











A Radar-Based Concept for Simultaneous High-Resolution Imaging and Pixel-Wise Velocity Analysis for Tracking Human Motion

JOHANNA BRÄUNIG ¹ (Student Member, IEEE), SIMON HEINRICH ², BIRTE COPPERS ^{3,4},
CHRISTOPH KAMMEL ¹ (Student Member, IEEE), VANESSA WIRTH ⁵ (Student Member, IEEE),
MARC STAMMINGER ⁵, SIGRID LEYENDECKER ², ANNA-MARIA LIPHARDT ^{3,4},
INGRID ULLMANN ¹ (Member, IEEE), AND MARTIN VOSSIEK ¹ (Fellow, IEEE)

(Regular Paper)

¹Institute of Microwaves and Photonics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany

²Institute of Applied Dynamics, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany

³Department of Internal Medicine 3 – Rheumatology and Immunology, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, 91054 Erlangen, Germany

⁴Deutsches Zentrum Immuntherapie, Friedrich-Alexander-Universität Erlangen-Nürnberg and Universitätsklinikum Erlangen, 91054 Erlangen, Germany

⁵Chair of Visual Computing, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany

CORRESPONDING AUTHOR: Johanna Bräunig (e-mail: johanna.braeunig@fau.de).

This work was supported in part by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1483 – Project-ID 442419336, EmpkinS; and in part by DFG which funded the other major instrument at the Institute of Applied Dynamics, Friedrich-Alexander-Universität Erlangen-Nürnberg, used in this study (reference number INST 90/985-1 FUGG).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Ethics Committee of Friedrich-Alexander-Universität Erlangen-Nürnberg, Application No. #357 20B, and performed in line with the principles of the revised Declaration of the World Medical Association of Helsinki.

ABSTRACT The radar-based analysis of human motion is actively being researched due to its contact- and markerless nature and ability to measure motion directly via the Doppler effect. Especially in medical and biomechanical fields, precise movement analysis is crucial. However, existing radar-based studies typically exhibit low lateral resolution, focusing on velocity evaluations and the tracking of scattering centers resolvable in the range or Doppler domains. In this work, we present a novel concept that enables a pixel-wise velocity analysis of human motion in radar near-field imaging scenarios. For this, we utilize the well-established back-projection technique to reconstruct consecutive radar images and perform a subsequent pixel-wise phase comparison. To accurately capture pixel-specific velocities along the depth dimension, this is followed by corrections of near-field geometry distortions accounting for aperture properties and pixel positions. Our theoretical derivations are supported by comprehensive point target simulations. To assess the performance of the proposed approach, we conducted a proof-of-concept study. We tracked a hand surface's movement while performing a finger tapping motion and compared the fingertip position and velocity determined by the radar with the respective values obtained from an optical marker-based system. The results showed a velocity measurement accuracy of 8.1 mm s^{-1} and a tracking accuracy of 1.4 mm, demonstrating the great potential of our approach. The high angular resolution of the velocity measurement enables the tracking of the entire illuminated body shell, extending the range of future applications of radar-based motion analysis.

INDEX TERMS Human motion tracking, near-field imaging, radar imaging, velocity measurement.

I. INTRODUCTION

The use of radar for the detection, tracking, and analysis of human motion offers a multitude of advantages

over conventional accelerometer-based and optical methods. What mainly sets it apart is its markerless, contactless, and light-independent working principle. Moreover, the

technology preserves privacy and tracks motion by directly measuring distances, angles, and velocities through phase evaluations, offering significant advantages over markerless camera-based systems. Radar signal returns from a moving human body are subject to Doppler modulations [1]. Due to the nonrigid nature of the human body, different body parts are characterized by specific motion trajectories and velocities, resulting in individual Doppler components. Over the last decade, this technology has rapidly gained prominence in sensing human motion, spanning research fields such as physical activity [2], [3], [4] and gesture [5] recognition, gait analysis [6], and hand or finger tracking [7]. In activity or gesture recognition, the focus is not on the precise extraction of parameters, such as speed or position data, to track different body parts. Instead, movement patterns are identified and classified using simplified representations of radar data that reveal specific patterns, such as micro-Doppler signatures [1], commonly in conjunction with machine learning or deep learning architectures. However, many applications—for instance, in the fields of medicine, biomechanics, sports, and robotics—as well as fine gesture control require the precise analysis of human motion in terms of position and velocity. State-of-the-art systems for precise human motion tracking are so-called marker-based optoelectronic measurement systems, whose accuracy is higher than that of markerless technologies [8]. Such systems currently reach submillimeter accuracy at capture frequencies up to 500 Hz [9].

However, radar technology is increasingly being used as a markerless alternative for precisely tracking human movements. One field of application is gait or lower limb motion analysis, in which simple (pulse-)Doppler radars are often used [6]. Evaluation is performed using spectrograms, which depict all Doppler components along the time axis by applying a short-time Fourier transform to the data to extract relevant parameters, such as torso, knee, foot, or ankle velocity [10], [11], [12], [13], [14], [15], [16], [17]. These studies have in common that different targets are resolved using high velocity resolution and distinguishable signal powers related to the respective radar cross-sections of the body parts. In [18] and [19], multiple-input multiple-output (MIMO) frequency-modulated continuous-wave (FMCW) radars consisting of multiple transmitting (Tx) and receiving (Rx) antennas were used to resolve individual legs within range-Doppler maps and to extract the corresponding foot velocities. Moreover, angle information obtained by applying digital beamforming to the resolved targets was included in the evaluation. More recent works have combined MIMO FMCW radars with low lateral resolution with deep-learning networks to estimate human postures based on multidimensional radar data streams [20], [21]. The authors of [22] did not only extract joint positions but also velocities from the radar data.

The radar-based tracking of hand motions is also being researched. Various studies have used radar to track one human hand as one target [23], [24], [25], [26]. In [7], Google introduced the project Soli, in which an FMCW radar with

very high bandwidth of 7 GHz and, consequently, a high radial resolution of approximately 2 cm, was used for gesture recognition and precise gesture sensing by tracking distinct scattering centers of a hand. To allow this tracking, the scattering centers first need to be resolved in the Doppler or range dimension, which restricts its application to certain hand poses and motions. However, the resolution of all fingers or different finger parts can be relevant in the context of human-computer interaction [27], [28]. Furthermore, it plays an important role in the analysis of human hand kinematics, with applications in the fields of biomechanics, prosthetics [29], [30], and robotics [31], as well as in the clinical context of musculoskeletal disorders [32], [33], [34], [35], [36], [37]. In [32], a simple Doppler radar was used to compare the fingertip velocity of patients with rheumatoid arthritis during a finger tapping motion with that of healthy controls for clinical hand function analysis. The spectrogram-based approach was found to be inadequate, highlighting the need for more fine-grained radar-based techniques for measuring hand motions.

All of the abovementioned works incorporate either a very limited or no lateral resolution. This limitation restricts their applications to tasks focused on tracking specific scattering centers rather than the entire body shell. As body parts must be resolved in the Doppler or range domain, such approaches are applicable to human limbs to some extent but hardly apply to individual fingers of the human hand.

Increasing the angular resolution offers two major advantages. First, it enables a more detailed representation of the human body that closely resembles our visual perception. This can be helpful for the segmentation of body parts, as image-based approaches can be more easily transferred to the radar domain. Second, high angular resolution offers new possibilities for tracking the entire illuminated body surface. High-resolution radar imaging is the state of the art in several near-field applications, such as ground-penetrating radar, radar cross-section imaging, and personnel security scanning [38], [39], [40]. Radar signals have primarily been used to capture high-resolution spatial information on target scenes without directly evaluating motion through phase changes over time. Low frame rates have been a limiting factor in this regard, especially in MIMO radar systems that use many Tx antennas, each transmitting in a time-division-multiplex (TDM) manner to create large synthetic apertures. Moreover, high signal bandwidths, used for achieving high radial resolution, result in longer measurement acquisition times, posing a challenge when evaluating Doppler shifts for motion analysis. To increase frame rates with existing hardware, a potential solution is to leverage the high angular resolution while reducing the radial resolution and the associated number of frequency samples that define the range ambiguity. In [41], we proposed a novel ultra-efficient method for the 3D reconstruction of the human body shell, specifically applied to hand surfaces, using millimeter wave near-field imaging techniques and a very low bandwidth.

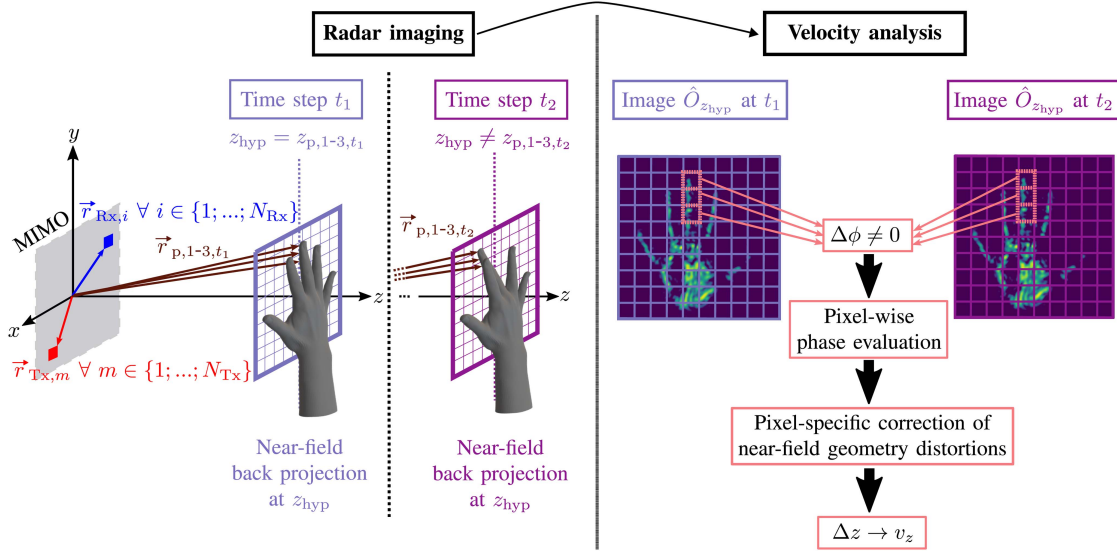


FIGURE 1. Overview of the pixel-wise velocity analysis along the depth dimension z of a MIMO radar based on near-field imaging techniques. At time step t_1 , the estimated depths z_{hyp} for three exemplary pixels of the back-projected image correspond to the depths of the scatterers $z_{p,1-3,t_1}$. At time step t_2 , the scatterers have moved along the depth dimension to $z_{p,1-3,t_2}$, while back projection is performed with the prior hypothesis value z_{hyp} . This leads to a changing residual phase of the complex pixel information between t_1 and t_2 . As we are in the near field of an imaging MIMO radar, this pixel phase change cannot be directly translated into classical “radial” motion (referring to a monostatic case). Instead, the phase change is converted into a velocity estimation v_z along the depth dimension, accounting for near-field geometry distortions.

Building on [41], in this paper we present a novel concept for the simultaneous radar-based near-field imaging and tracking of human motion. This approach enables velocity measurements and comprehensive tracking of the entire illuminated body surface. Integrating radar-based high-resolution imaging with velocity measurements necessitates considering near-field effects due to the large aperture dimensions. Hence, conventional Fast Fourier Transform (FFT)-based algorithms, which assume far-field conditions, cannot be employed. Consequently, the key contributions of this paper are as follows:

- We present a novel methodology that allows the analysis of pixel-specific laterally resolved velocities along the depth dimension in near-field radar imaging applications by evaluating pixel phase changes over time.
- We validated the developed approach through comprehensive simulations of moving point targets.
- We conducted a proof-of-concept measurement study to assess the performance of the proposed velocity estimation. We used the pixel-specific velocity information to track the motion of the hand surface and compared the velocity and position data obtained through the radar-based approach with those obtained by a gold-standard marker-based camera system.

II. PIXEL-WISE VELOCITY ANALYSIS

The basic principle used for the pixel-wise velocity analysis of the human body shell is illustrated in Fig. 1, using hand motion as an example. Initially, the hand surface is localized using, for instance, the principle described in [41]. For time step t_1 , a 2D image consisting of N_p pixels is reconstructed at the pixel-specific depths $z_{\text{hyp}}(x, y)$ by performing the near-field back-projection algorithm [38], [40]. Depending on the

specific hand pose, $z_{\text{hyp}}(x, y)$ can either vary across pixels or remain constant. The image is reconstructed for one frequency step or, to achieve a higher SNR, for multiple closely neighboring frequency steps. For time step t_2 , $z_{\text{hyp}}(x, y)$ is kept equal to the values from t_1 . If some parts of the hand have changed their locations, the hypotheses do not match the measured signals and the residual pixel phase that remains after performing back projection changes in relation to the motion. For monostatic radar systems, phase shifts can directly be translated into radial motion and, hence, a motion component along the radar’s line of sight. For imaging MIMO radars, defining one radial direction is challenging, as such radars consist of multiple antennas at different locations; therefore, each antenna measures a different Doppler shift. For this reason, we conduct a pixel-specific phase evaluation, followed by a correction of near-field distortions caused by the MIMO array geometry and pixel positions. This allows motion measurements along the depth dimension of the MIMO radar coordinate system. The derivation of this correction is described in the following subsection.

A. PIXEL-SPECIFIC GEOMETRY CORRECTION

We assume a MIMO radar consisting of N_{Rx} receive and N_{Tx} transmit antennas in positions $\vec{r}_{\text{Tx},m} = (x_m, y_m, z_m)$ for $m \in \{1; \dots; N_{\text{Tx}}\}$ and $\vec{r}_{\text{Rx},i} = (x_i, y_i, z_i)$ for $i \in \{1; \dots; N_{\text{Rx}}\}$. The target scene consists of a human body or a body part, which can be described as a distributed nonrigid target comprised by N_p point scatterers in positions $\vec{r}_{p,k}$ for $k \in \{1; \dots; N_p\}$ within the near field of the antenna array. Each Tx antenna sends N_f closely neighboring frequency steps denoted by f_s for $s \in \{1; \dots; N_f\}$. For simplicity, we assume $N_f = 1$ for the following derivation. The received beat signals for each Tx–Rx

combination, assuming one point target at $\vec{r}_{p,k,t_1} = (x, y, z)$, for time step t_1 are denoted as

$$s_{B,im}(\vec{r}_{p,k,t_1}) = A \exp \left(-j2\pi f_s \frac{d_{t_1,im}}{c} + \phi_{Rx} \right), \quad (1)$$

where A describes the amplitude of the received signal, and c is the speed of light. The phase of the received signal depends on $d_{t_1,im}$, which denotes the total path length from each Tx–Rx pair to the scatterer, and an additional unknown phase term ϕ_{Rx} . In the following, we assume $A = 1$ and $\phi_{Rx} = 0$ for the sake of readability.

As we are in the near field of the MIMO radar, FFT-based approaches are not sufficient to obtain spacial information about the target. Therefore, back projection is performed as in [38]. In our case, this is not done for an entire volume but for the estimated body surface, which can be described as a 2D depth map in which each pixel contains an estimated depth coordinate denoted as z_{hyp} . For the reconstruction of the pixel position $\vec{r}_{p,lk} = (x, y, z_{hyp})$ for $l \in \{1; \dots; N_{\bar{p}}\}$ that matches the x and y coordinates of the k th point target, the baseband signals are correlated with the matched filter weights $w_{im}(\vec{r}_{p,lk})$, denoted as

$$w_{im}(\vec{r}_{p,lk}) = \exp \left(j2\pi f_s \frac{d_{hyp,im}}{c} \right), \quad (2)$$

where $d_{hyp,im}$ describes the total path length from each Tx–Rx combination to the pixel. Assuming that we perfectly localize the body surface for t_1 , we get $z = z_{hyp}$ and, therefore,

$$d_{t_1,im}(\vec{r}_{p,k,t_1}) = d_{hyp,im}(\vec{r}_{p,lk}). \quad (3)$$

To reconstruct the pixel $\vec{r}_{p,lk} = (x, y, z_{hyp})$, we perform back projection [38] as

$$\hat{O}_{t_1}(\vec{r}_{p,lk}) = \sum_{s=1}^{N_f} \sum_{m=1}^{N_{Tx}} \sum_{i=1}^{N_{Rx}} s_{B,im}(\vec{r}_{p,k,t_1}) \cdot w_{im}(\vec{r}_{p,lk}). \quad (4)$$

For t_1 , we obtain the complex pixel information $\hat{O}_{t_1}(\vec{r}_{p,lk})$. Due to the relationship described in (3), the residual pixel phase $\phi_{res} = \arg \hat{O}_{t_1}(\vec{r}_{p,lk})$ is zero (or equal to a constant value caused by ϕ_{Rx}).

For t_2 , we assume that the point target has moved along the depth dimension. However, the depth value z_{hyp} for the reconstructed pixel $\vec{r}_{p,lk}$ remains equal to that at t_1 . This leads to a change in the residual pixel phase ϕ_{res} , which can be used to estimate the velocity along depth. The new point target position at t_2 is

$$\vec{r}_{p,k,t_2} = (x, y, z - \Delta z) = (x, y, z_{hyp} - \Delta z). \quad (5)$$

Hence, for the baseband signals at t_2 , we get

$$s_{B,im}(\vec{r}_{p,k,t_2}) = \exp \left(-j2\pi f_s \frac{d_{t_2,im}}{c} + \phi_{Rx} \right), \quad (6)$$

with

$$d_{t_2,im} = \sqrt{(x_m - x)^2 + (y_m - y)^2 + (z_m - z_{hyp} + \Delta z)^2} + \sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z_{hyp} + \Delta z)^2}. \quad (7)$$

In the following, we assume that all antenna positions are located at the same depth $z_{i/m} = 0$ of the MIMO radar coordinate system. Hence, (7) can be rewritten as

$$d_{t_2,im} = \sqrt{d_m^2 - 2z_{hyp}\Delta z + \Delta z^2} + \sqrt{d_i^2 - 2z_{hyp}\Delta z + \Delta z^2}, \quad (8)$$

where $d_{i/m} = \sqrt{(x_{i/m} - x)^2 + (y_{i/m} - y)^2 + z_{hyp}^2}$ describes the Rx/Tx term of the round-trip distance to the pixel position. At t_2 , the distances $d_{hyp,im}$ remain the same as at t_1 and can also be expressed by a Tx/Rx term as

$$d_{hyp,im} = d_m + d_i, \quad (9)$$

which allows us to separate the back projection for pixel $\vec{r}_{\bar{p}}$ for t_2 into a Tx- and a Rx-dependent part as

$$\begin{aligned} \hat{O}_{t_2}(\vec{r}_{p,lk}) &= \sum_{i=1}^{N_{Rx}} \exp \left(-j2\pi f_s \frac{f_s}{c} \sqrt{d_i^2 - 2z_{hyp}\Delta z + \Delta z^2} \right) \\ &\quad \cdot \exp \left(j2\pi f_s \frac{f_s}{c} d_i \right) \\ &\quad \cdot \sum_{m=1}^{N_{Tx}} \exp \left(-j2\pi f_s \frac{f_s}{c} \sqrt{d_m^2 - 2z_{hyp}\Delta z + \Delta z^2} \right) \\ &\quad \cdot \exp \left(j2\pi f_s \frac{f_s}{c} d_m \right), \end{aligned} \quad (10)$$

which gives

$$\begin{aligned} \hat{O}_{t_2}(\vec{r}_{p,lk}) &= \sum_{i=1}^{N_{Rx}} \sum_{m=1}^{N_{Tx}} \exp \left(j2\pi f_s \frac{f_s}{c} \right. \\ &\quad \cdot \left[d_i - \sqrt{d_i^2 - 2z_{hyp}\Delta z + \Delta z^2} \right. \\ &\quad \left. \left. + d_m - \sqrt{d_m^2 - 2z_{hyp}\Delta z + \Delta z^2} \right] \right). \end{aligned} \quad (11)$$

We approximate both square root terms in (11) using a first-order Taylor series expansion of the term $\sqrt{a+x}$ at $x = 0$ with $a = d_{i/m}^2$ and $x = -2z_{hyp}\Delta z + \Delta z^2$. As Δz yields small values and x is, therefore, close to 0, the distance term influencing $\arg \hat{O}_{t_2}(\vec{r}_{\bar{p}})$ can be approximated by

$$\begin{aligned} d_{approx} &= d_i - \left(\sqrt{d_i^2 - 2z_{hyp}\Delta z + \Delta z^2} \right) \\ &\quad + d_m - \left(\sqrt{d_m^2 - 2z_{hyp}\Delta z + \Delta z^2} \right) \end{aligned}$$

$$= \frac{1}{d_i} \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right) + \frac{1}{d_m} \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right). \quad (12)$$

Further, we rewrite d_{approx} as

$$d_{\text{approx}} = \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right) \cdot \left(\frac{1}{d_{\varnothing_{\text{Rx}}} + \Delta d_i} + \frac{1}{d_{\varnothing_{\text{Tx}}} + \Delta d_m} \right), \quad (13)$$

where $d_{\varnothing_{\text{Rx/Tx}}}$ represents the mean distance term over all Rx/Tx antennas. The term $\Delta d_{i/m}$ describes the respective deviations from the mean distance term for each Rx/Tx antenna. As $\Delta d_{i/m}$ yields small values compared to $d_{\varnothing_{\text{Rx/Tx}}}$, we use another first-order Taylor expansion for the term $\frac{1}{a+x}$ at $x = 0$, with $a = d_{\varnothing_{\text{Rx/Tx}}}$ and $x = \Delta d_{i/m}$. Hence, we get

$$d_{\text{approx}} = \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right) \cdot \left(\frac{1}{d_{\varnothing_{\text{Rx}}}} - \frac{\Delta d_i}{d_{\varnothing_{\text{Rx}}}^2} + \frac{1}{d_{\varnothing_{\text{Tx}}}} - \frac{\Delta d_m}{d_{\varnothing_{\text{Tx}}}^2} \right). \quad (14)$$

For the near-field radar imaging scenarios considered in this work, we can assume that $\frac{\Delta d_{i/m}}{d_{\varnothing_{\text{Rx/Tx}}}^2} \approx 0$. In the case of symmetric Tx and Rx apertures of equal length, we further assume $d_{\varnothing_{\text{Rx}}} = d_{\varnothing_{\text{Tx}}}$, which leads to

$$d_{\text{approx}} = \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right) \cdot \frac{2}{d_{\varnothing_{\text{Tx}}}}. \quad (15)$$

To calculate Δz , we evaluate the phase difference $\Delta\phi$ between consecutive time steps as

$$\Delta\phi = \arg \widehat{O}_{t_2}(\vec{r}_{\text{p},l_k}) - \arg \widehat{O}_{t_1}(\vec{r}_{\text{p},l_k}) \approx 2\pi \frac{f_s}{c} \frac{2}{d_{\varnothing_{\text{Tx}}}} \left(z_{\text{hyp}} \Delta z + \frac{1}{2} \Delta z^2 \right). \quad (16)$$

With $\sigma = \frac{c\Delta\phi}{4\pi f_s}$, we get

$$0 = \frac{1}{2d_{\varnothing_{\text{Tx}}}} \Delta z^2 + \frac{z_{\text{hyp}}}{d_{\varnothing_{\text{Tx}}}} \Delta z - \sigma, \quad (17)$$

which is a quadratic equation. Hence, Δz can be solved by applying the quadratic formula. With this, each phase difference can be used to calculate Δz , which translates into a z -velocity when divided by the frame duration.

With (16), it further becomes clear that $\Delta\phi$ and, therefore, the maximum unambiguous velocity along z , $v_{z,\text{unamb}}$, is influenced by the parameters $d_{\varnothing_{\text{Tx}}}$ and z_{hyp} . Hence, to calculate $v_{z,\text{unamb}}$, the quadratic equation in (17) must be solved for $\Delta\phi = 360^\circ$ and the pixel-specific parameters $d_{\varnothing_{\text{Tx}}}$ and z_{hyp} .

B. SIMULATION

We performed simulations to check the validity of the previously described derivations regarding the pixel-specific

evaluation of v_z . Specifically, we simulated two different scenes and geometries including one or multiple MIMO radar clusters consisting of 94 Tx and 94 Rx antennas each (see Fig. 2(a)). The scenes are illustrated in Fig. 2(b) and (c). Each simulation involves a MIMO radar and point scatterers moving along the z -axis. We simulated the point scatterers individually. Hence, each simulation consisted of a single point scatterer. All evaluated positions are shown in Fig. 2. Scene 1 simulates a single MIMO radar cluster. We also used this geometry for the proof-of-concept measurement. The MIMO radar is a single cluster of a commercially available radome tester [42]. The respective antenna positions are shown in Fig. 2(a). The physical aperture spans approximately $14 \text{ cm} \times 14 \text{ cm}$, with 3 mm spacing between adjacent antennas. Scene 1 was designed to demonstrate applicability to measurement scenarios of a moving hand and fingers. Hence, we set the target velocity to $v_z = 1 \text{ ms}^{-1}$, which, according to [32], is a valid velocity of a fast-moving finger. As a hand is a small target compared to an entire body, it can be brought close to the MIMO radar, at a depth of 30 cm, where the lateral resolution is about 4 mm [40]. To resolve fingers or finger parts in the lateral dimension, the resolution should be below 1 cm. Therefore, the lateral resolution of this MIMO configuration at close distances is well-suited for capturing hand motions. The scatterers' positions vary on a grid defined by $x/y \in [-10 \text{ cm}, 10 \text{ cm}]$. For the second simulation, we used three single clusters to increase resolution and illumination for a theoretical full-body measurement. Arranged at a depth of $z = 1.5 \text{ m}$ the point scatterers were moving at a velocity of $v_z = 3 \text{ ms}^{-1}$, which, as shown in [43], is a plausible toe velocity of a walking human. The scatterers' positions varied on a grid defined by $x \in [-30 \text{ cm}, 30 \text{ cm}]$ and $y \in [-1 \text{ m}, 1 \text{ m}]$. It should be noted that such a radar system with sufficient frame rates does not currently exist. The simulation scene was designed to demonstrate that our derivations are applicable to larger apertures and whole-body measurements, providing insight into potential applications. Full-body or gait measurements can also be performed by using only one single cluster with lower lateral resolution.

For both simulations, we assumed a single continuous-wave (CW) frequency of $f_s = f_c = 80 \text{ GHz}$ and a frame duration of $200 \mu\text{s}$ to ensure a sufficient value for $v_{z,\text{unamb}}$ in both cases. To illustrate the efficiency of the proposed concept, we compared the estimated velocities obtained by solving (17) for Δz and calculating v_z with the velocity estimation that directly translated $\Delta\phi$ into an estimate of v_z . The latter was inspired by the traditional method of calculating the radial velocity in far-field scenarios. To obtain $\Delta\phi$, we generated two radar images at two consecutive time steps by performing near-field back projection for f_c at depth z_{hyp} , which was $z = 0.3 \text{ m}$ and $z = 1.5 \text{ m}$ for Simulation 1 and 2, respectively. We then calculated $\Delta\phi$ for the pixel position that had the same x - and y -coordinates as the current target position. Finally, we translated $\Delta\phi$ into Δz by solving (17) and subsequently calculated v_z . The results for Simulation Scene 1 are shown in Fig. 3. For the far-field estimation

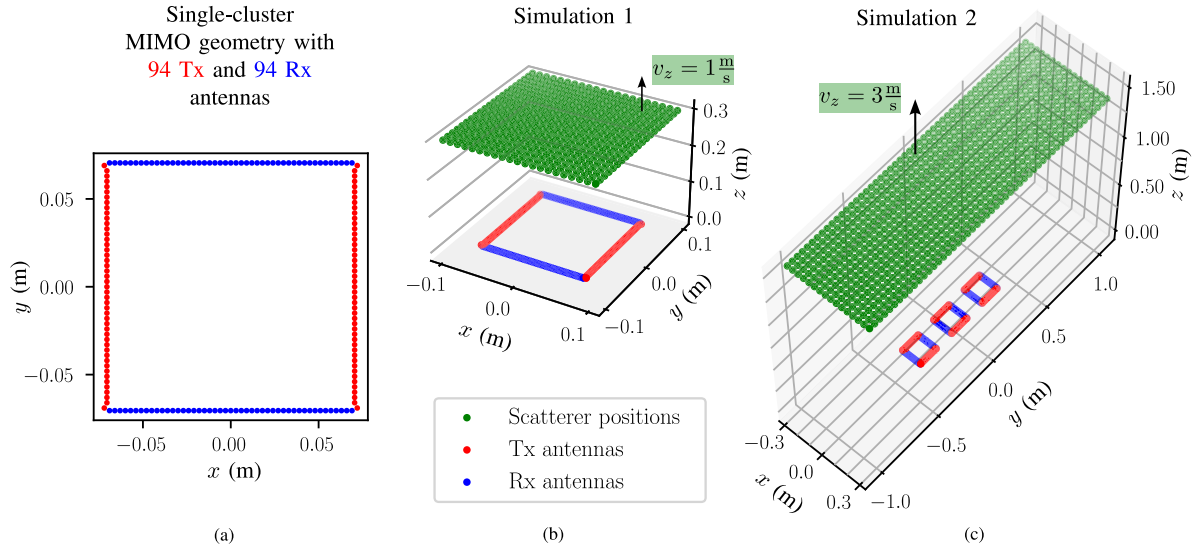


FIGURE 2. Overview of the simulation setup. (a) Single cluster MIMO geometry consisting of 94 Tx and 94 Rx antennas. The physical aperture is approximately $14 \text{ cm} \times 14 \text{ cm}$. (b) In Simulation 1, one single cluster is used. Point targets are placed at a depth of $z = 0.3 \text{ m}$ and move away from the aperture with a speed of $v_z = 1 \text{ ms}^{-1}$. This simulation is intended to prove the proposed method's applicability to hand and finger tracking scenarios. (c) In Simulation 2, three single clusters are vertically aligned. The simulated point targets are placed at a depth of $z = 1.5 \text{ m}$ and move at a speed of $v_z = 3 \text{ ms}^{-1}$. Simulation 2 aims to show how our derivations apply to larger apertures, relevant for full-body measurements.

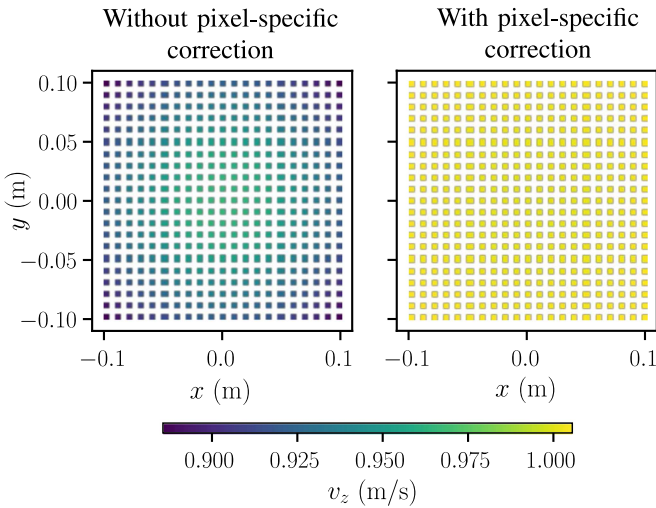


FIGURE 3. Velocity estimation for Simulation Scene 1 without and with pixel-specific correction of near-field geometry distortions. The ground-truth velocity of the point targets is $v_z = 1 \text{ ms}^{-1}$.

of v_z , the values ranged from 0.886 ms^{-1} to 0.966 ms^{-1} . Around $x = y = 0$, the estimated value was the closest to the simulated ground truth of 1.0 ms^{-1} but the true value was never reached, suggesting that the definition of a radial axis is problematic for this near-field application. These findings are further supported by the results from Simulation Scene 2, which is shown in Fig. 4. In this case, the estimated velocity values ranged from 2.461 ms^{-1} to 2.964 ms^{-1} , with the most accurate value also occurring around $x = y = 0$ not reaching the ground-truth value. Figs. 3 and 4 also show the estimated

velocities along the z -coordinate using the proposed pixel-specific correction. For Simulation 1, the values ranged from 1.0 ms^{-1} to 1.005 ms^{-1} , with the ground truth being perfectly matched around $x = y = 0$. In Simulation 2, v_z ranged from 3.002 ms^{-1} around $x = y = 0$ to 3.017 ms^{-1} . The deviations from the ground truth in both simulations when applying the correction were most likely related to the prior simplifications introduced by the Taylor approximations in (12) and (14). These simulation results demonstrate that the derivations described in Section II-A are valid and that, when assuming only a velocity component along the depth dimension, we can use radar imaging to detect its pixel-specific value.

III. MEASUREMENT VERIFICATION

To demonstrate the proposed method's applicability to realistic human motion, we conducted a proof-of-concept measurement study, including the measurement of simple hand movements. The study protocol was approved by the Ethics Committee of Friedrich-Alexander-Universität Erlangen-Nürnberg (no. #357_20B). The participants were included after given written informed consent. In this study, we assessed the accuracy of the proposed pixel-specific velocity-evaluation method and its potential to track the human body shell.

A. MEASUREMENT SETUP AND STUDY DESIGN

In this measurement study, we used the TDM MIMO radar shown in Fig. 5, which also provided the geometry used in Simulation 1. At a distance of 30 cm, this MIMO radar achieves a lateral resolution of approximately 4 mm, enabling the differentiation of individual finger parts. This makes the

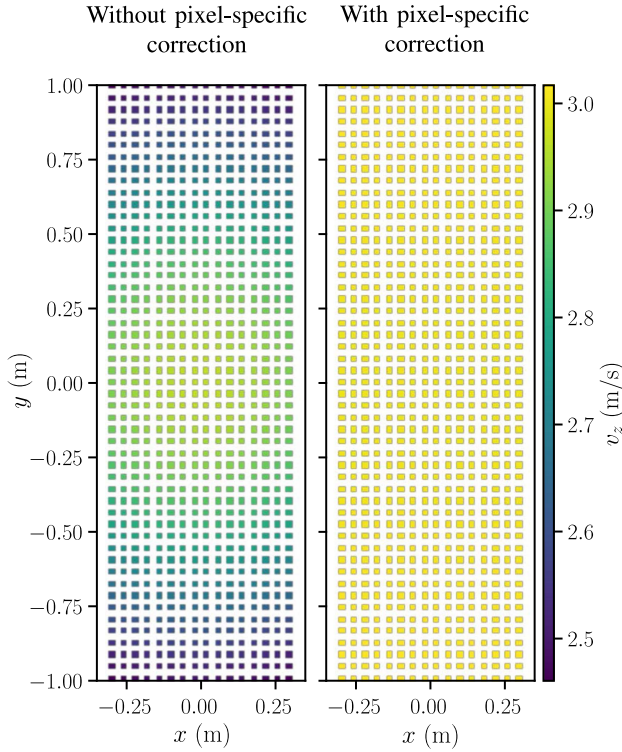


FIGURE 4. Comparison of the velocity estimation for Simulation Scene 2 without and with pixel-specific correction of near-field geometry distortions. The ground-truth velocity of the moving point targets is $v_z = 3 \text{ ms}^{-1}$.

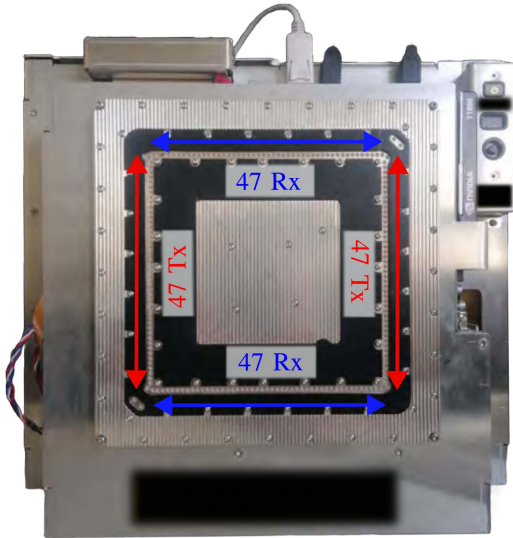


FIGURE 5. MIMO radar used in the measurement study. The array consists of 94 Tx and 94 Rx antennas.

MIMO architecture highly suitable for analyzing hand motion. The radar implements a stepped-frequency continuous-wave (SFCW) signal shape with an adjustable bandwidth and number of equidistant frequency steps. We used two frequency steps, $f_1 = 79.8 \text{ GHz}$ and $f_2 = 80 \text{ GHz}$, per radar frame. This resulted in a mean frame duration of 2.5 ms,

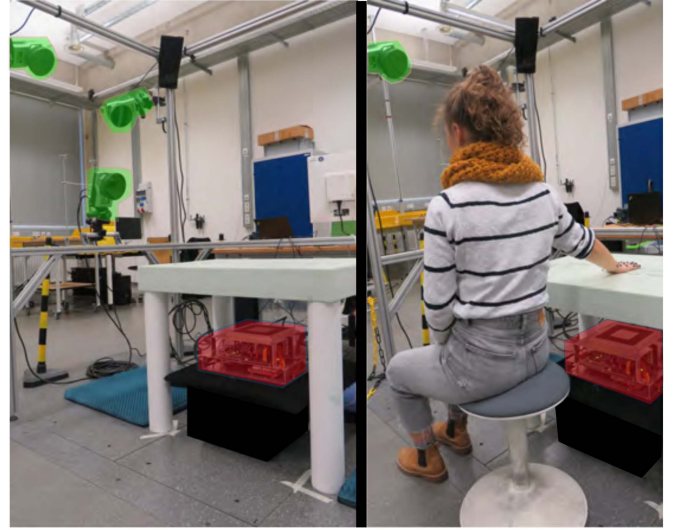


FIGURE 6. Measurement setup. The MIMO imaging radar (red) was positioned under a Styrodur table, which is transparent to millimeter waves. The marker-based camera system (green), composed of six Oqus Series 7+ cameras (12 MP), was arranged around the Styrodur table using a self-built aluminum frame. The participant sat next to the Styrodur table, placing the right hand in a designated area on the table, centered on the MIMO radar's coordinate system origin.

which, for example, for a pixel position at $(0,0,0.3)$ yields a maximum unambiguous z -velocity of $v_{z,\text{unamb}} = 0.77 \text{ ms}^{-1}$, when solving (17) for $z_{\text{hyp}} = 0.3 \text{ m}$, $d_{\text{Tx}} = 0.31 \text{ m}$, and $\Delta\phi = 2\pi$. It should be noted that the current firmware and data transfer protocols of the radar system are not optimized for achieving the maximum measurement rate. However, given that the switching time for all Tx antennas in the specified radar configuration is approximately $13 \mu\text{s}$, it is evident that the frame rate can be significantly increased with future firmware adaptations.

We used an optical marker-based camera system [44] (Qualisys AB, Sweden) consisting of six Oqus Series 7+ cameras (12 MP; 4096×3072 pixels) as a reference system. The measurement setup is shown in Fig. 6. The MIMO imaging radar was placed on a box under a Styrodur table, which is transparent to millimeter waves. The marker-based camera system was arranged around the table using a self-built aluminum frame for mounting. The participants sat at the Styrodur table and placed their right hand directly above the MIMO radar (right side of Fig. 6). The area of best illumination, which is centered on the origin of the MIMO radar's coordinate system, was marked on the table, and the hand was placed inside this area. The distance between the MIMO frontend and the hand was approximately 30 cm. We synchronized the radar and optical reference system using a self-designed synchronization board¹ so that both sensors had a common start trigger. As an exemplary movement, each participant performed a simple finger tapping motion of cyclically raising and lowering a straight finger by flexion

¹[Online]. Available: <https://github.com/empkins/empkins-sync-board>

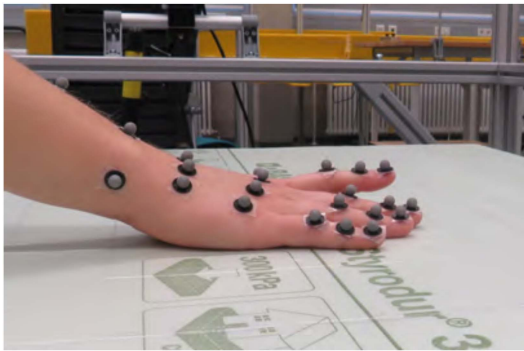


FIGURE 7. Finger tapping motion and optical marker placement. The tapping motion is performed by periodic flexion-extension of the metacarpophalangeal joint of the index or middle finger. The markers are placed on the dorsum of the hand to avoid interfering with the radar measurement.

and extension in the metacarpophalangeal joint while keeping the rest of the hand on the table. The marker placement and tapping motion are illustrated in Fig. 7. We selected this motion because it provided constant visibility of the entire finger in the radar image, allowing an easy comparison with the reference system. For more complex movements, some parts of the hand may not return a signal due to specular reflection effects, which may cause difficulty if one wants to map the marker-based data to certain parts of the radar image. The optical markers were placed on the top of the hand, to avoid interference with the radar measurement that captured the bottom of the hand. We measured the right hands of two healthy female participants aged 29 ± 1 years with no history of hand injuries and a hand length of 177.5 ± 12.5 mm, and a hand width of 85 ± 5 mm. Each participant performed the requested movement five times with both the index and middle fingers, resulting in a total of 20 measurements, each lasting about 10 s. To validate the radar measurements and to assess the accuracy of the proposed approach, we compared the fingertip velocity measured by the radar and reference systems. To ensure, that the fingertip movement would not exceed $v_{z,unamb}$, we instructed both participants to perform the tapping at a moderate speed.

B. RADAR SIGNAL PROCESSING

Each radar measurement consisted of 4000 radar frames with a mean frame duration of 2.5 ms. As the radar's frame rate varies to some extent, timestamps are included in the data to allow for better synchronization. For each radar frame, the two carrier frequencies, $f_1 = 79.8$ GHz and $f_2 = 80$ GHz, were sent by all Tx antennas. To calibrate the MIMO radar, we first conducted a measurement using a metal plate positioned parallel to the radar's frontend at a known distance. This initial step corrects for phase and amplitude variations in the signals transmitted and received by different channels. We then performed an "empty scene" measurement to address any residual effects from the measurement environment. This calibration method has proven effective for our radar

imaging application. A more sophisticated MIMO calibration approach that also addresses mutual coupling effects between channels can be found in [45]. An overview of the radar signal processing step is shown in Fig. 8 and will be described in detail in the following.

1) LOCALIZATION

To localize the hand, we followed the procedure described in [41]. The frequency difference between the two CW signals of $\Delta f = 200$ MHz forms the basis of this evaluation, allowing a high unambiguous range of the first coarse localization. Thus, the hand is, first, coarsely localized by reconstructing two single-tone images at the estimated depth of the hand z_e for the first radar frame included in the evaluation. Second, the phase difference between the two single-tone images of the pixel containing the strongest scatterer (from the f_1 image) is evaluated to adjust z_e . This coarse hand localization is performed only once per measurement by evaluating the first frame. The remaining frames are then divided into sequences of length $N_{fr} = 15$. For each frame within a sequence, two single-tone images are reconstructed at the adjusted depth z_e . Subsequently, the phase difference between these single-tone images is evaluated for every pixel with an amplitude surpassing the predefined threshold of -13 dB after normalization to the maximum occurring amplitude. Thus, a depth map $z(x, y)$ for each radar frame is obtained to estimate the 3D hand surface. These depth maps, along with velocity estimate per sequence for each pixel, are used to track the hand surface.

2) VELOCITY ESTIMATION

The first depth map of a sequence is used to define $z_{hyp}(x, y)$ for each pixel containing a target, forming the basis for the velocity estimation in the sequence. In the next step, $N_{fr} = 15$ consecutive radar two-tone images are reconstructed at the pixel-specific depths $z_{hyp}(x, y)$ by performing the back-projection method described in (4) using both CW frequencies ($N_f = 2$) to obtain a higher SNR. It should be noted that in the case of the tapping motion measured in this experiment, z_{hyp} was kept identical for all pixel positions, as the hand was placed flatly on the Styrodur table, and the range of motion of the moving finger was kept small. In general, z_{hyp} can vary across pixel positions, which becomes relevant if you want to maximize the SNR for hand poses that cover a wide depth range. After the reconstruction of all two-tone images in a respective sequence, the pixel-specific phase differences between consecutive images for all pixels that contain a target are evaluated and averaged for all frames within the sequence. This averaged phase difference is then translated into a Δz -value using the pixel-specific geometry correction described in Section II-A. Fig. 9 shows the evaluation results of an exemplary sequence, showing the locally distributed z -velocity across the hand during the tapping motion of the index finger. The figure demonstrates the system's ability to distinguish between velocities of different parts of the hand and depicts the entire illuminated hand surface. The linearly increasing

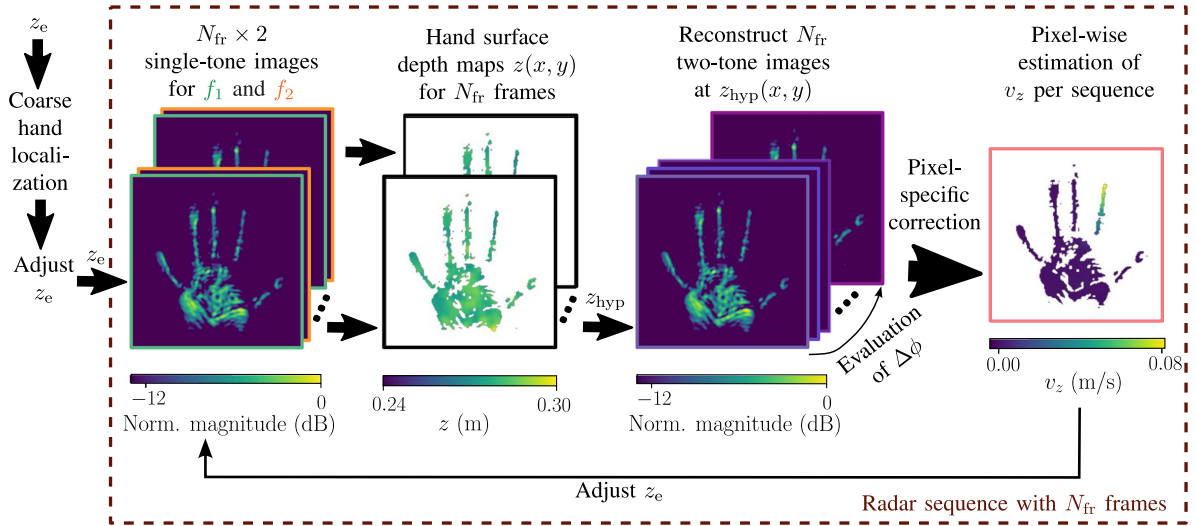


FIGURE 8. Overview of post-processing steps within one radar sequence (containing $N_{fr} = 15$ frames) to track motion. First, the hand is roughly localized following the procedure described in [41], yielding a first estimate of the depth of the hand surface z_e . For each frame within the sequence, two single-tone images are reconstructed at z_e , and the hand surface is localized as in [41]. The first frame of each sequence is used to obtain the depth z_{hyp} for each pixel, and N_{fr} two-tone images are reconstructed. Subsequently, the phase difference between corresponding pixels is evaluated for all frames per sequence. This phase difference is corrected and used to calculate a z -directed velocity v_z for each pixel. Finally, z_e can be adjusted for the next sequence based on the velocity evaluation.

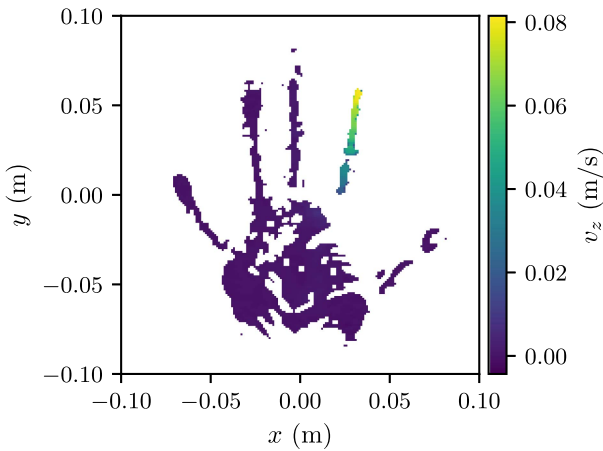


FIGURE 9. Locally distributed velocity during finger tapping movement.

velocity value from the palm toward the index fingertip shows the high lateral resolution of the velocity evaluation. In the next step, the estimated velocity can be used to adjust z_e . In the case of the specific measurement, z_e remained the same after the first coarse localization, since the hand itself was not moving. The process of hand surface depth map generation and pixel-specific velocity estimation is then repeated. Thus, for each radar sequence, 15 2D depth maps depicting the hand surface coordinates and a pixel-specific estimation of the z -directed velocity are obtained.

C. MOTION TRACKING

To evaluate the suitability of the proposed radar-based imaging approach for tracking motion, we compared the measurements of the position and speed of the tip of each moving finger with those obtained by the optical reference system.

1) α -FILTERING

To extract the pixel position that contained the fingertip of the moving finger from the radar data, the pixel exhibiting the highest absolute occurring velocity was automatically used for the comparison of the fingertip positions per frame and z -velocity per sequence. However, as the frequency difference $\Delta f = 200$ MHz between the two CW signals at f_1 and f_2 was used for hand surface localization rather than using the carrier frequency itself, the phase sensitivity regarding range measurements was relatively low, affecting the measurement accuracy [46]. Moreover, the evaluation of the fingertip position was based on a single radar frame. Therefore, the extraction of the fingertip position was more influenced by noise than the velocity estimation. For this reason, we used an α -filter [47] with $\alpha = 0.005$ to smoothen the position estimates. We took the first depth value from the first evaluated radar frame. For each radar sequence, we used the estimated velocity of the fingertip to predict the consecutive z -positions within a radar sequence to leverage all captured information and enable the evaluation of tracking accuracy for a greater number of samples. The α -filter then compared the predicted positions with the measured depth maps of the corresponding frames in the sequence, smoothing the z -position of the fingertip. For the next sequence, the next z -velocity was read, and the last z -position from the prior sequence served as a reference.

2) COMPARISON TO MARKER DATA

To compare the radar-based z -velocity and position estimation of the fingertip with the reference data, we extracted the marker position at the fingertip of the moving finger and transformed it into the radar coordinate system, which we determined by prior calibration measurements of markers that

TABLE 1 Accuracy Metrics of Radar-Based z -Position and z -Velocity Measurements of the Moving Fingertip

Measurement	Without Pixel-Specific Correction				With Pixel-Specific Correction			
	MAE		RMSE		MAE		RMSE	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
z -Velocity (mm/s)	9.0	4.4	12.3	5.8	8.1	3.9	11.0	5.2
z -Position (mm)	1.6	0.3	2.1	0.4	1.4	0.3	1.8	0.4

RMSE: Root mean square error; MAE: mean absolute error; SD: standard deviation. All values are rounded to one decimal point.

were attached to the radar's frontend surface. We then aligned the data from both sensors to a uniform time base. Because we obtained velocity data for each radar sequence but position data for each radar frame, the respective time base for the comparison between the radar and marker data differed in both cases. For 15 frames per sequence, we obtained an average sequence duration of 37.5 ms, corresponding to a sampling rate of 27 Hz. However, it should be noted that the frame duration and, hence, the sampling rate varied to some extent. We received a time stamp for the beginning of each radar frame and, therefore, for each radar sequence. The 4000 radar frames yielded a measurement duration of more than 10 s. For the optical reference system, the 3D position of each marker was provided at a constant frame rate of 200 Hz. We filtered these 3D positions using a 10 Hz low-pass Butterworth filter (4th order). We calculated the velocities in the x , y , and z directions using a central differences approximation, and disregarded the velocities in the first and last frames of the measurement. To compare the z -velocity of the fingertip extracted from the radar data with the velocity obtained by the reference system, we interpolated the marker data at the corresponding radar sequence time stamps between 0.005 and 9.995 s.

To determine the fingertip's position, we interpolated the radar data obtained for each frame and smoothed by the α -filter to a constant frame rate of 400 Hz to allow filtering by a 10 Hz low-pass filter according to the filtering of the marker data. As the optical system measured the dorsal side of the finger, whereas the radar system measured its distance to the palmar side, there was an inherent offset in the absolute positions measured by the two sensors. Moreover, the radar's frontend surface, which we used to align the radar data with the marker-based system's data, did not represent $z = 0$, which caused an additional offset in the measurement. We removed this offset by determining the difference between the mean of the local minima in the position data of the two sensors. To compare the radar- and marker-based position data, we then downsampled the radar signal to 200 Hz.

D. RESULTS

To evaluate the accuracy of the proposed procedure, we compared the z -directed velocities of the fingertip estimated by the radar and reference system for 262 samples in a time span from 0.042 s to 9.99 s after the measurement start. Moreover we compared the z -distance from 0.005 s to 10 s in 2000 samples. We calculated the root mean square error (RMSE) and the mean absolute error (MAE) for both parameters across

19 measurements (one measurement was excluded due to incorrect movement execution). We used the optical reference system for ground-truth comparison. Table 1 shows the RMSE and MAE means and standard deviations of all 19 position and velocity measurements without and with the pixel-specific correction that we introduced in Section II-A. Without the pixel-specific correction of geometry distortions, we obtained a mean MAE of 9 mm/s and a mean RMSE of 12.3 mm/s for the fingertip's z -velocity. With the proposed pixel-specific correction, we were able to reduce these measures by approximately 10 % to a mean MAE of 8.1 mm/s and a mean RMSE of 11.0 mm/s, indicating very low error rates and demonstrating the effectiveness of the proposed evaluation. The most significant difference between the two evaluations occurs at the extreme points, where the velocity reaches its highest absolute values. Consequently, it should be noted that the discrepancy in accuracy between the evaluation with and without the pixel-specific correction will be even greater when measuring faster motions. The standard deviations of the mean MAE and RMSE of the z -velocity measurements with pixel-specific corrections were 3.9 mm/s and 5.2 mm/s, respectively, suggesting some error variability but overall consistency. For the measurement of the z -position, the mean MAE was 1.4 mm, and the mean RMSE was 1.8 mm. The standard deviations of both the RMSE and MAE were minimal, indicating stable and precise measurements after α -filtering. As the tracking accuracy of the fingertip's z -position depends on the accuracy of the velocity evaluation, the results without pixel-specific correction also show increased error values. Exemplary radar-based measurements of the fingertip's z -velocity, after applying pixel-specific geometric distortion corrections, and the z -position, after α -smoothing and offset removal, are compared with marker-based measurements in Figs. 10 and 11, respectively. Both graphs show a close correspondence between the radar and the reference measurements. Minor deviations can be seen at the extreme points where the direction of movement reversed. This can most likely be attributed to averaging effects within the radar measurement, as we used several radar frames to evaluate speed and to smooth the position data. It should also be noted that the low-pass filtering of the marker position data had the strongest effect at the movement reversal and that we calibrated the marker-based system with a residual error of 0.5 mm, which is why there was also some uncertainty in the reference measurement. Fig. 12 shows an exemplary comparison of the estimated depth of the fingertip before and after α -filtering, demonstrating how this signal processing can help

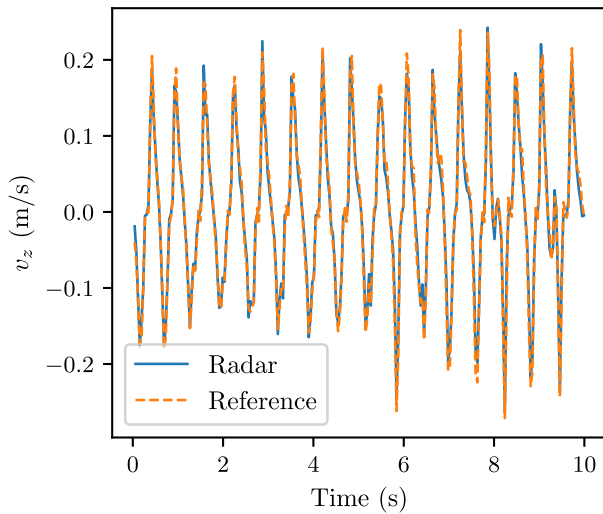


FIGURE 10. Comparison of exemplary measurements of the fingertip z -velocity v_z between the reference system and the radar after applying the pixel-specific geometry correction.

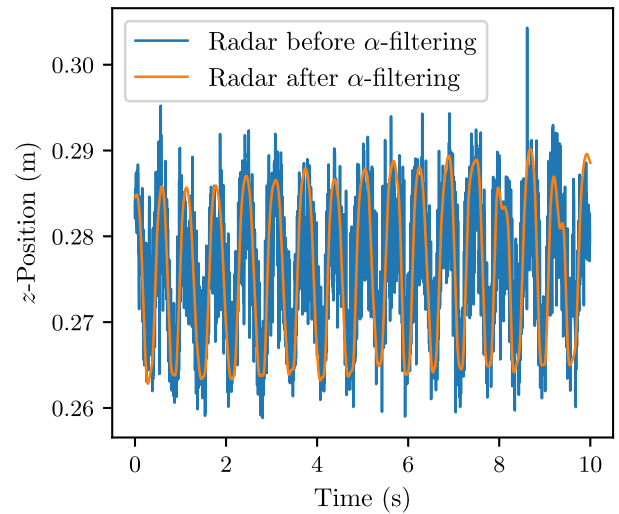


FIGURE 12. Comparison of radar-based fingertip z -coordinate estimation before and after smoothing with an α -filter.

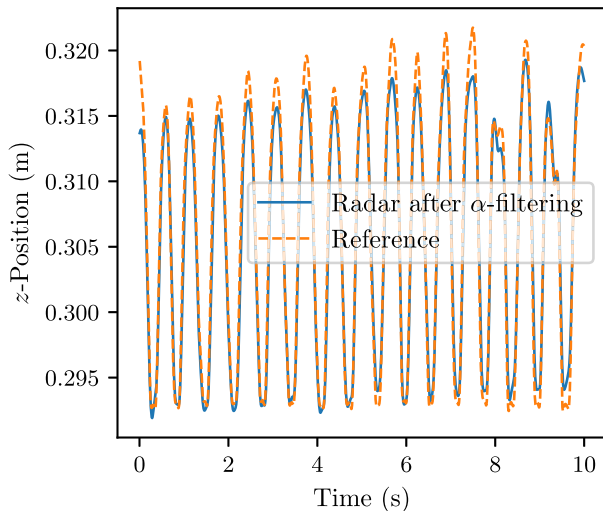


FIGURE 11. Comparison of an exemplary measurement of the moving fingertip's z -position between the radar and reference system. The radar data were smoothed using an α -filter and the offset in depth was corrected.

to reduce noise and smoothen the depth curve by leveraging the knowledge of the pixel-wise velocity estimation for each radar sequence. These results demonstrate the effectiveness of the proposed imaging radar system in accurately capturing the z -position and z -velocity of a moving human body shell.

IV. DISCUSSION

Despite the promising results presented in Section III-D, certain aspects of the proposed approach should also be discussed. First, it should be noted that human motion is inherently complex and multidimensional, encompassing movements across all three Cartesian coordinates (x , y , and z). Although the scope of this study was confined to motion along the depth dimension, we can gain useful insights into specific

types of motion in which the depth dimension is prominently involved, including gait analysis. When a person walks directly toward or away from a radar system, the predominant movement component is along the depth dimension. This is also true for many hand motions—for example, when the palm is facing the radar. The hand itself can be tracked, and finger motions that occur, for instance, while grasping, can be partly analyzed. Combined with machine or deep learning approaches, such evaluations can also be used in the context of gesture control. This work also forms the basis for measuring vectorial motions of the body shell using radar near-field imaging techniques. This can be achieved by integrating multiple MIMO radar systems positioned at different viewing angles. With such a configuration, motion components can be captured along all three spatial dimensions.

Second, a limitation of the imaging radar hardware used in this study is its frame rate, which hinders measurements with a high unambiguous velocity. This limitation is crucial when considering full-body measurements. However, as already noted, the current radar firmware is not optimized for high frame rates. Firmware optimization can enhance the frame rate even without hardware changes. Moreover, as autonomous driving research is increasingly oriented toward high-resolution 4D millimeter wave imaging [48], and given that high frame rates are required for measurements with a high unambiguous velocity, novel hardware architectures that incorporate high data rates and fast transmission schemes are being developed. Therefore, the low frame rate limitation of imaging radar systems is likely to be overcome in the future.

Third, in the context of full-body measurements, the influence of clothing needs to be investigated, as the proposed approach assumes only one reflection along the range dimension. However, assuming the strongest reflection occurs at the body surface, this method is still applicable to accurate velocity measurements of the body shell.

Fourth, if the entire post-processing workflow for tracking from Fig. 8 is extended to full-body motions, additional aspects, related to the 3D surface estimation using the approach from [41], must be considered, as this method has only been analyzed for hand contours so far. When reconstructing a single depth slice, the SNR could be reduced if a body part is located at a significant distance from the focal depth. This issue can be mitigated by iteratively estimating the surface contour, as proposed in [46]. However, two factors that require careful consideration are the width of the single-tone point spread function along the depth dimension and the unambiguous depth difference between the reconstructed and actual depth coordinates, which is determined by the frequency step Δf . The latter must be adapted to the requirements of a full-body measurement. The proposed velocity estimation then builds upon a depth map of the body surface, allowing the depth coordinates for the back projection of each pixel to vary, thereby maximizing the SNR.

Finally, we address the aspects of computation time and real-time capability, which are crucial not only for applications in human-computer interaction but also for scenarios where immediate feedback is essential, such as during medical examinations or dynamic adjustments in gait analysis. In [41], we achieved a significant reduction in the computational complexity of radar-based 3D reconstruction of the body shell by a factor of 1000. Currently, evaluating a single radar sequence (see Fig. 8), which involves reconstructing 15 radar images that correspond to a measurement duration of approximately 40 ms, takes about 20 s. Further acceleration can be achieved through fully optimized GPU reconstruction algorithms or by incorporating additional efficient reconstruction techniques, as discussed in [49]. This demonstrates the proposed method's substantial potential for future real-time evaluations.

V. CONCLUSION

In this paper, we introduced a novel concept for the pixel-wise velocity estimation of depth-directed human motion using near-field radar imaging. The proposed approach enables the tracking of the entire illuminated body shell, although we targeted the hand surface in this study. The proposed technique represents a substantial improvement over previous works, which have primarily focused on specific scattering centers of the body resolved in the Doppler or range dimension. The evaluation includes a pixel-specific correction that accounts for geometry distortions in the near field that would be ignored by far-field Doppler evaluations. We confirmed our theoretical derivations through suitable simulations. Moreover, we conducted a proof-of-concept measurement study comprising 19 measurements with a duration of 10 s each. During this study, the participants performed a finger tapping motion while the entire hand surface and its laterally resolved velocities were tracked by the imaging radar. A comparison of the radar-based fingertip velocity and position measurements with those obtained by a reference marker-based system confirmed the

feasibility of the approach. We achieved an outstanding accuracy in the pixel-wise velocity measurement with an MAE of 8.1 mm s^{-1} and an RMSE of 11.0 mm s^{-1} when applying the presented pixel-specific corrections. Furthermore, by using the hand surface localization method presented in [41], we were able to accurately track the motions of the hand. Incorporating the velocity information obtained through radar sequence evaluations allowed us to achieve very high tracking accuracy with an MAE of 1.4 mm and an RMSE of 1.8 mm. This novel method offers entirely new possibilities for contactless human motion assessment. The ability to track the surface of a body part with such high resolution is particularly beneficial for applications in biomechanics, medical diagnostics, rehabilitation, and human-computer interaction. Our approach provides a foundational framework that can be expanded to the estimation of full-body motion and other complex movements, significantly advancing the field of human motion analysis using radar technology.

ACKNOWLEDGMENT

The authors would like to thank the Rohde & Schwarz GmbH & Co. KG (Munich, Germany) for providing the radar imaging devices and technical support that made this study possible.

REFERENCES

- [1] V. C. Chen, "Detection and analysis of human motion by radar," in *Proc. IEEE Radar Conf.*, Rome, Italy, 2008, pp. 1–4.
- [2] S. Z. Gurbuz and M. G. Amin, "Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 16–28, Jun. 2019.
- [3] X. Li, Y. He, and X. Jing, "A survey of deep learning-based human activity recognition in radar," *Remote Sens.*, vol. 11, no. 9, May 2019, Art. no. 1068.
- [4] I. Ullmann, R. G. Guendel, N. C. Kruse, F. Fioranelli, and A. Yarovsky, "A survey on radar-based continuous human activity recognition," *IEEE J. Microwaves*, vol. 3, no. 3, pp. 938–950, Apr. 2023.
- [5] S. Ahmed, K. D. Kallu, S. Ahmed, and S. H. Cho, "Hand gestures recognition using radar sensors for human-computer-interaction: A review," *Remote Sens.*, vol. 13, no. 3, Feb. 2021, Art. no. 527.
- [6] F. Quaiyum, N. Tran, J. E. Piou, O. Kilic, and A. E. Fathy, "Noncontact human gait analysis and limb joint tracking using Doppler radar," *IEEE J. Electromagn. RF Microw. Med. Biol.*, vol. 3, no. 1, pp. 61–70, Mar. 2019.
- [7] J. Lien et al., "Soli," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–19, Jul. 2016.
- [8] E. van der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *Eur. J. Sport Sci.*, vol. 18, no. 6, pp. 806–819, May 2018.
- [9] M. Topley and J. G. Richards, "A comparison of currently available optoelectronic motion capture systems," *J. Biomech.*, vol. 106, 2020, Art. no. 109820.
- [10] J. Geisheimer, W. Marshall, and E. Grenaker, "A continuous-wave (CW) radar for gait analysis," in *Conf. Rec. 35th Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, 2001, pp. 834–838.
- [11] P. van Dorp and F. Groen, "Human walking estimation with radar," in *Proc. Inst. Elect. Eng., - Radar, Sonar, Navigation*, vol. 150, pp. 356–365, Nov. 2003.
- [12] P. E. Cuddihy et al., "Radar walking speed measurements of seniors in their apartments: Technology for fall prevention," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 260–263.

- [13] F. Wang, M. Skubic, M. Rantz, and P. E. Cuddihy, "Quantitative gait measurement with pulse-Doppler radar for passive in-home gait assessment," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2434–2443, Apr. 2014.
- [14] T. Yardibi et al., "Gait characterization via pulse-Doppler radar," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, Seattle, WA, USA, 2011, pp. 662–667.
- [15] P. van Dorp and F. Groen, "Feature-based human motion parameter estimation with radar," *IET Radar Sonar Navigation*, vol. 2, no. 2, pp. 135–145, Apr. 2008.
- [16] K. Saho, M. Fujimoto, Y. Kobayashi, and M. Matsumoto, "Experimental verification of micro-Doppler radar measurements of fall-risk-related gait differences for community-dwelling elderly adults," *Sensors*, vol. 22, no. 3, Jan. 2022, Art. no. 930.
- [17] K. Saho, K. Shioiri, S. Kudo, and M. Fujimoto, "Estimation of gait parameters from trunk movement measured by Doppler radar," *IEEE J. Electromagn. RF Microw. Med. Biol.*, vol. 6, no. 4, pp. 461–469, Dec. 2022.
- [18] D. Wang, J. Park, H.-J. Kim, K. Lee, and S. H. Cho, "Noncontact extraction of biomechanical parameters in gait analysis using a multi-input and multi-output radar sensor," *IEEE Access*, vol. 9, pp. 138496–138508, 2021.
- [19] P. Held, D. Steinhauser, A. Koch, T. Brandmeier, and U. T. Schwarz, "A novel approach for model-based pedestrian tracking using automotive radar," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 7082–7095, Jul. 2022.
- [20] M. A. Alanazi et al., "Towards a low-cost solution for gait analysis using millimeter wave sensor and machine learning," *Sensors*, vol. 22, no. 15, Jul. 2022, Art. no. 5470.
- [21] S. Hu, S. Cao, N. Toosizadeh, J. Barton, M. G. Hector, and M. J. Fain, "mmPose-FK: A forward kinematics approach to dynamic skeletal pose estimation using mmWave radars," *IEEE Sens. J.*, vol. 24, no. 5, pp. 6469–6481, Mar. 2024.
- [22] S. An and U. Y. Ogras, "Mars: MmWave-based assistive rehabilitation system for smart healthcare," *ACM Trans. Embed. Comput. Syst.*, vol. 20, no. 5s, pp. 1–22, Sep. 2021.
- [23] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3D tracking via body radio reflections," in *Proc. 11th USENIX Symp. Netw. Syst. Des. Implement.*, Seattle, WA, USA, Apr. 2014, pp. 317–329.
- [24] M. Arsalan and A. Santra, "Character recognition in air-writing based on network of radars for human-machine interface," *IEEE Sens. J.*, vol. 19, no. 19, pp. 8855–8864, Oct. 2019.
- [25] S. K. Leem, F. Khan, and S. H. Cho, "Detecting mid-air gestures for digit writing with radio sensors and a CNN," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1066–1081, Apr. 2020.
- [26] P. Wang et al., "A gesture air-writing tracking method that uses 24 GHz SIMO radar SoC," *IEEE Access*, vol. 8, pp. 152728–152741, 2020.
- [27] H.-S. Yeo, B.-G. Lee, and H. Lim, "Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware," *Multimedia Tools Appl.*, vol. 74, no. 8, pp. 2687–2715, Apr. 2015.
- [28] L. Guo, Z. Lu, and L. Yao, "Human-machine interaction sensing technology based on hand gesture recognition: A review," *IEEE Trans. Hum. Mach. Syst.*, vol. 51, no. 4, pp. 300–309, Aug. 2021.
- [29] M. Lapresa, V. Corradini, A. Iacca, F. D. S. Luzio, L. Zollo, and F. Cordella, "A comprehensive analysis of task-specific hand kinematic, muscle and force synergies," *Biocybern. Biomed. Eng.*, vol. 44, no. 1, pp. 218–230, Jan. 2024.
- [30] J. Ma'touq, T. Hu, and S. Haddadin, "Sub-millimetre accurate human hand kinematics: From surface to skeleton," *Comput. Methods Biomech. Biomed. Eng.*, vol. 21, no. 2, pp. 113–128, Jan. 2018.
- [31] Y. Li, P. Wang, R. Li, M. Tao, Z. Liu, and H. Qiao, "A survey of multifingered robotic manipulation: Biological results, structural evolutions, and learning methods," *Front. Neurobot.*, vol. 16, Apr. 2022, Art. no. 843267.
- [32] U. Phutane et al., "Evaluation of optical and radar based motion capturing technologies for characterizing hand movement in rheumatoid arthritis A pilot study," *Sensors*, vol. 21, no. 4, Feb. 2021, Art. no. 1208.
- [33] V. Gracia-Ibáñez, M.-J. Agost, V. Bayarri-Porcar, P. Granell, M. Vergara, and J. L. Sancho-Bru, "Hand kinematics in osteoarthritis patients while performing functional activities," *Disabil. Rehabil.*, vol. 45, no. 7, pp. 1124–1130, Apr. 2023.
- [34] J. Bek, E. Gowen, S. Vogt, T. J. Crawford, and E. Poliakoff, "Observation and imitation of object-directed hand movements in Parkinson's disease," *Sci. Rep.*, vol. 13, no. 1, Oct. 2023, Art. no. 18749.
- [35] P. R. Sánchez-Laulhé et al., "An exercise and educational and self-management program delivered with a smartphone app (carehand) in adults with rheumatoid arthritis of the hands: Randomized controlled trial," *JMIR Mhealth Uhealth*, vol. 10, no. 4, Apr. 2022, Art. no. e35462.
- [36] L. W. X. Cejnog, T. de Campos, V. M. C. Elui, and R. M. Cesar Jr., "A framework for automatic hand range of motion evaluation of rheumatoid arthritis patients," *Inform. Med. Unlocked*, vol. 23, Mar. 2021, Art. no. 100544.
- [37] M. Köprülüoğlu, İ. Naz, D. Solmaz, and S. Akar, "Hand functions and joint position sense in patients with psoriatic arthritis— A comparison with rheumatoid arthritis and healthy controls," *Clin. Biomech.*, vol. 95, May 2022, Art. no. 105640.
- [38] D. Sheen, D. McMakin, and T. Hall, "Near-field three-dimensional radar imaging techniques and applications," *Appl. Opt.*, vol. 49, no. 19, pp. E83–E93, Jul. 2010.
- [39] D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," *IEEE Trans. Microw. Theory Techn.*, vol. 49, no. 9, pp. 1581–1592, Sep. 2001.
- [40] S. S. Ahmed, "Microwave imaging in security – Two decades of innovation," *IEEE J. Microwaves*, vol. 1, no. 1, pp. 191–201, Jan. 2021.
- [41] J. Bräunig et al., "An ultra-efficient approach for high-resolution MIMO radar imaging of human hand poses," *IEEE Trans. Radar Syst.*, vol. 1, pp. 468–480, 2023.
- [42] Rohde & Schwarz, "Rohde & Schwarz QAR50 quality automotive radome tester," 2022. [Online]. Available: <https://www.rohde-schwarz.com/qar50/>
- [43] A.-K. Seifert, M. Grimmer, and A. M. Zoubir, "Doppler radar for the extraction of biomechanical parameters in gait analysis," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 2, pp. 547–558, Feb. 2021.
- [44] Qualisys, "Qualisys cameras 5+, 6+, and 7+," Bata sheet for Qualisys cameras, 2024. [Online]. Available: <https://www.qualisys.com/cameras/5-6-7/>
- [45] J. Geiss, E. Sippel, and M. Vossiek, "A practical concept for precise calibration of MIMO radar systems," in *Proc. 18th Eur. Radar Conf.* London, U.K., 2022, pp. 405–408.
- [46] J. Bräunig, V. Wirth, M. Stamminger, I. Ullmann, and M. Vossiek, "An efficient yet high-performance method for precise radar-based imaging of human hand poses," 2024, *arXiv:2406.13464*.
- [47] P. Kalata, "The tracking index: A generalized parameter for – and – target trackers," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-20, no. 2, pp. 174–182, Mar. 1984.
- [48] L. Fan, J. Wang, Y. Chang, Y. Li, Y. Wang, and D. Cao, "4D mmWave radar for autonomous driving perception: A comprehensive survey," *IEEE Trans. Intell. Veh.*, vol. 9, no. 4, pp. 4606–4620, Apr. 2024.
- [49] J. Moll, P. Schops, and V. Krozer, "Towards three-dimensional millimeter-wave radar with the bistatic fast-factorized back-projection algorithm-potential and limitations," *IEEE Trans. THz Sci. Technol.*, vol. 2, no. 4, pp. 432–440, Jul. 2012.



JOHANNA BRÄUNIG (Student Member, IEEE) was born in Bamberg, Germany, in 1993. She received the M.Sc. degree in electrical engineering from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2020. Since 2020, she has been working toward the Ph.D. degree with the Institute of Microwaves and Photonics, FAU. Her research interests include radar imaging and signal processing. She is particularly interested in radar-based motion tracking and the use of radar technology in medical engineering and machine learning applications. In 2023, she received the Young Engineer Prize at the European Radar Conference.



SIMON HEINRICH was born in Tirschenreuth, Germany, in 1996. He received the M.Sc. degree in mechanical engineering from the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2021. Since 2021, he has been working toward the Ph.D. degree with the Institute of Applied Dynamics, FAU. His research focuses on optimal data inclusion in optimal control simulations of biomechanical systems. He is particularly interested in the optimal control simulation of hand motions to obtain objective and reliable biomarkers for rheumatoid arthritis.



BIRTE COPPERS was born in Pforzheim, Germany in 1995. She received the B.A. degree in sports science from the University of Konstanz, Konstanz, Germany, and the M.Sc. degree in sports science (specializing in movement and technology) from the Karlsruher Institut of Technology, Karlsruhe, Germany, in 2021, and conducted her master's thesis with the Laboratory of Functional Biomechanics, University Hospital, Basel, Switzerland. Since 2021, she has been working toward the Ph.D. degree with the Department of

Internal Medicine 3, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, and University Hospital Erlangen, Erlangen. She was a sports therapist with the Neurological Rehabilitation Clinic, Zihlschlacht, Switzerland. Her research interests include applied biomechanics and motion analysis for clinical applications.



CHRISTOPH KAMMEL (Student Member, IEEE) was born in Nuremberg, Germany, in 1996. He received the M.Sc. degree in electrical engineering from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2020. He is currently working toward the Ph.D. degree with the Institute of Microwaves and Photonics, FAU. He joined the Institute of Microwaves and Photonics at FAU. His current research focuses on radar imaging and signal processing in the space sector and radar-based motion tracking in medical applications.

He was the recipient of the Semikron Promotion Prize in 2016, the Rohde & Schwarz Prize in 2018, the Baumüller Master Prize in 2021, and a Student Paper Award at the IEEE Radar Conference in 2023.



VANESSA WIRTH (Student Member, IEEE) was born in Schwabach, Germany, in 1996. She received the M.Sc. degree in computer science from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in collaboration with the Technical University of Munich, Munich, Germany. She is currently working toward the Ph.D. degree with the Chair of Visual Computing, FAU. Her research interests include the intersection of computer graphics, computer vision, and machine learning. She is particularly

interested in surface reconstruction and motion tracking of nonrigid objects using radar-based and optical sensor systems.



MARC STAMMINGER is currently the Head of the Chair of Visual Computing, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. He has authored or coauthored about 200 papers in leading journals and conference proceedings. His research interests include visual computing, such as real-time rendering, visualization, 3D reconstruction, object tracking, video manipulation, and virtual and augmented reality. He has been elected to the Eurographics Executive Committee and currently is the Chair of

the Eurographics Professional Board.



SIGRID LEYENDECKER received the Diploma in applied mathematics and the Ph.D. degree in mechanics from Universität Kaiserslautern, Kaiserslautern, Germany, in 2002 and 2006, respectively. She has been a Professor of mechanics/dynamics with Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, since 2011. Her research interests include computational mechanics, with a particular focus on dynamics and optimal control, addressing contemporary questions in life sciences and engineering.



ANNA-MARIA LIPHARDT received the Diploma and Ph.D. degree in sports science from Deutsche Sporthochschule Köln, Cologne, Germany, in 2003 and 2008, respectively. She is a member of the Medical Faculty, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, and leads an FAU-Independent Research Group, Department of Internal Medicine 3, Universitätsklinikum Erlangen, Erlangen. Her research focuses on identifying functional and biochemical biomarkers that indicate changes in musculoskeletal homeostasis in

response to disease and immobilization.



INGRID ULLMANN (Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree from Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Erlangen, Germany, in 2016 and 2021, respectively. She is currently a Postdoctoral Student and Head of the Research Group Wave-Based Imaging Systems, Institute of Microwaves and Photonics, FAU. In 2022, she spent one month as a visiting researcher with the Microwave Sensing, Signals and Systems Group, Delft University of Technology, Delft,

The Netherlands. Her research interests include radar imaging and radar signal processing for nondestructive testing, security screening, medical radar, and automotive applications. She is a reviewer for the European Radar Conference (EuRAD) and various journals in the field of microwaves and is a member of the IEEE CRFID Technical Committee on Motion Capture & Localization. Since 2022, she has been an Associate Editor for IEEE TRANSACTIONS ON RADAR SYSTEMS. She was the recipient of the Argus Science Award (sponsored by Airbus Defense and Space, now Hensoldt) in 2016 and the EuRAD Prize in 2019.



MARTIN VOSSIEK (Fellow, IEEE) received the Ph.D. degree from Ruhr-Universität Bochum, Bochum, Germany, in 1996. In 1996, he joined Siemens Corporate Technology, Munich, Germany, where he was the Head of the Microwave Systems Group from 2000 to 2003. Since 2003, he has been a Full Professor with Clausthal University, Clausthal-Zellerfeld, Germany. Since 2011, he has been the Chair of the Institute of Microwaves and Photonics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. He has authored or coauthored more than 400 publications. His research has led to more than 100 granted patents. His current research interests include radar, microwave systems, wave-based imaging, transponders, RF identification and communication, and wireless sensor and locating systems. Dr. Vossiek is the Spokesman of the Collaborative Research Centre (CRC 1483) EmpkinS, where more than 80 researchers aim to develop innovative wireless and wave-based sensor technologies for medicine and psychology. He is a member of the German National Academy of Science and Engineering and the Review Board of the German Research Foundation (DFG). He is a member of the IEEE Microwave Theory and Technology (MTT) Technical Committees for MTT-24 Microwave/mm-wave Radar, Sensing, and Array Systems; MTT-27 Connected and Autonomous Systems (as Founding Chair); and MTT-29 Microwave Aerospace Systems. He also serves on the Advisory Board of the IEEE CRFID Technical Committee on Motion Capture & Localization. He was the recipient of numerous best paper prizes and other awards. In 2019, he was awarded the Microwave Application Award by the IEEE MTT Society (MTT-S) for Pioneering Research in Wireless Local Positioning Systems. Dr. Vossiek has been a member of organizing committees and technical program committees for many international conferences and has served on the review boards of numerous technical journals. From 2013 to 2019, he was an Associate Editor for IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES. Since October 2022, he has been an Associate Editor-in-Chief for IEEE TRANSACTIONS ON RADAR SYSTEMS.