PASSIVE DETECTION AND DISCRIMINATION OF BODY MOVEMENTS IN THE SUB-THZ BAND: A CASE STUDY

Sanaz Kianoush, Stefano Savazzi, Vittorio Rampa *

National Research Council of Italy (C.N.R.)
Institute of Electronics, Computers and Telecommunication Engineering (I.E.I.I.T.)

ABSTRACT

Passive radio sensing technique is a well established research topic where radio-frequency (RF) devices are used as real-time virtual probes that are able to detect the presence and the movement(s) of one or more (non instrumented) subjects. However, radio sensing methods usually employ frequencies in the unlicensed $2.4-5.0$ GHz bands where multipath effects strongly limit their accuracy, thus reducing their wide acceptance. On the contrary, sub-terahertz (sub-THz) radiation, due to its very short wavelength and reduced multipath effects, is well suited for high-resolution body occupancy detection and vision applications. In this paper, for the first time, we adopt radio devices emitting in the 100 GHz band to process an image of the environment for body motion discrimination inside a workspace area. Movement detection is based on the real-time analysis of body-induced signatures that are estimated from sub-THz measurements and then processed by specific neural network-based classifiers. Experimental trials are employed to validate the proposed methods and compare their performances with application to industrial safety monitoring.

***Index Terms*—** Terahertz communication, passive activity recognition, human-robot collaboration, machine learning, feed-forward networks, long short-term memory networks (LSTM).

1. INTRODUCTION

Terahertz (THz) band communications [1] lie in the frequency gap between 0.1 THz and 10 THz and it is envisioned to satisfy the increasing demand for high speed wireless communication in the very short range ($1-10$ m). Recent technological innovation in designing sensors, detectors and antennas in the THz band, enables available research results to be widely applied in more scenarios, especially automotive and industrial applications [2]. In particular, the sub-THz band, from 0.1 THz up to 1 THz, is expected as a promising option to be utilized for communication due to a tractable level of signal attenuation. In addition, a considerable amount of work has been done in various applications due to this frequency range [3], and special-purpose devices have been demonstrated for imaging [4] as well as sensing [5] and biology [6]. Compared to the microwave radiation in the $1-50$ GHz band, the sub-THz frequency range achieves a fairly good spatial resolution required for precise imaging [4]. The diffusion of sub-THz radios for next generation of personal communication and IoT devices [7] is expected to pave the way towards novel dense communication, sensing and computing paradigms, able to provide new functionalities by leveraging existing hardware devices, networks, and building infrastructures. Transforming deployed sub-THz-enabled systems into sensing infrastructures might become a promising opportunity to develop high precision human-scale vision applications. Radio vision systems [8] for the recognition of human body movements generally exploit measurements of the stray radiation emitted by unmodified radio-frequency (RF) devices to recover a rough 2D/3D image of the environment. Most of the existing systems leverage 2.4 GHz and 5.0 GHz ISM bands, and generally cm-scale wavelengths [8, 9]. However, RF sensing techniques in such bands have shown to have the limitations of rich signal multi-path [10], low accuracy [11], complex signal processing methods [12] and necessity to maintain a controlled environment [13]. The adoption of sub-THz bands, and, generally, mm-scale wavelengths (mmW), is expected to provide reduced multi-path effects. This kind of scenarios, characterized by a more tractable measurement setup and a simpler analysis of reflections, in turns increases the accuracy of tracking and activity monitoring at the cost of a reduced coverage area. Finally, diffusion of consumer-grade devices for 5G and automotive applications will provide the wide adoption of passive RF techniques for the sub-THz and mmW bands.

2. LITERATURE REVIEW

Several models have been proposed to describe body-induced multipath radio propagation [14] up to 100 GHz band, while body effects in the sub-THz band are still rather unexplored with only a few available works [15]. The development of sources and detectors for the high-frequency range has been driven by applications such as spectroscopy, imaging and impulse ranging [16]. However, with sensing applications and

*Funded by the CyberSort project CNR - Regione Lombardia 2016-2018 and ERANET-COFUND (H2020) CHIST-ERA III, RadioSense.

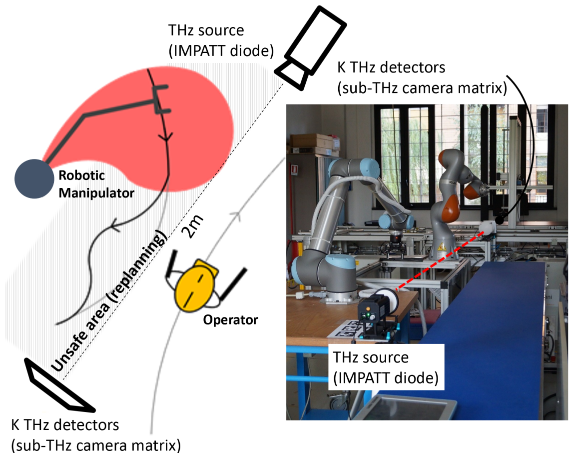

Fig. 1. HRC setup in the experimental test plant: emitter and receiver are 2 m apart; the operator is inside the shared area implementing some activities.

short-range communications possible in the $1-10$ m range, the technology and applications for sub-THz waves are growing rapidly. In [17] authors investigated the mmW band, specially 60 GHz for localization task. The results demonstrate a strong potential for mmW as a localization service with decimeter level accuracy and high availability. In [18], authors proposed the use of 60 GHz radios for human mobility tracking and activity monitoring. The properties of the sub-THz radiation make people screening solutions extremely valuable for many industrial applications where human safety is critical. In line with the next generation of smart manufacturing systems (i.e., Industry 4.0), in this paper, we exploit the sub-THz radio technology to implement a virtual safety fence with the goal of isolating a human worker from robots that are cooperating in implementing HRC (Human-Robot Collaboration) tasks [19]. In particular, we explore the problem of segmentation, processing and classification of signal sequences measured from the stray sub-THz radiations emitted for the detection and discrimination of body movements inside an indoor industrial HRC test plant. As depicted in Fig. 1, the sub-THz electromagnetic (EM) radiation is collected using a sub-THz camera consisting in a 2D array of K detectors. We address here the sensitivity of the system under different environmental conditions, and the accuracy for the discrimination of safe vs. unsafe body movements inside a HRC shared space [19]. Discrimination of human activities is based on machine learning tools, while a comparative analysis of different techniques for performance assessment is provided and results are discussed.

3. PROBLEM STATEMENT

This section focuses on the statistical model adopted for the selection of the features obtained from stray sub-THz radiation measurements that can be processed for body detection and activity discrimination. This model is verified by experiments in Sect. 4. The proposed virtual safety fence system consists of a sub-THz source and a sub-THz receiver that is composed by a 2D array of K single detectors, namely a sub-THz camera available in the laboratory. Let $i_{t,k}$ be a measure of the intensity, or strength, of the radiated sub-THz field observed, at time t, by the k-th detector (with $k \in [1, ..., K]$) that composes the camera. The intensity is proportional to the electric field power of the received electromagnetic waves normalized to the $[0-1]$ value range. The sampling interval $\triangle t$ is chosen according to some criteria detailed in Sect. 4. Vector $\mathbf{I}_t := \{i_{t,k}\}_{k=1}^{K}$ collects the measurements of the sub-THz radiation for all K detectors expressed in dB-scale. The virtual safety fence is programmed to detect up to $j = 1, ..., M$ worker activities in the surroundings of the line-of-sight (LOS) path connecting the sub-THz source and camera, since body activities leave a characteristic footprint on the received radiation. A safety controller continuously monitors the collaborative task and use the output of the detection system to replan, stop or resume the robot activity, depending on the specific actions taken by the human worker. In general, the multi-ray propagation between the fixed source and receiver consists of line-of-sight (LOS), reflected, scattered and diffracted rays [20]. When no obstacles are present near the LOS path and the sub-THz setup is directional (e.g., using dielectric lenses as in the experimental setup of Sect. 4), the propagation is mostly due to the main LOS ray that depends only on the free-space loss and the absorption loss. Therefore, the radiation intensity is modeled as

$$\mathbf{I}_t = \mathbf{s}_t(j) + \mathbf{w}_t, \tag{1}$$

where $\mathbf{s}_t(j) := \{s_{t,k}(j)\}_{k=1}^{K}$ corresponds to the body-induced signature corresponding to the j-th activity while the Gaussian noise $\mathbf{w}_t \sim \mathcal{N}(\mathbf{b},\mathbf{C})$ models the background radiation (i.e., observed in the *empty* environment, namely with the worker outside the detection area) including the measurement noise due to the sub-THz camera. Considering that environmental changes (i.e., due to concurrently robot movements) might alter the LOS propagation of the sub-THz signal, the average component $\mathbf{b} := \{b_k\}_{k=1}^{K}$ is a function of LOS terms: absorption and free-space loss. The absorption loss accounts for the attenuation that a propagating wave suffers because of molecular absorption and depends on the concentration of molecules encountered along the path [20]. This is considered a constant at the given center frequency (here equal to 100 GHz) and constant temperature. The free-space loss accounts for the attenuation due to the wave propagation through the medium. Reflection, scattering and diffraction effects due to a worker or an obstacle (e.g., robot) near the LOS path (i.e., approaching/crossing the virtual fence) introduce time-varying changes in (1). The covariance $\mathbf{C}$ includes such environmental changes in addition to those due to the measurement noise sources.

3.1. Change detection and segmentation

Discrimination of body activities is based on the analysis of the sequences $\mathbf{I}_{t_A,T} = [\mathbf{I}_{t_A}, \mathbf{I}_{t_A+1} ..., \mathbf{I}_{t_A+T}]$ of T samples. We define at first a change detection algorithm to segment anomalous fragments of data in real-time. It is used to detect the optimal change point, or beginning t_A of the segment, considering all K detectors. Then, in Sect. 3.2, we analyze some features that can be extracted from sequences in order to apply machine learning tools for classification of body movements. Considering a typical time-varying indoor environment, we model the fluctuations of the detected radiation $i_{t,k}$ by a linear first order auto-regressive (AR-1) model

$$i_{t,k}(\boldsymbol{\theta}_k) = a_k \times i_{t-1,k} + r_k + n_{t,k}, \tag{2}$$

with parameters $\boldsymbol{\theta}_k = [a_k, \sigma_k, r_k]$, namely the AR parameter a_k, the static component $r_k = (1 - a_k)b_k$ and the residual $n_{t,k} \sim \mathcal{N}(\,0, \sigma_k^2)$ with deviation $\sigma_k^2 = \mathrm{E}\left[(i_{t,k} - b_k)^2\right] \times \left(1 - a_k^2\right)$. The detection problem is implemented iteratively by exploiting the cumulative sum (CUSUM) paradigm [21]. Assuming an inspection interval T (see also Sect. 4), the CUSUM g_t is designed to identify the intensity fluctuations

$$g_t = \frac{1}{T} \sum_{k=1}^{K} \frac{1}{\sigma_k} \sum_{j=t}^{t+T} \left(\frac{n_{j,k}^2}{\sigma_k^2} - 1 \right)^+ \tag{3}$$

with $(\cdot)^+ = \max[0, \cdot]$ and $n_{t,k} = i_{t,k} - a_k \times i_{t-1,k} - r_k$. According to the detection threshold h_0 (obtained during the training stage), the beginning t_A of the anomalous sequence (i.e., the change point) is estimated in real-time as

$$t_A = \min\{t\!:\; g_t \geq h_0\}. \tag{4}$$

3.2. Feature extraction and machine learning tools

Detection of worker activities is herein based on the real-time analysis of body-induced signatures $\mathbf{s}_t$ that are estimated from measurements $\mathbf{I}_t$ through background subtraction as

$$\hat{\mathbf{s}}_t = \mathbf{C}^{-\frac{1}{2}}(\mathbf{I}_t - \mathbf{b}). \tag{5}$$

Discrimination of activity (to classify its safety relevance) is based on the real-time analysis of the sequence

$$\mathbf{S}_{t_A,T} = [\hat{\mathbf{s}}_{t_A}, \hat{\mathbf{s}}_{t_A+1} ..., \hat{\mathbf{s}}_{t_A+T}] \tag{6}$$

over the segmented interval $[t_A,\;\; t_A + T]$. For comparative evaluation, we examined the Maximum Likelihood Estimation (MLE) [22] approach against two artificial neural networks (ANN), namely the Feed-Forward neural network (FF-NN) and the Long Short-Term Memory (LSTM) network [23]. MLE takes as features the mean and standard deviation of signatures $\mathbf{S}_{t_A,T}$ obtained during the training phase. Based on (5), signature $\hat{\mathbf{s}}_t$ observed at time t is assumed to be conditionally independent given the j-th activity near the virtual fence. The measurement $\hat{\mathbf{s}}_t = \hat{\mathbf{s}}_t(j)$ conditioned to activity j is thus assumed to be an uncorrelated Gaussian vector with mean $\mathbf{h}_S(j) = [h_1(j), ..., h_K(j)]^T$ and covariance $\mathbf{C}_S(j) = \mathrm{diag}\left(\sigma_1^2(j), ..., \sigma_K^2(j)\right)$. Conditional likelihood is

$$p(\hat{\mathbf{s}}_t|j) = \frac{1}{\sqrt{(2\pi)^K\,|\mathbf{C}_S(j)|}} \exp\left\{-\frac{1}{2}\,\|\hat{\mathbf{s}}_t - \mathbf{h}_S(j)\|^2_{\mathbf{C}_S^{-1}(j)}\right\} \tag{7}$$

where $\|\mathbf{s}\|^2_{\mathbf{C}} = \mathbf{s}^T\mathbf{C}\;\mathbf{s}$ denotes the square norm of the vector $\mathbf{s}$ weighted by the matrix $\mathbf{C}$. Snapshot MLE of activity $\hat{j}$, based on the segmented sequence $\mathbf{S}_{t_A,T}$, can be obtained by maximizing $\hat{j}_t = \mathrm{argmax}_{j\in[1,...,M]}p(\hat{\mathbf{s}}_t|j)$ while the final decision about safe/unsafe activity within the considered interval T is based on major voting. Both neural network models take the segmented sequences (6) of length T as inputs for classification. FF input layer is followed by a single hidden layer with 10 neurons, a softmax layer and an output classification layer with $M = 3$ classes (see Sect. 4 for details), for discriminating 3 different activities. LSTM architecture is specifically designed to model temporal sequences and, unlike hidden Markov models, includes long-range dependencies. LSTM architectures have been explored for large-scale acoustic modeling in speech recognition, language translation, and handwriting recognition [24]. The methodology is here used to track long- and short-term dependencies on the segmented sequences. For both FF-NN and LSTM, 80% of data is used for training the network, 10% for validation and 10% for testing the classifiers. Results and comparative analysis are summarized in Sect. 4.

4. PRELIMINARY VALIDATION

For the implementation of the safety fence, we adopted a sub-THz source (about 100 mW at 100 GHz) that uses an IMPATT diode (impact ionization avalanche transit-time diode) technology [25] and a sub-THz camera consisting of $K = 1024$ pixel detectors arranged over a 2D array (32×32). The camera, developed by Terasense, has a room-temperature responsivity up to 50 kV/W (with pixel-to-pixel deviation responsivity within a 20% range) [26], a noise-equivalent power of 1 nW/$\sqrt{Hz}$ in the frequency range $0.01 - 1$ THz and an acquisition rate up to 50 fps (frames per second). As shown in Fig. 1, the experimental activity has been conducted inside a test plant over an indoor area of about 6×4 square meters size. The sub-THz source and camera are placed 2 m apart. A worker stands inside the monitored area and performs non rigid body motions (i.e., working activities) in close proximity with a robotic arm. In what follows, the MLE, FF-NN and LSTM approaches are compared. The virtual fence system is designed to discriminate between 3 activities, namely: *a*) a worker moving towards the unsafe space occupied by robotic arms, and crossing the virtual fence ($j = 1$); *b*) a worker traversing the virtual fence with both arms thus making un-

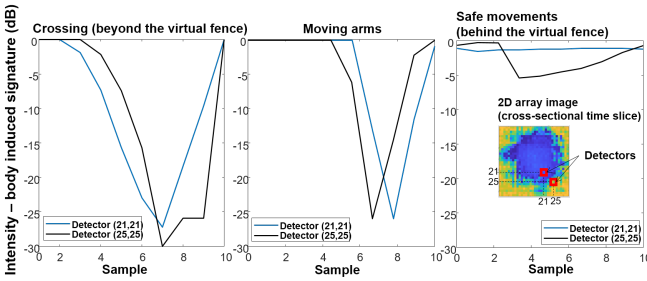

Fig. 2. Body-induced time-domain signatures after background removal vs. time for a worker (from the left): a) crossing the virtual fence (unsafe); b) moving both arms (unsafe); c) performing safe movements behind the virtual fence.

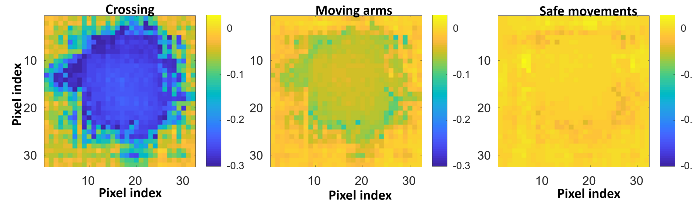

Fig. 3. Intensity images obtained after background removal for the 3 monitored activities involving a worker (from the left): a) crossing the virtual fence; b) moving both arms; c) performing safe movements behind the virtual fence.

safe movements while keeping the torso/body inside the the safe area ($j = 2$); c) a worker performing safe activities (random movements) behind the virtual fence (safe movements, $j = 3$). Fig. 2 shows some examples of the segmented time-domain signatures. They consist of $T = 10$ samples (200 ms) for all three cases; the change point detection adopts a threshold of $h_0 = 0.1$. In the same figure, we represent the time series considering 2 detectors from the 2D array, namely corresponding to the superimposed image pixels (21,21) and (25,25), respectively. Notice that detectors in the surrounding of pixel (16,16) mostly capture the LOS radiation field, while pixels located in the external sides of the image are more sensitive to obstructions in the surrounding of the LOS path (Obstructed LOS - OLOS). Both sub-THz source and camera mount PTFE (PolyTetraFluoroEthylene) lenses. According to this setup, the time-domain signatures after background removal abruptly change when the worker is entering the unsafe zone (i.e., about 30 dB of attenuation); a similar change, although characterized by a distinctive time-domain footprint and a smaller attenuation, is observed when the worker is performing unsafe arms movements. Safe movements behind the fence cause small attenuations (up to 5 dB) that are caused by OLOS propagation. These small alterations correspond to human torso approaching and partially obstructing the LOS path. Fig. 3 shows examples of cross-sectional time slices of the radiation field, after background subtraction, from all K detectors and for all activities. Fig. 4 shows performances of the classifier using MLE, FF and LSTM methods. While the MLE approach provides low accuracy for the discrimination of arms movements, FF and LSTM provide an accuracy higher than 95% for all activities. Fig. 4 also compares performance figures considering different pixel subsets of the 2D camera where only detectors corresponding to the zones outside the hatched areas are used for activity recognition. As shown in Fig. 4.c, removing the detector elements whose positions is in the center of the field of view and pixels at corners improves the accuracy and the computational time as well.

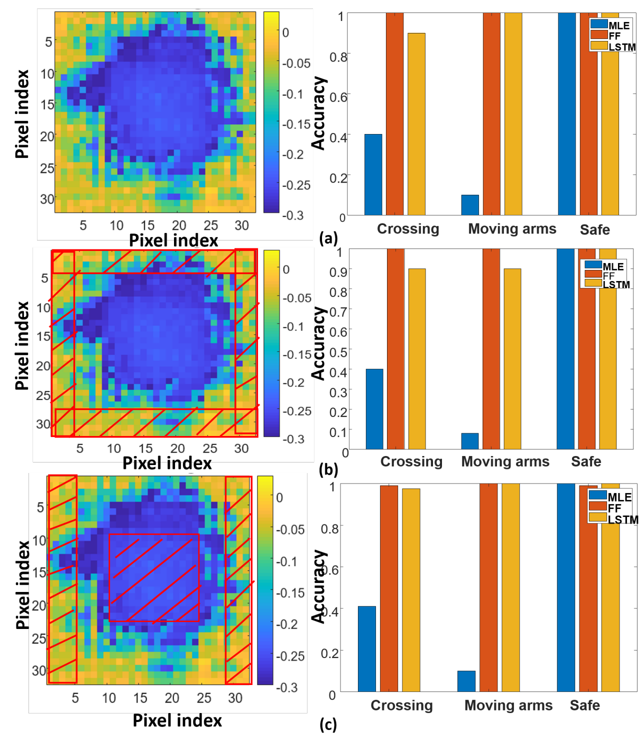

Fig. 4. Accuracy evaluation using MLE, FF-NN and LSTM networks (right) to discriminate the same activities detailed in Fig. 2 and 3. We considered (left, from the top): a) all pixels of the camera; b) only the pixels located in the central area; and c) only the pixels outside the highlighted areas.

5. CONCLUSIONS

In this paper, we propose a novel approach using sub-THz radiation for passive human sensing and particularly body movement discrimination. To reduce the complexity of training phase we also use sensing task in unsupervised mode using machine learning tools. Experimental validation is conducted using sub-THz sensor to discriminate different activities. Accuracy results show that the sub-THz technology and NN tools are well suited for human activity detection.

6. REFERENCES

[1] I. F. Akyildiz et al., "Terahertz band: Next frontier for wireless communications," *Physical Communication*, vol. 12, pp. 16 – 32, 2014.

[2] T. Nagatsuma, "Generating millimeter and terahertz waves," *IEEE Microwave Magazine*, vol. 10, no. 4, pp. 64–74, June 2009.

[3] M. Shur, "Subterahertz and terahertz sensing of biological objects and chemical agents," 2018.

[4] P.U. Jepsen et al., "Terahertz spectroscopy and imaging modern techniques and applications," *Laser and Photonics Reviews*, vol. 5, no. 1, pp. 124–166, 2010.

[5] E. Grossi et al., "Opportunistic radar in ieee 802.11ad networks," *IEEE Transactions on Signal Processing*, vol. 66, no. 9, pp. 2441–2454, May 2018.

[6] P. H. Siegel, "Terahertz technology in biology and medicine," *IEEE Transactions on Microwave Theory and Techniques*, vol. 52, no. 10, pp. 2438–2447, Oct. 2004.

[7] S.A. Busari et al., "5g millimeter-wave mobile broadband: Performance and challenges," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 137–143, 2018.

[8] S. Savazzi et al., "Device-free radio vision for assisted living: Leveraging wireless channel quality information for human sensing," *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 45–58, Mar. 2016.

[9] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 5, pp. 621–632, May 2010.

[10] V. Rampa et al., "Physical modeling and performance bounds for device-free localization systems," *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1864–1868, Nov 2015.

[11] J. Wilson et al., "See-through walls: Motion tracking using variance-based radio tomography networks," *IEEE Transactions on Mobile Computing*, vol. 10, no. 5, pp. 612–621, May 2011.

[12] S. Kianoush et al., "Leveraging mimo-ofdm radio signals for device-free occupancy inference: system design and experiments," *EURASIP Journal on Advances in Signal Processing*, vol. 2018, no. 1, pp. 44, Jul. 2018.

[13] F. Viani et al., "Passive real-time localization through wireless sensor networks," in *Geoscience and Remote Sensing Symposium,2009 IEEE International,IGARSS 2009*, July 2009, vol. 2, pp. II–718–II–721.

[14] N. Chahat et al., "On-body propagation at 60 ghz," *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 4, pp. 1876–1888, Apr. 2013.

[15] K. Yang et al., "Numerical analysis and characterization of thz propagation channel for body-centric nano-communications," *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 3, pp. 419–426, May 2015.

[16] M. Fitch et al., "Terahertz waves for communications and sensing," vol. 25, 10 2004.

[17] F. Lemic et al., "Localization as a feature of mmwave communication," in *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Sept. 2016, pp. 1033–1038.

[18] Y. Zeng et al., "Poster abstract: Human tracking and activity monitoring using 60 ghz mmwave," in *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, April 2016, pp. 1–2.

[19] S. Savazzi et al., "Device-free human sensing and localization in collaborative human robot workspaces: A case study," *IEEE Sensors Journal*, vol. 16, no. 5, pp. 1253–1264, March 2016.

[20] C. Han et al., "Propagation modeling for wireless communications in the terahertz band," *IEEE Communications Magazine*, vol. 56, no. 6, pp. 96–101, June 2018.

[21] J. Bilmes et al., "The graphical models toolkit: An open source software system for speech and time-series processing," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2002, pp. 3916–3919.

[22] S. Kianoush et al., "Device free rf human body fall detection and localization in industrial workplaces," *IEEE Internet of Things Journal*, vol. 4, no. 2, pp. 351–362, Apr. 2017.

[23] R. Zazo et al., "Age estimation in short speech utterances based on lstm recurrent neural networks," *IEEE Access*, vol. 6, pp. 22524–22530, 2018.

[24] D. Keysers et al., "Multi-language online handwriting recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016.

[25] M. Mukherjee et al., "Gan impatt diode: a photosensitive high power terahertz source," *Semiconductor Science and Technology*, vol. 22, no. 12, pp. 1258, 2007.

[26] V. Muravev et al., "Novel relativistic plasma excitations in a gated two-dimensional electron system," *Phys. Rev. Lett.*, vol. 114, pp. 106805, Mar. 2015.