly lower than in the user-dependent scenario, reaching 90.25% with data from the time gating stage (N=39, AP=(1, 2, 3, 6, 8, 9)), and 40.17% with data from the filtering stage (N=25, AP=(1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12)). The execution time is longer, reaching 203 ms and 12 ms with time-domain and inversion data, respectively. Confusion matrices in this scenario are similar to the user-dependent scenario, with similar observations.

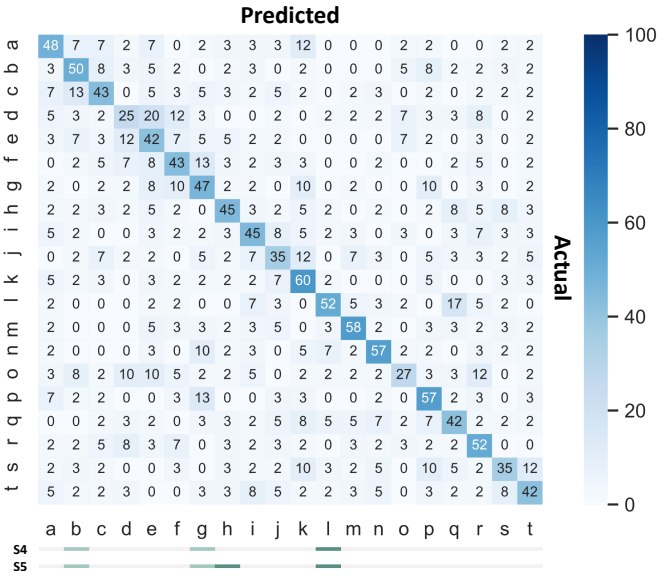
Overall, the best performance is achieved in user-dependent and mixed scenarios, using data from the time gating stage of our processing strategy. For real-time execution in a mixed scenario, the number of training samples should be carefully selected to maximize accuracy while minimizing the response time of the system, which depends mainly on the execution time of the recognizer. The good performance of the system in the mixed scenario shows that it can accurately recognize gestures from one user, even when the training set contains samples from other users. This property makes the combination of our pipeline and the Jackknife recognizer ideal for building gestural interfaces that are trained and used by a small set of users, such as a gesture-based TV remote. In addition, the use of a template matching recognizer enables quick and easy customization of the gesture set.

G. Subsets of Gestures

This experiment aims to identify well-differentiated subsets of our gesture dataset that can be recognized accurately in a user-dependent scenario.



(a) Time gating stage.



(b) Filtering stage.

Fig. 13: Confusion matrices of the best performing configurations in the user-dependent scenario on the complete dataset.

1) Procedure: The procedure is similar to the first experiment (section IV-F1). For each subset, the Jackknife recognizer is trained with data from the same sets of antenna pairs, numbers of sampling points, and stages of the pipeline, in a user-dependent scenario. Three gesture subsets were created based on our taxonomy: hand-level, arm-level, and body-level gestures. In addition, subsets of highly differentiable gesture classes were created as follows: (1) identify all subsets with zero (resp. one) confusion between pairs of gesture classes in the results from the previous user-dependent testing, (2) order them by number of gesture classes and total number of confusions, and (3) select the largest subset(s) with zero (resp. one)

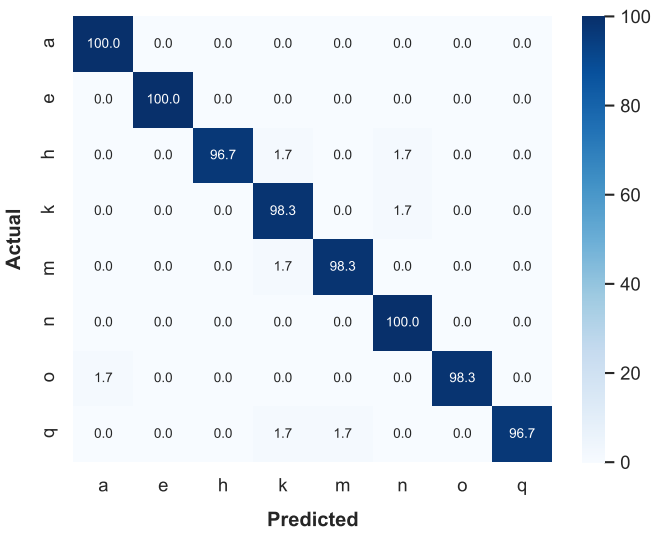


Fig. 14: Confusion matrix of the best performing configuration on gesture subset S6.

Id	Description	Gestures
S1	Hand-level	a b c d e f g
S2	Arm-level	h i j k l m n o p q r
S3	Body-level	s t
S4	Filtering, C=0	b g l
S5	Filtering, C=1	b g h l
S6	Time gating, C=0 (a)	a e h k m n o q
S7	Time gating, C=0 (b)	a e h k n o q t
S8	Time gating, C=1 (a)	b f h i j l m n p q r t
S9	Time gating, C=1 (b)	b f h i k l m n p q r t

TABLE II: The nine selected subsets of gestures. C represents the maximum number of confusion between pairs of gestures. Gestures are colored based on the taxonomy in Fig. 8.

confusion between pairs of gesture classes. We thus have both larger subsets with low confusion between gesture classes and smaller subsets with little to no confusion. The selected subsets are summarized in Table II and highlighted in Fig. 13. The time-gating subsets (S6 to S9) consist mainly of arm-level gesture classes, with only two hand-level gesture classes per subset. This indicates that the larger motion of arm-level gestures was easier to differentiate than the smaller-scale hand-level gestures. The filtering subsets (S4 and S5) contain fewer gesture classes than the time-gating subsets, which reflects the high confusion between gesture classes when using data from the filtering stage.

2) Results and discussion: Despite containing fewer gestures, the classification accuracy on subset S1 (hand-level) is lower than on the full set of gestures, at 87.86% (N=35, AP=$(1,2,3,4,5,6,7,8,9,10,11,12)$). Conversely, the accuracy on the larger S2 subset (arm-level) is higher, at 96.21% (N=37, AP=$(1,2,3,6,8,9)$). This indicates that arm-level gestures may be easier to differentiate with our technique than smaller-scale gestures. Accuracy on body-level gestures (S3) is even higher at 99.17% (N=20, AP=$(4,5,7,10,11,12)$), but the small size of this subset leaves less room for confusion between gestures. As expected, classification accuracy is lower on data from the filtering stage, with the accuracy for hand-, arm-, and body-level gestures reaching 51.90% (N=26, AP=$(1,2,3,4,5,6,7,8,9,10,11,12)$), 55.45% ($N$=34, AP=$(1,2,3,4,5,6,7,8,9,10,11,12)$), and 80.00% ($N$=28, AP=$(4,5,7,10,11,12)$), respectively. A similar pattern emerges, however, with arm-level gestures being classified more accurately than hand-level ones despite the greater size of the subset.

The accuracy on subsets S4 and S5 with data from the filtering stage is relatively low considering the size of the subsets: 86.11% (N=22, AP=$(1,2,3,4,5,6,7,8,9,10,11,12)$) and 77.50% ($N$=23, AP=$(1,2,3,4,5,6,7,8,9,10,11,12)$) for three and four gesture classes, respectively. Conversely, Jackknife managed a high accuracy of 98.54% on subsets S6 (N=38, AP=$(4,5,7,10,11,12)$) and S7 (N=38, AP=$(4,5,7,10,11,12)$) using data from the time gating stage. The accuracy was still high at 97.78% and 97.50% on the larger S8 (N=38, AP=$(4,5,7,10,11,12)$) and S9 (N=37, AP=$(4,5,7,10,11,12)$) subsets, respectively. Other subsets could perform even better but testing them all would be unrealistic due to the large number of subsets.

Experiment 2 (Section IV-E) highlights that, even under controlled laboratory conditions, the outcome of the full-wave inversion is constrained by the signal-to-noise ratio (SNR) and by the presence of additional moving reflectors in the scene. In real-world deployments, low SNR or motion of objects behind or in front of the user can yield erroneous estimates of both distance and apparent permittivity.

To mitigate these effects, we apply a lightweight filtering stage after inversion (Section III-F). First, when the estimated apparent permittivity is close to that of air, we interpret this as an absence of a detectable hand in front of the radar and replace both the distance and permittivity values with default constants (d=75 cm, $\varepsilon_{r,e}$=1.0). This prevents spurious measurements from being propagated to downstream components. Second, we apply a moving median filter to the distance and permittivity time series. The median filter efficiently suppresses large, abrupt variations that are not physically plausible, while preserving gradual changes due to genuine hand motion.

This design deliberately trades some temporal resolution for robustness. More capable radar hardware—for instance with a larger dynamic range, higher SNR, or a wider field of view—would reduce the need for aggressive filtering, as hand reflections would be more clearly separated from noise and from other moving scatterers even when the hand is slightly off-axis. Importantly, however, the filtering stage operates on the two-dimensional inversion output $(d, \varepsilon_{r,e})$ only; its computational complexity per frame is therefore constant, $\mathcal{O}(1)$, irrespective of the original size of the radar signal. This makes it suitable for real-time gesture recognition on resource-constrained devices.

V. Conclusion

This paper presented and experimentally assessed a six-stage, physics-informed radar signal-processing pipeline for gesture recognition. The pipeline combines radar–antenna calibration, background-scene subtraction, time gating, full-wave inversion, and filtering to produce either normalized radar signals or low-dimensional physical descriptors. The evaluation comprised three experiments. First, the radar-system-invariance experiment showed that two distinct radar systems observing the same controlled scene produced closely aligned normalized Green's functions after removal of radar-source and antenna effects. Second, the sensitivity experiment characterized the influence of target distance, target angle, and strong static background reflectors on the full-wave inversion. Third, the gesture-recognition experiment evaluated a dataset of 20 gestures performed by six participants in user-dependent, user-independent, and mixed-user conditions.

Using the Jackknife gesture recognizer, the proposed processing achieved an accuracy of 90.58% on the full set of 20 gestures, 97.78% on a subset of 12 gestures, and 98.54% on a subset of eight gestures using signals from the time-gating stage in the user-dependent condition. Performance decreased in the user-independent condition, indicating that interpersonal variability remains an important challenge, whereas the mixed-user condition showed that a recognizer trained with examples from multiple intended users can provide a practical compromise. The inversion results further showed that the apparent-permittivity estimate is sensitive to target distance and off-axis angle because the current model approximates the finite hand as an infinite planar medium. Together, the experiments support the feasibility of the proposed normalization strategy while identifying the conditions under which its physical assumptions limit performance. Relevant resources, including the gesture set, dataset [71], and detailed experimental results, are available through the project website at https://sites.uclouvain.be/ingenious/2023/08/04/electromagnetic-model-and-inversion/.

Beyond laboratory gesture classification, the proposed processing architecture suggests several application domains in which contactless, camera-free interaction and stability across sensor deployments are advantageous. Because radar does not capture optical images, it can reduce the exposure of visually identifiable information relative to camera-based interfaces while continuing to operate under poor illumination. Depending on the operating frequency and on the electromagnetic properties, thickness, and geometry of an intervening layer, radar may also sense gestures through nonconductive dielectric covers, enabling sealed, concealed, or easy-to-clean interfaces [24], [25], [36]. In clinical environments, contactless gesture control could support the manipulation of medical images or other digital systems without requiring medical personnel to touch shared input devices, thereby helping to preserve sterile workflows [72], [73]. Related interaction constraints arise in isolated, confined, and extreme environments, including space and hypogravity contexts, where confined workspaces, protective equipment, and changes in body orientation and motor control can affect interaction [74], [75]. In such settings, the separation of one-time radar calibration from classifier training and the removal of stationary-background reflections could reduce the need for radar-specific data recollection after sensor replacement or deployment in a new static scene. For sensing through a fixed cover integrated with the sensor, the cover may be included consistently in the radar–antenna calibration; for a separate intervening material, its dielectric properties and geometry should instead be represented as layers in the electromagnetic forward model. These application scenarios remain prospective and require validation in their intended operational environments.

The findings of this paper suggest three main avenues for future exploration. First, there is potential in investigating how various external components could benefit from integration with our signal processing pipeline. For instance, training DNNs using data from the time gating stage (Section III-D) could significantly enhance their adaptability across diverse environments with less training data and little to no retraining required. In addition, the normalization performed by the second stage of our pipeline (Section III-B) could facilitate the accurate extraction of information such as point clouds or RDMs from radar signals. Finally, leveraging data from the full-wave inversion and filtering stages (Sections III-E and III-F) could help accurately retrieve the 3D position of the hand in space through trilateration. This information could serve to display a pointer on the screen and improve the recognition of hand trajectories. Second, focusing on improving individual stages, particularly the full-wave inversion stage (Sections III-E), could result in better performance in many scenarios. Potential improvements include accounting for the finite size of the hand to provide relative permittivity estimates independent of hand-radar distance, and normalizing this relative permittivity metric with respect to users' hand size. Third, exploring alternative radar systems and gesture recognition techniques could help further validate our approach. For example, we could employ data fusion techniques to combine signals from multiple radar antennas and investigate radar systems operating at different frequency ranges, with a better signal-to-noise ratio, or featuring more antennas.

Acknowledgments

The authors thank study participants for taking part in the collection of the gesture set employed in this paper. Arthur Sluÿters was funded by the "Fonds de la Recherche Scientifique - FNRS" under Grants n°40001931 and n°40011629. We also acknowledge the support of the COSMIC (COntinuouS MICrogesture acquisition and sensing for natural interaction) project funded by the "Fonds de la Recherche Scientifique - FNRS" under Grant n°40035986.

References

[1] N. Magrofuoco, P. Roselli, and J. Vanderdonckt, "Two-dimensional stroke gesture recognition: A survey," *ACM Comput. Surv.*, vol. 54, no. 7, Jul. 2021. [Online]. Available: https://doi.org/10.1145/3465400

[2] N. T. Attygalle, J. Vanderdonckt, and V. Parthiban, "Turning gesture sequences into gesture composition by macro-instructions: Definition, framework, and application in smart home," *Proc. ACM Hum.-Comput. Interact.*, vol. 10, no. 4, Jun. 2026, available: https://doi.org/10.1145/3816768.

[3] B.-F. Gheran, J. Vanderdonckt, and R.-D. Vatavu, "Gestures for smart rings: Empirical results, insights, and design implications," in *Proceedings of the 2018 Designing Interactive Systems Conference*, ser. DIS '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 623–635. [Online]. Available: https://doi.org/10.1145/3196709.3196741

[4] C. Ardito, P. Buono, M. F. Costabile, and G. Desolda, "Interaction with large displays: A survey," *ACM Comput. Surv.*, vol. 47, no. 3, Feb. 2015. [Online]. Available: https://doi.org/10.1145/2682623

[5] E. Velloso, D. Schmidt, J. Alexander, H. Gellersen, and A. Bulling, "The feet in human–computer interaction: A survey of foot-based interaction," *ACM Comput. Surv.*, vol. 48, no. 2, Sep. 2015. [Online]. Available: https://doi.org/10.1145/2816455

[6] R.-D. Vatavu, "Beyond features for recognition: Human-readable measures to understand users' whole-body gesture performance," *International Journal of Human–Computer Interaction*, vol. 33, no. 9, pp. 713–730, 2017. [Online]. Available: https://doi.org/10.1080/10447318.2017.1278897

[7] 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," *International Journal of Human–Computer Interaction*, vol. 39, no. 12, pp. 2492–2523, 2023. [Online]. Available: https://doi.org/10.1080/10447318.2022.2078464

[8] K. Sun, T. Zhao, W. Wang, and L. Xie, "Vskin: Sensing touch gestures on surfaces of mobile devices using acoustic signals," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 591–605. [Online]. Available: https://doi.org/10.1145/3241539.3241568

[9] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: A survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1–54, Jan. 2015, available: https://doi.org/10.1007/s10462-012-9356-9.

[10] N. H. Dardas and N. D. Georganas, "Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 11, pp. 3592–3607, 2011, available: https://doi.org/10.1109/TIM.2011.2161140.

[11] H. Cheng, L. Yang, and Z. Liu, "Survey on 3d hand gesture recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 9, pp. 1659–1673, 2016, available: https://doi.org/10.1109/TCSVT.2015.2469551.

[12] 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 *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3411764.3445367

[13] S. Chamorro, J. Collier, and F. Grondin, "Neural network based lidar gesture recognition for realtime robot teleoperation," in *2021 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2021, pp. 98–103, available: https://doi.org/10.1109/SSRR53300.2021.9597855.

[14] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking*, ser. MobiCom '13. New York, NY, USA: Association for Computing Machinery, 2013, p. 27–38. [Online]. Available: https://doi.org/10.1145/2500423.2500436

[15] Z. Yang, Z. Zhen, Z. Li, X. Liu, B. Yuan, and Y. Zhang, "Rf-cgr: Enable chinese character gesture recognition with rfid," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–16, 2023, available: https://doi.org/10.1109/TIM.2023.3322505.

[16] 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. 176 982–176 997, 2019, available: https://doi.org/10.1109/ACCESS.2019.2951349.

[17] 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," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 6, pp. 7967–7981, Jun. 2023, available: https://doi.org/10.1007/s12652-023-04606-9.

[18] C. Gu, C. Li, J. Lin, J. Long, J. Huangfu, and L. Ran, "Instrument-based noncontact doppler radar vital sign detection system using heterodyne digital quadrature demodulation architecture," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 6, pp. 1580–1588, 2010, available: https://doi.org/10.1109/TIM.2009.2028208.

[19] B. Li, W. Li, Y. He, W. Zhang, and H. Fu, "Radarnet: Noncontact ecg signal measurement based on fmcw radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–9, 2024, available: https://doi.org/10.1109/TIM.2024.3476545.

[20] W. Xia, Y. Li, and S. Dong, "Radar-based high-accuracy cardiac activity sensing," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021, available: https://doi.org/10.1109/TIM.2021.3050827.

[21] G. Li, Y. Ge, Y. Wang, Q. Chen, and G. Wang, "Detection of human breathing in non-line-of-sight region by using mmwave fmcw radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022, available: https://doi.org/10.1109/TIM.2022.3208266.

[22] 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, Jul. 2016. [Online]. Available: https://doi.org/10.1145/2897824.2925953

[23] X. Wang, H. Liu, X. Wang, V. C. Chen, M. G. Amin, and K. Cai, "Sensing, tracking, and recognition of macro–micro hand gestures using interferometric mimo radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024, available: https://doi.org/10.1109/TIM.2024.3412196.

[24] L. A. Leiva, M. Kljun, C. Sandor, and K. Copic Pucihar, "The wearable radar: Sensing gestures through fabrics," in *22nd International Conference on Human-Computer Interaction with Mobile Devices and Services*, ser. MobileHCI '20. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: https://doi.org/10.1145/3406324.3410720

[25] 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. 5771, 2021, available: https://doi.org/10.3390/s21175771.

[26] N. T. Attygalle, M. Kljun, and K. Čopič Pucihar, "Gesture recognition on deformable objects using millimeter-wave radar," in *Companion Proceedings of the 17th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, ser. EICS '25 Companion. New York, NY, USA: Association for Computing Machinery, 2025, pp. 50–58, available: https://doi.org/10.1145/3731406.3734979.

[27] F. Adib, C.-Y. Hsu, H. Mao, D. Katabi, and F. Durand, "Capturing the human figure through a wall," *ACM Trans. Graph.*, vol. 34, no. 6, Nov. 2015. [Online]. Available: https://doi.org/10.1145/2816795.2818072

[28] Y. Yao, J. Chen, S. Guo, X. Hao, N. Li, G. Cui, and L. Kong, "Target localization and wall parameters estimation via distributed through-wall imaging radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–11, 2025, available: https://doi.org/10.1109/TIM.2024.3502875.

[29] J. Cao, Y. Ding, Y. Peng, Y. Chen, and F. Ouyang, "A machine learning-based algorithm for through-wall target tracking by doppler twr," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–9, 2024, available: https://doi.org/10.1109/TIM.2024.3369133.

[30] C. Wang, D. Zhu, L. Sun, C. Han, and J. Guo, "Real-time through-wall multiperson 3-d pose estimation based on mimo radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–11, 2024, available: https://doi.org/10.1109/TIM.2024.3373058.

[31] A. Pandharipande, C.-H. Cheng, J. Dauwels, S. Z. Gurbuz, J. Ibanez-Guzman, G. Li, A. Piazzoni, P. Wang, and A. Santra, "Sensing and machine learning for automotive perception: A review," *IEEE Sensors Journal*, vol. 23, no. 11, pp. 11 097–11 115, 2023, available: https://doi.org/10.1109/JSEN.2023.3262134.

[32] 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, Aug. 2019, available: https://doi.org/10.1145/3241383. [Online]. Available: https://doi.org/10.1145/3241383

[33] S. Skaria, A. Al-Hourani, M. Lech, and R. J. Evans, "Hand-gesture recognition using two-antenna doppler radar with deep convolutional neural networks," *IEEE Sensors Journal*, vol. 19, no. 8, pp. 3041–3048, 2019, available: https://doi.org/10.1109/JSEN.2019.2892073.

[34] 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 Journal*, vol. 20, no. 18, pp. 10 706–10 716, 2020, available: https://doi.org/10.1109/JSEN.2020.2994292.

[35] J. Liu, Y. Liu, Y. Wang, Y. Chen, T. Zhou, and Y. Huang, "Deep learning approach for gesture recognition on millimeter- wave radar," in *2022 International Conference on Microwave and Millimeter Wave Technology (ICMMT)*, 2022, pp. 1–3, available: https://doi.org/10.1109/ICMMT55580.2022.10023429.

[36] 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, no. EICS, Jun. 2022, available: https://doi.org/10.1145/3532212.

[37] K. Čopič Pucihar, D. Salami, and N. T. Attygalle, "Current state and future research directions of radar-based human-computer interaction," in *Radar-Based Human-Computer Interaction*. Cham: Springer Nature Switzerland, 2026, pp. 235–267, available: https://doi.org/10.1007/978-3-032-04327-6_11.

[38] K. Čopič Pucihar, M. Kljun, and N. T. Attygalle, "Radar-based gesture recognition on deformable objects," in *Radar-Based Human-Computer Interaction*. Cham: Springer Nature Switzerland, 2026, pp. 107–122, available: https://doi.org/10.1007/978-3-032-04327-6_6.

[39] N. T. Attygalle, M. Kljun, U. Vuletić, and K. Čopič Pucihar, "Comparative testing of radar signal representations when sensing through materials," in *Radar-Based Human-Computer Interaction*. Cham: Springer Nature Switzerland, 2026, pp. 47–61, available: https://doi.org/10.1007/978-3-032-04327-6_3.

[40] B. Li, J. Yang, Y. Yang, C. Li, and Y. Zhang, "Sign language/gesture recognition based on cumulative distribution density features using uwb radar," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021, available: https://doi.org/10.1109/TIM.2021.3092072.

[41] Y. Wang, Y. Shu, X. Jia, M. Zhou, L. Xie, and L. Guo, "Multifeature fusion-based hand gesture sensing and recognition system," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022, available: https://doi.org/10.1109/LGRS.2021.3086136.

[42] Z. Zhang, Z. Tian, Y. Zhang, M. Zhou, and B. Wang, "u-deephand: Fmcw radar-based unsupervised hand gesture feature learning using deep convolutional auto-encoder network," *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6811–6821, 2019, available: https://doi.org/10.1109/JSEN.2019.2910810.

[43] P. Goswami, S. Rao, S. Bharadwaj, and A. Nguyen, "Real-time multi-gesture recognition using 77 ghz fmcw mimo single chip radar," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, 2019, pp. 1–4, available: https://doi.org/10.1109/ICCE.2019.8662006.

[44] 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. Mob. Wearable Ubiquitous Technol.*, vol. 5, no. 1, Mar. 2021. [Online]. Available: https://doi.org/10.1145/3448110

[45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015. [Online]. Available: https://arxiv.org/abs/1512.03385

[46] Z. Wang, F. Liu, X. Li, M. Ma, X. Feng, and Y. Guo, "A survey of hand gesture recognition based on fmcw radar," in *Proceedings of the 8th International Conference on Communication and Information Processing*, ser. ICCIP '22. New York, NY, USA: Association for Computing Machinery, 2023, p. 73–79. [Online]. Available: https://doi.org/10.1145/3571662.3571674

[47] H. Liu and Z. Liu, "A multimodal dynamic hand gesture recognition based on radar–vision fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–15, 2023, available: https://doi.org/10.1109/TIM.2023.3253906.

[48] S. Lambot, E. Slob, I. van den Bosch, B. Stockbroeckx, and M. Vanclooster, "Modeling of ground-penetrating radar for accurate characterization of subsurface electric properties," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2555–2568, 2004, available: https://doi.org/10.1109/TGRS.2004.834800.

[49] 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, Sep. 2023. [Online]. Available: https://doi.org/10.1145/3589645

[50] "Waves and Fields in Inhomogenous Media | IEEE eBooks | IEEE Xplore," https://ieeexplore.ieee.org/book/5270998.

[51] K. Michalski and D. Zheng, "Electromagnetic scattering and radiation by surfaces of arbitrary shape in layered media. i. theory," *IEEE Transactions on Antennas and Propagation*, vol. 38, no. 3, pp. 335–344, 1990, available: https://doi.org/10.1109/8.52240.

[52] E. Slob and J. Fokkema, "Coupling effects of two electric dipoles on an interface," *Radio Science*, vol. 37, no. 5, pp. 6–1–6–10, 2002. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2001RS002529

[53] O. Lopera, E. C. Slob, N. Milisavljevic, and S. Lambot, "Filtering soil surface and antenna effects from gpr data to enhance landmine detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 707–717, 2007, available: https://doi.org/10.1109/TGRS.2006.888136.

[54] S. Lambot, L. Weihermüller, J. A. Huisman, H. Vereecken, M. Vanclooster, and E. C. Slob, "Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content," *Water Resources Research*, vol. 42, no. 11, 2006. [Online]. Available: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006WR005097

[55] D. W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963, available: https://doi.org/10.1137/0111030.

[56] M. Ousmer, A. Sluÿters, N. Magrofuoco, P. Roselli, and J. Vanderdonckt, "Recognizing 3d trajectories as 2d multi-stroke gestures," *Proc. ACM Hum.-Comput. Interact.*, vol. 4, no. ISS, Nov. 2020. [Online]. Available: https://doi.org/10.1145/3427326

[57] S. Lambot and F. André, "Full-wave modeling of near-field radar data for planar layered media reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2295–2303, 2014, available: https://doi.org/10.1109/TGRS.2013.2259243.

[58] "Near-field or far-field full-wave ground penetrating radar modeling as a function of the antenna height above a planar layered medium," *Progress In Electromagnetics Research*, vol. 141, pp. 415–430, 2013, available: https://doi.org/10.2528/PIER13053106.

[59] A. De Coster and S. Lambot, "Fusion of multifrequency gpr data freed from antenna effects," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 2, pp. 664–674, 2018, available: https://doi.org/10.1109/JSTARS.2018.2790419.

[60] G. Zhang, S. Lan, K. Zhang, and L. Ye, "Temporal-range-doppler features interpretation and recognition of hand gestures using mmw fmcw radar sensors," in *2020 14th European Conference on Antennas and Propagation (EuCAP)*, 2020, pp. 1–4, available: https://doi.org/10.23919/EuCAP48036.2020.9135694.

[61] S. Hazra and A. Santra, "Short-range radar-based gesture recognition system using 3d cnn with triplet loss," *IEEE Access*, vol. 7, pp. 125 623–125 633, 2019, available: https://doi.org/10.1109/ACCESS.2019.2938725.

[62] 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. 137 122–137 135, 2019, available: https://doi.org/10.1109/ACCESS.2019.2942305.

[63] N. T. Attygalle, L. A. Leiva, M. Kljun, and K. Čopič Pucihar, "Efficient material selection for training occluded mmWave radar-based gesture recognisers," *Scientific Reports*, Jun. 2026, available: https://doi.org/10.1038/s41598-026-56018-2.

[64] N. T. Attygalle, U. Vuletic, M. Kljun, and K. Č. Pucihar, "Towards hand gesture recognition prototype using the IWR6843ISK radar sensor and leap motion," in *Proceedings of the 8th Human-Computer Interaction Slovenia (HCI SI) Conference 2023*. Maribor, Slovenia: HCI SI, 2024, pp. 78–88.

[65] K. Čopič Pucihar, R.-D. Vatavu, D. Salami, and N. T. Attygalle, Eds., *Radar-Based Human-Computer Interaction*, 1st ed., ser. Human–Computer Interaction Series. Cham: Springer Nature Switzerland, 2026, available: https://doi.org/10.1007/978-3-032-04327-6.

[66] 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 Research, Tech. Rep. MSR-TR-2012-111, Nov. 2012.

[67] R.-D. Vatavu, "ifad gestures: Understanding users' gesture input performance with index-finger augmentation devices," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580928

[68] ADAM. KENDON, "Gesticulation and speech: Two aspects of the process of utterance," in *The Relationship of Verbal and Nonverbal Communication*. Berlin, New York: De Gruyter Mouton, 1980, pp. 207–228, available: https://doi.org/10.1515/9783110813098.207.

[69] M. R. Mahmoudzadeh Ardekani and S. Lambot, "Full-wave calibration of time- and frequency-domain ground-penetrating radar in far-field conditions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 664–678, 2014, available: https://doi.org/10.1109/TGRS.2013.2243458.

[70] 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 *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 5850–5861. [Online]. Available: https://doi.org/10.1145/3025453.3026002

[71] A. Sluÿters, S. Lambot, and J. Vanderdonckt, "Radar-based hand, arm, and body gestures," 2023. [Online]. Available: https://dx.doi.org/10.21227/5n08-v130

[72] M. G. Jacob, J. P. Wachs, and R. A. Packer, "Hand-gesture-based sterile interface for the operating room using contextual cues for the navigation of radiological images," *Journal of the American Medical Informatics Association*, vol. 20, no. e1, pp. e183–e186, Jun. 2013, available: https://doi.org/10.1136/amiajnl-2012-001212.

[73] A. Mewes, B. Hensen, F. Wacker, and C. Hansen, "Touchless interaction with software in interventional radiology and surgery: A systematic literature review," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 2, pp. 291–305, Feb. 2017, available: https://doi.org/10.1007/s11548-016-1480-6.

[74] J. Vanderdonckt, P. Lefevre, R.-D. Vatavu, N. T. Attygalle, L.-B. Bilius, G. Calvary, S. Dupuy-Chessa, M. Fairhurst, T. Nilsson, L. Oppermann, P. Roselli, G. Saggio, and V. Swaen, "The impact of graviception on human-computer interaction," in *Proceedings of the Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '26. New York, NY, USA: Association for Computing Machinery, 2026. [Online]. Available: https://doi.org/10.1145/3772363.3799280

[75] H. Hilgers, J. Vanderdonckt, and R.-D. Vatavu, "Human-AI Interaction in Space: Insights from a Mars Analog Mission with the Harmony Large Language Model," in *Advancing Human-Computer Interaction for Space Exploration (SpaceCHI 2025)*, ser. Open Access Series in Informatics (OASIcs), L. Bensch, T. Nilsson, M. Nisser, P. Pataranutaporn, A. Schmidt, and V. Sumini, Eds., vol. 130. Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2025, pp. 1:1–1:20. [Online]. Available: https://drops.dagstuhl.de/entities/document/10.4230/OASIcs.SpaceCHI.2025.1



Arthur Sluÿters (SM'21) received a Master's degree in Computer Science and Engineering in 2020 with the magna cum laude designation and a Ph.D. degree in Computer Science in 2024 from Université catholique de Louvain, Belgium. He is a scientific collaborator at Louvain Research Institute in Management and Organizations (LouRIM), Université catholique de Louvain (UCLouvain). His research interests include gesture interaction, radar-based gesture recognition, and software engineering.



Nuwan T. Attygalle received a B.Sc. in Physical Science (2017) from the University of Sri Jayewardenepura, Sri Lanka, and Ph.D. in Computer Science (2024) from the University of Primorska, Slovenia. He is a postdoc in computer science at Louvain Research Institute in Management and Organizations (LouRIM), Université catholique de Louvain (UCLouvain). His research interests include interaction techniques, Machine Learning and Human-Computer Interaction.



Sébastien Lambot is a Full Professor (part-time) and FNRS Research Director at the Université catholique de Louvain (UCLouvain), Belgium. He received the M.Sc. and Ph.D. degrees in agricultural and environmental engineering from UCLouvain in 1999 and 2003, respectively, with his doctoral research focusing on full-wave modeling and inversion for ground-penetrating radar (GPR). Following his Ph.D., he was awarded a European Marie Curie Postdoctoral Fellowship and conducted research at Delft University of Technology, The Netherlands, from 2004 to 2005. From 2006 to 2012, he held dual appointments at UCLouvain, where he was appointed Professor, and at Forschungszentrum Jülich, Germany, where he served as a Research Group Leader.

For more than 25 years, Prof. Lambot has been at the forefront of GPR research, with a particular emphasis on advancing full-wave modeling and inversion techniques for digital soil mapping and non-destructive testing. His research interests also include drone-borne GPR, remote sensing, hydrogeophysics, road inspection, and gesture recognition. He has authored or coauthored over 118 peer-reviewed journal articles and has made numerous other contributions to the field. Prof. Lambot has played an active role in the organization of international conferences and workshops, serving as General Chair of the 3rd International Workshop on Advanced Ground Penetrating Radar in 2005 and the 15th International Conference on Ground Penetrating Radar in 2014, and more recently co-chairing IGARSS 2021. He has also served as a Guest Editor for several special issues of *Near Surface Geophysics*, *Vadose Zone Journal*, and the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS)*.



Jean Vanderdonckt (M'04-SM'08) 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, and is a permanent Invited Professor with the Universitat Politècnica de Valencia. He is currently a Full Professor at the 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 (TiiS) and the Co-Editor-in-Chief of the Springer Series of Human–Computer Interaction and the Springer Briefs in Human–Computer Interaction.