On Indoor Human Sensing Using Commodity Radar

Mohammed Alloulah
Nokia Bell Labs
Cambridge, UK
mohammed.alloulah@nokia-bell-labs.com

Fahim Kawsar
Nokia Bell Labs
Cambridge, UK
fahim.kawsar@nokia-bell-labs.com

Anton Isopoussu
Nokia Bell Labs
Cambridge, UK
anton.isopoussu@nokia-bell-labs.com

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Copyright held by the owner/author(s). Publication rights licensed to ACM.
UbiComp/ISWC'18 Adjunct,, October 8–12, 2018, Singapore, Singapore
ACM 978-1-4503-5966-5/18/10.
https://doi.org/10.1145/3267305.3274180

Abstract
Radio frequency radar indoors is gaining traction owing to its promise for extended coverage and device-free operation. However, while the well-behaved radar sensing model affords clear advantages, the cluttered indoor environment presents numerous challenges for reliable human sensing. Classic radar techniques are hard to call upon since the kinematic and clutter behaviours in aerospace are vastly different from their indoor counterparts. We demonstrate the peculiarities of indoor radar using a commercial 2D array commodity device in the 6 to 8.5 GHz band. We then present a set of processing tools suited for indoor radar human sensing. We show that excessive indoor clutter and erratic human kinematics can be largely mitigated building on such processing tools without resorting to much low-level techniques unsupported by commercial commodity radars.

Author Keywords
Human Sensing, Indoor Radar

ACM Classification Keywords
10003120.10003138 [Human-centered computing]: Ubiquitous and mobile computing

Introduction
Radar sensing is undergoing a renaissance. Emerging mission-critical applications, spearheaded by the autonomous

vehicle initiative, are driving development across all aspects radar: from analogue and mixed-signal integrated circuit (IC) design [7], sophisticated target tracking schemes [8], to artificial intelligence (AI) aided inference and decision-making [10].

A recent beneficiary of this renaissance is the indoor environment. Radar has been repurposed to sense respiration [12], see through-wall a human skeletal figure [1], localise a small number of co-located people [2], and as far as detect emotion based on physiological vital signs [16].

Despite novelty, the reported indoor radar use cases achieve their sensing tasks by a combination of: (1) low-level radar configuration e.g. a form of frequency sweep coding in [2], and (2) carefully orchestrated setups e.g. directly facing the radar in [12, 16] for respectively accurate sleep stage monitoring and emotion recognition. As a result, it would seem that further open innovation in indoor radar human sensing is hindered by two barriers:

Radar accessibility. In the short-term, further innovation in indoor radar for human sensing requires accessible research platforms. However, many commodity indoor radar vendors employ propriety architectures and algorithms to which users have neither the insight nor the low-level configurability. One such vendor we trialled is Walabot [15]. Apart from the high-level parametrisation of the sensing arena, no provisions for modifying the signalling architectures and algorithms will invariably stifle innovations while users attempt application-specific customisations. Other commodity indoor radar vendors such as XeThru do provide open platforms for development [11]. However, such platforms demand a certain level of expert knowhow scattered across the radio stack which cannot be assumed on the part of new entrants into the indoor radar human sensing field.

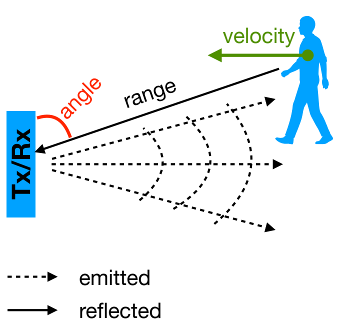


Figure 1: Radar sensing system.

Environmental robustness. Longer-term, it is desirable to transition early compelling radar results [16, 3] from the lab and into the wild. This will in turn fosters the further innovations needed for the technology to reach the level of maturity needed for impacting people's quality of life—tangibly and on an every day basis. For this to happen, it is crucial to begin to tackle problems arising from the notoriously cluttered indoor environment not only on a low-level—pertaining to often times inaccessible radar signalling architectures and algorithms—but also on a high-level for added resilience against residual environmental noise and dynamics.

In this paper, we set out to address the short-term radar accessibility challenge with the view to help foster long-term future research. To this end, we describe a sensing pipeline built on commodity indoor radars for human sensing in the wild. Our methodological stance is to help address environmental robustness without low-level radar modifications, which we believe will make our findings reproducible in a wider community of new entrants to the field of radar technology. Our early results show that this approach is able to tackle some of the complexities of using indoor radar for human sensing while allowing for avenues of further research.

Radar Primer

In a nutshell, a radar consists of an antenna array as shown if figure 1. The array scans the environment by using separate designated transmitter and receiver antennas. Alternatively, antennas belonging to the array may alternate between transmission and reception. During a scanning interval, transmitters emit radio energy omnidirectionally. Transmission energy can also be steered towards a spatial sector in the environment by making the transmitted pulses interfere constructively at a given angle, and destructively at others. Echoes reflect off inanimate obstacles or peo-

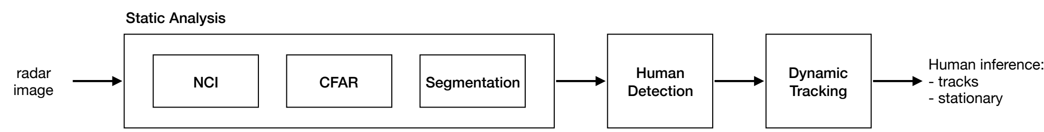


Figure 2: Human sensing using accessible radar.

ple back to the receiving array elements. The round-trip time-of-flight (TOF) allows for computing the distance to an object i.e. range. By making use of spatial sampling, the receiver array is also able to determine the angle at which reflected echoes arrive back in azimuth and elevation. Doppler shift can be used to enable the inference of an object's velocity with respect to that of the array. As such, well-behaved radar models make possible the simultaneous determination of range, velocity, and angle associated with a given moving object in the sensing environment. Many radar architectures with various pros and cons exist. Examples include ultra wideband (UWB), pulse-based signalling and frequency sweep-based approaches such as frequency modulation continuous wave (FMCW). Typically, commodity radars are shipped with API's that give end users access to a parametrisable[1] detection image corresponding to the radar's field-of-view (FOV) response to the unfolding sensing scene. Commodity radar vendors may also provide specialised API's targeted at added functionalities on top of basic target detection such as respiration or sleep monitoring.

In terms of radar behaviour indoors versus other more established media, there are a number of fundamental and noteworthy differences. For instance, the human body has a much lower radar cross section (RCS) characteristics[2] when compared to a metallic vehicle body in automotive radar [6]. Also, the kinematic behaviour of an automotive vehicle is constrained largely by physics, which in turn allows for simplifying assumptions such as constant velocity or constant acceleration motion models [4]. In contrast, it is unclear if such behaviour can be assumed on the part of moving people indoors. Many other open questions surrounding an indoor radar operation exist, which presents avenues for future research.

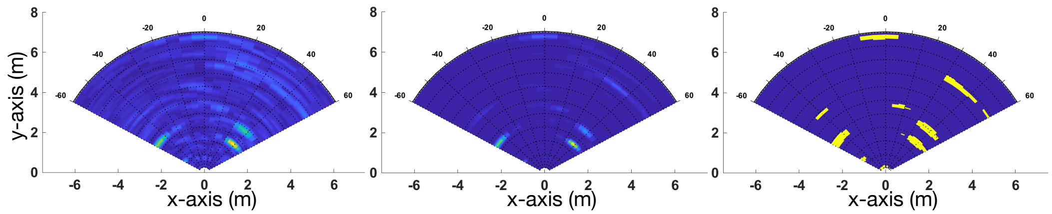
Human Sensing using Accessible Radar

We next present a sequence of processing steps suited for isolating and tracking humans from cluttered indoor images by commodity radar. A high-level block diagram of these steps in shown in figure 2. In what follows, we explain these processing stages and the rationale behind them.

(1) Static analysis. This refers to few subblocks which together make up the static analysis processing. When a person walks into a room and remains stationary in one location for a period of time, stronger reflections off indoor reflective radio surfaces such as walls and furniture will dominate the sensing scene. We therefore need to employ a series of transformations in order to isolate the human object of interest from inanimate targets and background clutter. To this end, we first apply non-coherent integration (NCI) to a number of elevation scan planes. The idea here is that depending on a limb RCS, such as torso and legs, the human body will be detected with various voxel intensities. For

[1] see table 1 for instance

[2] A radar cross section refers to the reflective characteristics of the material shone with radio energy.



(a) raw scene scanned by radar (b) NCI scene (c) segmented mask

Figure 3: Static person facing radar.

	res	limits
range	5cm	[0.25, 7]m
azimuth	5°	$\pm 60^{\circ}$
elevation	10°	$\pm 10^{\circ}$

Table 1: 3D scan parameters

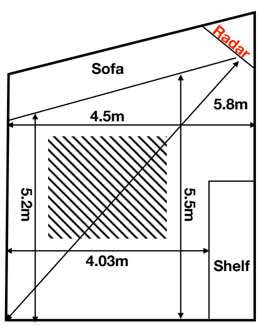


Figure 4: Test room layout

example, the radar unit may be positioned such that three elevation scans would target a person's head, torso, and legs. The middle elevation scan hitting the torso is likely to have the highest voxel intensity. Integrating elevation planes helps normalise for such behaviour to enhance sensitivity. Second, we apply a constant false alarm (CFAR) detector [14] across range bin scans in order to derive a detection mask to coarsely estimate background noise. An image segmentation algorithm [13] is then applied on the elevation integrated scene bootstrapped by the CFAR noise estimate. That is, the CFAR noise estimate adapts the image segmentation algorithm to changing scene dynamics. The overall output of the processing stage is a segmented image mask for direct human detection.

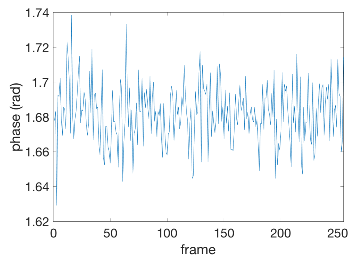
(2) Human detection in slow-time. The notion of slow-time in radar refers to conducting analysis in time across frames as opposed to within a single scene frame [5]. Slow-time processing enables target speed determination. In the context of indoor radar human sensing, slow-time target *phase* processing is what allows for vital sign estimation [12, 3].

Even after the application of proprietary low-level target detection algorithms, we make the observation that slow-time analysis across frames affords us the possibility of extracting humans targets from inanimate ones. Accordingly, the segmented mask outputted by previous static analysis stage is used to analyse the phase of targets in slow-time—possibly after further target centroid estimation [9]. It has been observed in prior art that human presence causes a target phase to change smoothly and periodically in slow-time compared to that of an inanimate object. Therefore, in this stage, static analysis is followed by a phase search across slow-time for all segmented targets. The output of this search is a labelling of targets as either static background-related or human objects of interest for further tracking.

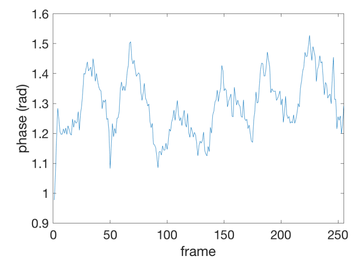
(3) Dynamic analysis. When humans move indoors, detection is somewhat simplified since we can employ differential scene techniques [5]. Additionally, similar image segmentation procedure can be followed. The targets can then be tracked with a variety of tools ranging from the simplistic, power-based to the sophisticated. Examples include probabilistic methods such as Kalman-based track initiation and maintenance variants [4] or neural networks variants [10].

Early Experimental Evaluation

We conduct experiments with a 2D commodity radar that consists of 24 antenna elements operating in the band 6



(a) inanimate object



(b) human

Figure 5: Slow-time phase for an inanimate object and a human.

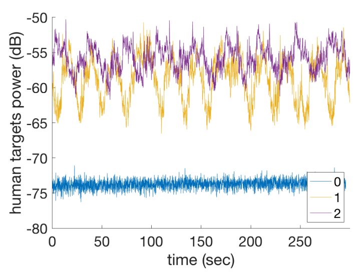


Figure 6: Aggregate human target power levels for moving people indoors across three occupancy cases.

to 8.5 GHz. The device is a customised version of the 18-element "Developer" device by Walabot [15]. Only high-level parametrisation is exposed to end developers in the form of crude arena configuration. By means of this arena configuration, we control resolutions and maximum sensing limits for: range, azimuth, and elevation. Indirectly, the 3D scanned volume also determines the frame rate since the radar need to process and stream all voxels to a host computer in a finite time. We configure the 2D radar with the 3D scan parameters shown in table 1. Under such scan settings, we obtain around 10 measurement frames per second.

Turning to human detection, phase analysis in slow-time is depicted in figure 5. It is readily observed that human presence causes a target phase to change smoothly and periodically in slow-time compared to that of an inanimate object.

We perform an experiment in which one human subject walks into a room and stands in the middle of it facing the radar unit. The designated standing zone is highlighted within the room layout of figure 4 as a diagonal pattern. The room contains many fixtures not indicated on figure 4 such as a lamp, plant, table, etc. 3D radar frames with the arena parameters enumerated in table 1 are recorded and fed to our proposed human sensing pipeline offline. As shown in figure 3a, when examining the raw detection image, the presence of a person in the middle is indiscernible.[3] The output of our pipeline's series of operations is depicted in figures 3b and 3c. The human target now appears in the segmented mask along with a number of false targets from office fixtures.

[3]Walabot does provide other API's for tunable detectors using their propriety pipelines should users prefer an out-of-the-box experience.

Finally we provide a simple example of dynamic tracking using a power-based technique for demonstration purposes. We conduct an experiment in the same room in which up to two people were asked to move continuously for 5 minutes. In figure 6, we track the aggregate contribution of these human targets for 1 person and 2 people moving, relative to the noise floor of an empty room. The periodicity of waveforms 1 and 2 are due to human subjects moving in circles within the room. When considering the entire test interval, the case of 2 people moving contains more dynamic target power than 1 person moving. It is worth re-emphasising that in the case of unmoving human subjects, slow-time analysis can isolate such stationary presence which will appear as a constant power offset over that of the noise floor—discounting false inanimate targets and as estimated by CFAR. Using these simple power metrics, classifiers based on static and dynamic analyses allow applications such as crowd counting and occupancy detection to be realised atop minimally invasive indoor radar and with limited or no site-specific calibration subject to application requirements.

Outlook

In this brief, we showcase how commodity radars can be used, with minimal insight into underlying technology, to build a processing pipeline suited for indoor human sensing. The outlined processing steps are accessible to new entrants to the field and are readily extensible and customisable for a variety of sensing tasks. We hope that our brief treatment serves to provide pointers to others wishing to conduct research in this newly reinvigorated body of literature. We expect indoor radar to play a major role in today's and future smart environments.

REFERENCES

1. Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. 2015a. Capturing the Human Figure Through a Wall. *ACM Trans. Graph.* 34, 6, Article 219 (Oct. 2015), 13 pages.
2. Fadel Adib, Zachary Kabelac, and Dina Katabi. 2015b. Multi-person Localization via RF Body Reflections. In *Proc. of the 12th USENIX Conf. on Networked Systems Design and Implementation (NSDI'15)*. USENIX Association, Berkeley, CA, USA, 279–292.
3. Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C. Miller. 2015c. Smart Homes That Monitor Breathing and Heart Rate. In *Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 837–846.
4. Yaakov Bar-Shalom, Peter K Willett, and Xin Tian. 2011. *Tracking and Data Fusion*. YBS publishing Storrs, CT, USA:.
5. Gregory L. Charvat. 2014. *Small and Short-Range Radar Systems* (1 ed.). CRC Press, Inc., Boca Raton, FL, USA.
6. Joaquim Fortuny-Guasch and Jean-Marc Chareau. 2013. *Radar Cross Section Measurements of Pedestrian Dummies and Humans in the 24/77 GHz Frequency Bands: Establishment of a Reference Library of RCS Signatures of Pedestrian Dummies in the Automotive Radar Bands*. Technical Report.
7. B. Ginsburg. 2018. F3: Circuits and architectures for wireless sensing, radar and imaging. In *2018 IEEE Int'l ISSCC*. 508–510.
8. Karl Granstrom, Christian Lundquist, and Omut Orguner. 2012. Extended target tracking using a Gaussian-mixture PHD filter. *IEEE Trans. Aerospace Electron. Systems* 48, 4 (2012), 3268–3286.
9. Anil Kumar, Yaakov Bar-Shalom, and Eliezer Oron. 1995. Precision tracking based on segmentation with optimal layering for imaging sensors. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 17, 2 (1995), 182–188.
10. Wenjie Luo, Bin Yang, and Raquel Urtasun. 2018. Fast and Furious: Real Time End-to-End 3D Detection, Tracking and Motion Forecasting With a Single Convolutional Net. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3569–3577.
11. Novelda AS. 2018a. XeThru X4M03 custom sensor development. (2018). `https://www.xethru.com/xethru-development-platform.html`.
12. Novelda AS. 2018b. XeThru X4M200 respiration sensor. (2018). `https://www.xethru.com/x4m200-respiration-sensor.html`.
13. Eliezer Oron, Anil Kumar, and Yaakov Bar-Shalom. 1993. Precision tracking with segmentation for imaging sensors. *IEEE Trans. on aerospace and electronic systems* 29, 3 (1993), 977–987.
14. Hermann Rohling. 1983. Radar CFAR thresholding in clutter and multiple target situations. *IEEE Trans. on Aerospace and Electronic Systems* 4 (1983), 608–621.
15. Vayyar Imaging Ltd. 2018. Walabot. (2018). `https://walabot.com/makers`.
16. Mingmin Zhao, Fadel Adib, and Dina Katabi. 2016. Emotion Recognition Using Wireless Signals. In *Proc. of the 22Nd Annual Int'l. Conf. on Mobile Computing and Networking (MobiCom '16)*. ACM, New York, NY, USA, 95–108.