

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/297608745>

Device-Free Radio Vision for Assisted Living: Leveraging wireless channel quality information for human sensing

ARTICLE *in* IEEE SIGNAL PROCESSING MAGAZINE · MARCH 2016

Impact Factor: 5.85 · DOI: 10.1109/MSP.2015.2496324

READS

6

6 AUTHORS, INCLUDING:



[Stefano Savazzi](#)

Italian National Research Council

62 PUBLICATIONS 644 CITATIONS

[SEE PROFILE](#)



[Vittorio Rampa](#)

Italian National Research Council

49 PUBLICATIONS 273 CITATIONS

[SEE PROFILE](#)



[Sanaz Kianoush](#)

University of Pavia

7 PUBLICATIONS 7 CITATIONS

[SEE PROFILE](#)



[Umberto Spagnolini](#)

Politecnico di Milano

227 PUBLICATIONS 2,872 CITATIONS

[SEE PROFILE](#)

Copyright notice

This is the accepted version of the IEEE-copyrighted article “Device-Free Radio Vision for Assisted Living: Leveraging wireless channel quality information for human sensing”, already available online at the link:

<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7426567>.

It will be published in the IEEE Signal Processing Magazine vol. 33, no. 2, pp. 45-58, March 2016, DOI: 10.1109/MSP.2015.2496324.

© © 20xx IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Device-Free Radio Vision for Assisted Living

[Leveraging wireless channel quality information for human sensing]

Stefano Savazzi, Stephan Sigg, Monica Nicoli, Vittorio Rampa, Sanaz Kianoush, Umberto Spagnolini

Wireless propagation is conventionally considered as the enabling tool for transporting information in digital communications. However, recent research has shown that the perturbations of the same electromagnetic fields that are adopted for data transmission can be used as powerful sensing tools for *device-free* radio vision. Applications range from human body motion detection and localization to passive gesture recognition. In line with the current evolution of mobile phone sensing [1], radio terminals are not only ubiquitous communication interfaces, but they also incorporate novel or augmented sensing potential, capable of acquiring an accurate human-scale understanding of space and motion. This article shows how radio-frequency signals can be employed to provide a device-free environmental vision, and investigates the detection and tracking capabilities for potential benefits in daily life.

1. INTRODUCTION

“It’s not difficult. Every time I lift my arm, it distorts a small electromagnetic field that is maintained continuously across the room. Slightly different positions of my hand and fingers produce different distortions and my robots can interpret these distortions as orders. I only use it for simple orders: Come here! Bring tea! and so on.” Isaac Asimov, The Robots of Dawn, 1983.

Device-free radio vision is an augmented functionality provided by radio transceivers – typically heterogeneous, densely distributed and networked – that monitor the fluctuations of the electromagnetic (EM) field across the space. These monitoring devices may be pre-existing, deployed at arbitrary (or optimized) locations for communication purposes in the area of interest, and exchange digital information by any wireless communication protocol. Radio vision systems leverage diffraction, reflection and scattering phenomena that affect radio-frequency (RF) propagation for ubiquitous sensing. RF signals can be either narrowband or wideband, in licensed or unlicensed frequency bands, with carrier frequencies ranging from MHz to GHz, and above. The presence, position and motion of a human body in the network area affect the EM field in a predictable way,

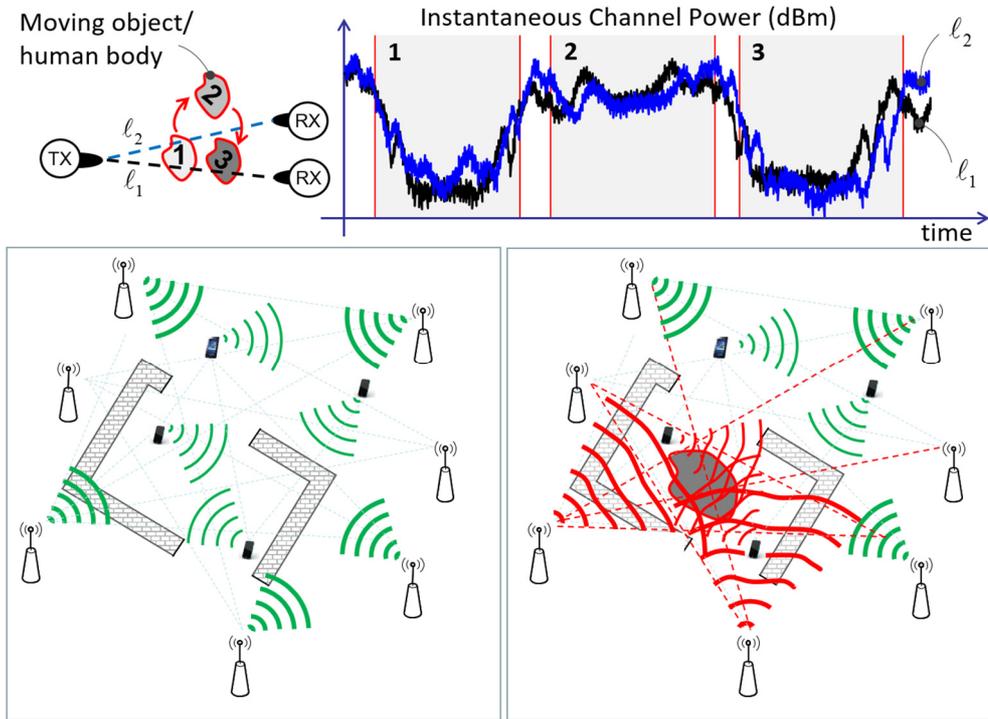


Figure 1: Device-free radio vision is based on tracking the perturbations of RF field sensed by dense networks of radio-interacting wireless devices.

making it possible to estimate and track its activity without the need to deploy and calibrate any additional wearable sensor (*sensor-free detection*), nor to ask for specific user actions (*non-cooperative detection*). This passive sensing approach has been experimented with in heterogeneous networks but it is also appropriate for most of the emerging low-power wireless standards, and for personal and device-to-device (D2D) communications [2], including WiFi, Bluetooth low energy (BLE), ZigBee and D2D enabled long term evolution (LTE Advanced) [3].

Tracking and recognition of human motions and activities are done through real-time processing of the wireless channel quality information (CQI). In this paper, leading edge research and developments are discussed with special focus on assisted living applications [4].

2. LEVERAGING RF SIGNALS FOR SENSING: DEVICE-FREE VISION

Personal sensing is the current scale at which these technologies are being studied by the research community: they are designed for sensing a single (or a limited number of) individual(s) based on real-time analysis of CQI. As depicted in Fig. 1, radio-based vision systems track RF field perturbations by dense networks of air-interacting wireless devices and process CQI data for the purpose of human sensing. In order to support “vision” functions, three key distinctive technological features are incorporated:

- *Sensor-less Interaction and Anonymous Tracking.* Gesture-based interactions of the user with the environment are detected without instrumenting the human body (device-free) or deploying sensors calibrated for each user (sensor-less). Subject(s) are anonymously tracked and localized, in contrast to privacy intrusive video-cameras, inferring the EM perturbations from CQI.
- *Ubiquitous Monitoring.* Unlike existing infrared (IR) recognition platforms [4], device-free radio vision systems support ubiquitous user detection in complex non-line-of-sight (NLOS) indoor spaces [5], using both fixed (e.g., WiFi access points, ZigBee/Bluetooth devices) and nomadic (e.g., smart-phones, tablets) radio devices (see Fig. 1) that are interacting over mixed line-of-sight (LOS) and NLOS), or through-the-wall links [6]. RF signals with wavelengths that are long enough to penetrate dense objects, such as doors or walls, can be exploited to recognize human motion and gestures even if these gestures are visually in shadow or in a different room adjacent to the one where the RF device itself is operating.
- *Scalable CQI (Big-Data) Analytics.* The technology typically requires information aggregation, processing and computation of massive amounts of CQI data generated from, and delivered to, highly distributed and heterogeneous wireless devices. CQI data for real-time processing are often produced at high rates, in the order of tens of thousands of observations per second to cover large spaces. Learning and running analytics from these large volumes of data requires the use of signal processing tools designed to efficiently work on high-dimensional and often incomplete data-sets [7] (e.g., due to random power cycling of devices or communication failures).

ACTIVE AND PASSIVE CONFIGURATIONS. Device-free radio vision systems can be based on active or passive configurations as illustrated in Fig. 2. The distinction between active and passive systems differentiates systems in which the active part (the transmitter) is under the control of the system from those where it is not [8]. *Passive* systems capitalize on a pre-existing network infrastructure where densely air-interacting devices are exposed to some EM fields (e.g., FM radio [8], WiFi [5][9]) and capture those ambient RF signals. CQI processing might be carried out distributedly or centrally. *Active* systems exploit dense communications with fixed/nomadic transmitters acting as interconnected mobile probes. These systems typically rely on a decentralized architecture where user data are propagated in direct mode instead of through a remote service provider

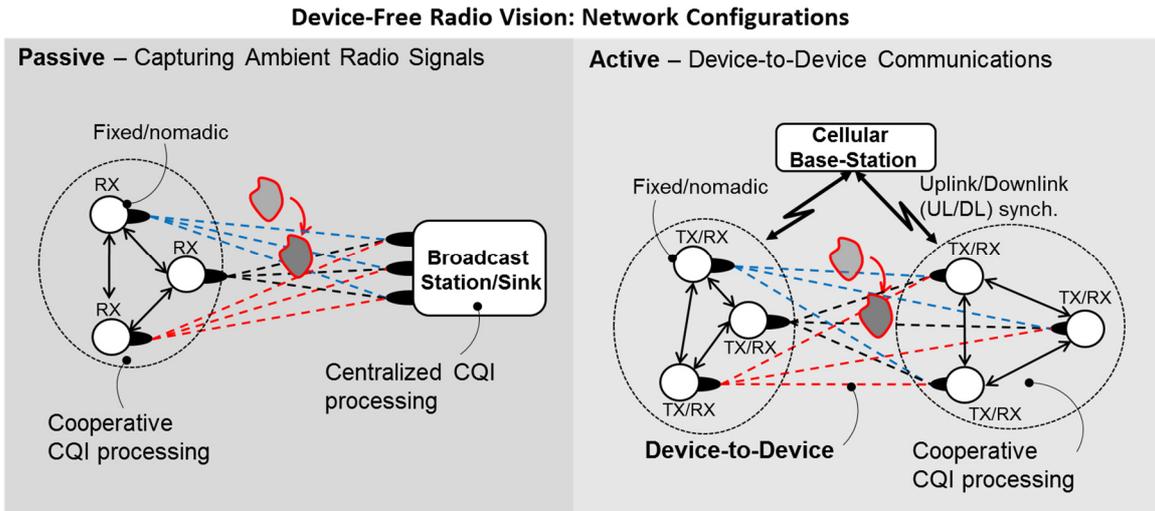


Figure 2: Active/passive network configurations for device-free radio vision. Transmitter (TX), receiver (RX).

(e.g., cellular base stations, WiFi access points), even if providers might trigger the first device connection, for logging, uplink/downlink (UL/DL) synchronization, etc. This concept is in line with the current trend [3] of enabling small/femto-cell deployments with smartphones able to discover other phones in proximity, and overhear RF signals from device-to-device links.

3. MODELING OF RF SIGNALS FOR RADIO VISION

In radio vision systems, CQI measurements used for recognition can be either in the form of physical (PHY) layer values, e.g., the baseband radio channel state information (CSI) sampled at symbol level, or received signal strength (RSS) data extracted at upper layers.

Let us consider a wireless transmission organized into periodic frames consisting of groups of adjacent symbols, and a human body, located at position \mathbf{x} inside the wireless link area, performing an activity δ defined as an ensemble of non-rigid body motions [10]. The user state $\Theta = [\mathbf{x}, \delta]$ defines a generic combination of user location and activity. The effects of the user state Θ on the channel response are observed over a set $\mathcal{T} = \{1, \dots, T\}$ of consecutive received symbols (or frames). For a static environment where the human body does not obstruct the link (i.e., human-free state $\Theta = \emptyset$), the equivalent base-band channel response $h(\tau|\emptyset) = \sum_{k=0}^N \alpha_k g_{\tau-\tau_k} e^{-j\phi_k}$ can be modeled as multipath with a combination of N delayed paths: α_k and ϕ_k are the amplitude and the phase shift of the k -th ray, respectively, and $g_{\tau-\tau_k}$ models the received pulse waveform with delay τ_k . Human in state Θ modifies the channel response at symbol time $t \in \mathcal{T}$ as

$$h_t(\tau|\Theta) = \sum_{k=0}^N \alpha_k(t|\Theta) g_{\tau-\tau_k(t|\Theta)} e^{-j\phi_k(t|\Theta)}, \quad (1)$$

where the amplitude $\alpha_k(t|\Theta) = \alpha_k + \Delta\alpha_k(t|\Theta)$, the phase shift $\phi_k(t|\Theta) = \phi_k + \Delta\phi_k(t|\Theta)$ and the augmented delay $\tau_k(t|\Theta) = \tau_k + \Delta\tau_k(t|\Theta)$ of the k -th ray highlight the human-induced perturbations compared to the human-free state $\Theta = \emptyset$. Amplitude $\alpha_k(t|\Theta)$ and phase shift $\phi_k(t|\Theta)$ incorporate human-induced micro-Doppler effects.

In what follows, the effect of human body motion on CQI is experimentally evaluated, either for single and multicarrier digital communication systems. An introductory case-study is shown in Fig. 3, where detection of human motion is based on RSS (Fig. 3.a), and CSI measurements extracted from an orthogonal frequency division multiplexing (OFDM) implementation (Fig. 3.b).

RECEIVED SIGNAL STRENGTH. The RSS is a practical metric to assess CQI at frame level, and it is commonly adopted for transmitter (TX) – receiver (RX) link adaptation and *link-layer* transmission scheduling tasks. Power estimators, or peak detectors, are commonly used to acquire information about signal strength as depicted in Fig. 3.a: the automatic gain control (AGC) loop exploits RSS information to adapt the dynamic range before A/D conversion. At frame time t , the RSS s_t can be modelled in logarithmic (dB) scale as

$$s_t(\Theta) = s(\emptyset) + \Delta s_t(\Theta), \quad (2)$$

where the additive deviation $\Delta s_t(\Theta)$ from $\Theta = \emptyset$ models the body-induced perturbation and $s(\emptyset) = \mathbb{E}[s_t(\Theta = \emptyset)]$ is the (average) RSS observed in the human-free state. The sequence $\mathbf{s}(\Theta) = [s_t(\Theta)]_{t \in \mathcal{T}} \in \mathbb{R}^{T \times 1}$ collects the human-induced RSS footprint observed over T frames. Likewise, the *RSS profile* is the deviation with respect to $\Theta = \emptyset$: $\Delta \mathbf{s}(\Theta) = [s_t(\Theta) - s(\emptyset)]_{t \in \mathcal{T}} \in \mathbb{R}^{T \times 1}$. In IEEE 802.15.4 standard-compliant devices, the digital RSS indicator (RSSI) \hat{s}_t can be used as estimator of the RSS with 8-bit resolution. Other radios also implement the link quality indicator (LQI) that correlates with packet reception rate (see [11] and references therein), and provides an indirect estimate of RSS values. In WiFi standards, RSS estimation can be obtained from the received channel power indicator (RCPI). In LTE, the reference signal received power (RSRP) measures the power over the reference signals.

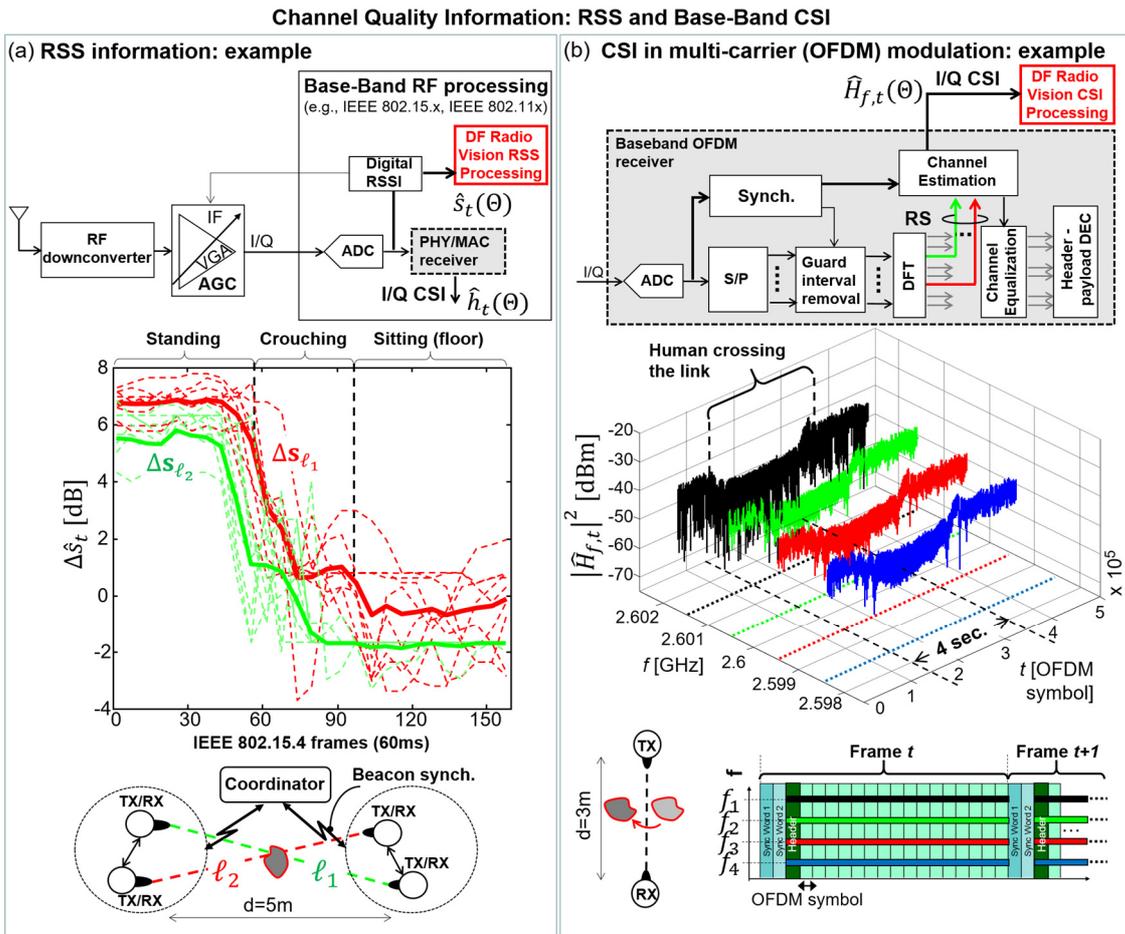


Figure 3.a: RSS processing for device-free radio vision: (top) digital/analog RSSI extraction, (bottom) RSSI data (dashed curves) – and corresponding average profiles (solid curves) - for human standing, crouching and sitting on the floor ($T = 160$ IEEE 802.15.4 frames). Active configuration is considered with two IEEE 802.15.4 links. **Figure 3.b:** baseband processing example for OFDM modulation: (top) CSI extraction, (bottom) CSI power footprints over $K=4$ OFDM pilot subcarriers ($T = 223000$ symbols) corresponding to human body crossing the link in 4 sec. OFDM implementation: 2.6 GHz, 64 subcarriers, cyclic prefix 16 samples, baseband sampling 5MHz with 16 payload symbols per frame.

Recognition of human activity can be also based on multi-link processing. Fig. 3.a (bottom) shows the profiles $\Delta \mathbf{s}_\ell(\theta) \in \mathbb{R}^{T \times 1}$ over two IEEE 802.15.4 links $\ell \in \{\ell_1, \ell_2\}$, for $T = 160$ frames, based on 20 RSSI independent measurements $\hat{\mathbf{s}}_\ell(\theta) = [\hat{s}_{\ell,t}(\theta)]_{t \in \mathcal{T}} \in \mathbb{R}^{T \times 1}$ featuring a human body standing in the surroundings of both links (located at 0.5 m above the ground), then crouching and sitting on the floor. Since the human torso causes more attenuation than waist and ankles, when a person is standing, there is a larger RSS attenuation with respect to the same body lying on the floor. The profiles $\Delta \mathbf{s}_\ell(\theta)$, superimposed in solid lines, average out noise and time-warping effects. Detection of the human state can be based on matching (e.g., using simple time-domain features) the observed entries $\hat{\mathbf{s}}_\ell(\theta)$, or the estimated deviations $\Delta \hat{\mathbf{s}}_\ell(\theta)$, with the corresponding RSS profiles learned during a training procedure. Human state estimation possibly entails de-noising, time-

warping and reconstruction of missing RSSIs observations (i.e., by interpolation methods) [12][13]. Missing or incomplete data can be represented as $\wp_{\Omega}[\hat{\mathbf{S}}_{\ell}(\Theta)]$ over the set of $t \in \Omega \subseteq \mathcal{T}$ received frames, where $\wp_{\Omega}(\cdot)$ is the sampling operator nulling the entries of $\hat{\mathbf{S}}_{\ell}(\Theta)$ not in Ω [7].

BASEBAND MODELING OF C.S.I. Baseband CSI measures the channel response at symbol-level: CSI estimation is typically obtained from training/reference signals (RS) multiplexed with information symbols and periodically placed in every frame. Therefore, in contrast to RSS, processing of CSI information for the purpose of radio vision can leverage on multiple independent measurements at frame level and can be used to capture fast human body movements and gestures. Assuming frequency-flat channel as for narrowband communication but time-varying for moving multipath environments, the received RS $r_t = h_t(\Theta)\omega_t + n_t$ at symbol time $t \in \mathcal{T}$ (with ω_t and n_t the transmitted RS and the noise term, respectively), captures the moving body in state Θ through the corresponding complex channel envelope adopted from (1)

$$h_t(\Theta) = \sum_{k=0}^N \alpha_k(t|\Theta) e^{-j\phi_k(t|\Theta)} = h(\emptyset) + \Delta h_t(\Theta). \quad (3)$$

Human body effects on channel response are now embedded into a characteristic footprint of channel variations over T received symbols $\mathbf{h}(\Theta) \in \mathbb{C}^{T \times 1} = [h_t(\Theta)]_{t \in \mathcal{T}}$. The *CSI profile* set is $\Delta \mathbf{h}(\Theta) \in \mathbb{C}^{T \times 1} = [\Delta h_t(\Theta) = h_t(\Theta) - h(\emptyset)]_{t \in \mathcal{T}}$, with $h(\emptyset) = \mathbb{E}[h_t(\Theta = \emptyset)]$ being the average response for the human-free state. Noisy profiles $\Delta \hat{\mathbf{h}}(\Theta) = [\hat{h}_t(\Theta) - \hat{h}(\emptyset)]_{t \in \Omega}$ with estimated channels $\hat{h}_t(\Theta)$ and human-free response $\hat{h}(\emptyset)$, are typically observed over a subset of times (or symbol indexes) $\Omega \subseteq \mathcal{T}$ accounting for the training/data multiplexing, and missing symbols.

The use of multi-carrier (OFDM) modulation enables multi-dimensional processing of CSI over the time-frequency grid and thus allows a fine-grained classification of human-motion [12]. As depicted in Fig. 3.b, the CSI estimation is carried out by periodic transmission of RSs over standard defined time-frequency patterns [3]. The received RSs \mathbf{r}_t over the K pilot subcarriers $\{f_1, \dots, f_K\}$ inside OFDM symbol t can be written as $\mathbf{r}_t = \text{diag}[\boldsymbol{\omega}_t] \cdot \mathbf{h}_t(\Theta) + \mathbf{n}_t$ with vector $\boldsymbol{\omega}_t$ collecting the transmitted RSs, and baseband channel vector $\mathbf{h}_t(\Theta) = [H_{f,t}(\Theta)]_{f=f_1}^{f_K}$ containing the Fourier Transform $\mathcal{F}(\cdot)$ of channel $h_t(\tau|\Theta)$

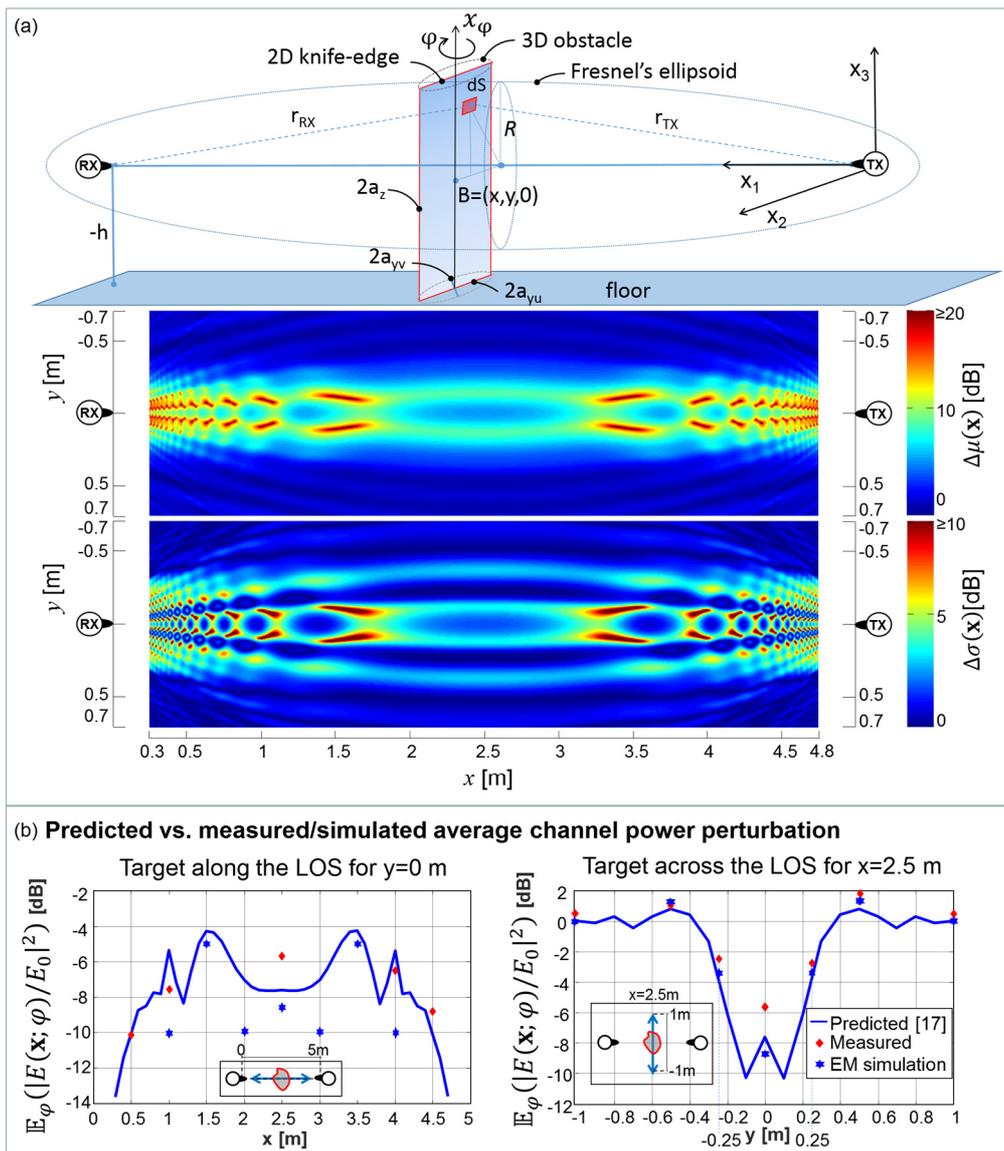


Figure 4.a: Link layout (top) and perturbation maps (bottom) of RF attenuation for mean and standard deviation. Figure 4.b: Predicted vs. measured and EM simulated average channel power perturbations (along and across the LOS path). Measurements are obtained with a person standing in x with varying posture. Predicted values are obtained for a rotating cylinder moving inside a 10×10 cm bin centered in x .

$$H_{f,t}(\theta) = H_f(\emptyset) + \underbrace{\mathcal{F}(\Delta h_t(\tau|\theta))|_f}_{\Delta H_{f,t}(\theta)} \quad (4)$$

The CSI footprint is the matrix $\mathbf{H}(\theta) \in \mathbb{C}^{K \times T} = [\mathbf{h}_1(\theta), \dots, \mathbf{h}_T(\theta)]_{t \in \mathcal{T}}$ with human-induced profile $\Delta \mathbf{H}(\theta) = [\mathbf{H}(\theta) - \mathbf{H}(\emptyset)]$. The estimate $\hat{\mathbf{H}}(\theta)$ is evaluated over the time-frequency set Ω now accounting for framing structure and irregular time-frequency RSs spacing. In Fig. 3.b (bottom) an OFDM transmission over 2.6 GHz is implemented in-lab using software defined radio (SDR) devices: a person is crossing the link and standing for 4s, causing an average attenuation of 5 dB. The CSI power footprint estimates $|\hat{H}_{f,t}(\theta)|^2$ are shown for $K = 4$ pilots and $T = 223000$ symbols.

A crucial problem for quantitative evaluation of radio vision system performance is the availability of a simple but realistic model to describe human body-induced shadowing. Ray-tracing [14], EM/stochastic [15][16], and geometric-based (see [2] for a review) models have been investigated to predict the correlation between the human body position \mathbf{x} and the corresponding channel perturbations. EM methods, that exploit geometric/uniform theory of diffraction (GTD/UTD) and ray-tracing algorithms as well, can be employed for their ability to accurately evaluate the EM field at the receiver, but they are usually very complex, time consuming and, above all, require perfect knowledge of the shape, composition and properties of the obstacle. In the tutorial “Diffraction based modeling of human body shadowing”, we consider a simplified but effective framework based on the Fresnel-Kirchoff diffraction theory and shown in the scenario of Fig. 4.a.

TUTORIAL: DIFFRACTION-BASED MODELING OF HUMAN BODY SHADOWING

We consider here the framework proposed in [16] and summarized according to the link scenario of Fig. 4.a (top), where the human body is sketched as a 3D homogeneous cylinder, having height $2a_z$ and an elliptical base with semi-axes a_{yu} and a_{yv} . The cylinder rotates along the vertical axis x_φ with azimuth φ thus modeling a human body standing in a specific position but possibly changing its posture. As a trade-off between model simplicity and accuracy, the 3D cylinder is reduced to an equivalent 2D rectangular knife-edge absorbing surface [11][16] that is modeled according to the Fresnel-Kirchoff diffraction theory. The obstacle has the same height $2a_z$ but variable traversal semi-size $a_y(\varphi)$ with $a_{yv} \leq a_y(\varphi) \leq a_{yu}$ while its barycenter B is placed in position $\mathbf{x} = [x_1 = x, x_2 = y]$. If the obstacle location \mathbf{x} is near the LOS path, and if multipath body-induced effects are neglected, by exploiting the paraxial approximation, the attenuation term in dB scale

$$\Delta E_{\text{dB}}^2(\mathbf{x}; \varphi) = -10 \log_{10} |E(\mathbf{x}; \varphi)/E_0|^2 \quad (5)$$

due to the obstacle is derived analytically in [16]. $\Delta E_{\text{dB}}^2(\mathbf{x}; \varphi)$ depends on the received electric field $E(\mathbf{x}; \varphi)$ and the free-space case E_0 . Focusing on body localization, attenuation can be considered as random due to the varying orientations φ of the obstruction body. Fig. 4.a shows the mean

$$\Delta\mu(\mathbf{x}) = \mathbb{E}_\varphi [\Delta E_{\text{dB}}^2(\mathbf{x}; \varphi)] \text{ and the standard deviation } \Delta\sigma(\mathbf{x}) = \sqrt{\mathbb{E}_\varphi \left[\left(\Delta E_{\text{dB}}^2(\mathbf{x}; \varphi) - \Delta\mu(\mathbf{x}) \right)^2 \right]}$$

perturbation maps of the RF attenuation for the link of length d as a function of the obstacle position \mathbf{x} after averaging with respect to azimuth ($d = 5$ m, 2.48 GHz, with $a_z = 90$ cm, $a_{yu} = 27.5$ cm, $a_{yv} = 12$ cm). With the given geometrical constants, the *sensitivity area* \mathcal{X} due to the obstacle has a width of about 0.7 m around the LOS path. The model (5) neglects the true shape, complex composition and EM properties of the obstacle (e.g., polarization, permittivity and conductivity), but it is accurate enough to model human-induced attenuation effects. Fig. 4.b shows the comparison of the average channel power perturbation $\mathbb{E}_\varphi[(|E(\mathbf{x}; \varphi)/E_0|^2)_{\text{dB}}]$ induced by a person against the values predicted by the model (5) and the ones obtained by simulating the obstacle as a perfect conductor, having the same size of the person, placed over a concrete floor. As expected, the impact of the target presence is higher along the LOS path and close to the transmitting/receiving devices.

4. RESEARCH ON RADIO VISION: A SURVEY

Recently, there has been an increasing research interest in wireless human tracking via RF devices. This broadly defined domain encompasses different research areas such as signal processing, computer vision, communication networks, and human-machine interfaces. The first experimental activity dates back to the works [17][18] showing that body motions leave a characteristic footprint on RSS patterns [17], while RSS fluctuations can be effectively used for body localization [18].

Focusing on device-free human body localization, the radio tomography imaging (RTI) proposed in [6][19][20] adopts computed tomography methods to reconstruct an image of the object(s) inside the network area. The technology has been now transferred to a commercial product (i.e., Xandem system) targeting assisted living applications. The methods introduced in [5][9][11] allow to explicitly track the position of target(s) using a Bayesian approach that jointly process the RSS mean and standard deviations [11]. More recently, device-free systems based on Bayesian tracking of RSS profiles have been also designed for obstacle/object 2D mapping [21], detection of human breathing [22][13], and fall detection [23][24]. Human gesture recognition and body motion detection have been addressed in recent research projects (SenseWaves, E-eyes, WiSee and Wi-Vi) targeting both RSS [25] and baseband CSI analytics using radio devices operating at 900 MHz

[8][26], 2.4 GHz with 20 MHz band WiFi-compliant RF frontends, [27][12] and 5.8 GHz [28]. Signal processing methodologies for these systems, most recent developments and open issues will be discussed in the following sections.

5. DEVICE-FREE LOCALIZATION AND MOTION TRACKING (DF-L)

In the context of assisted living, knowledge of the user location [5] is important for a number of services ranging from monitoring daily activities, forecasting user tendencies to remote control of appliances (e.g., lights, doors, windows, air conditioners). RSS-based device-free localization (DF-L) [11][18][19] has emerged in the last few years for passive localization of people movements in areas covered by pre-existing wireless mesh networks. Considering a single target at location \mathbf{x} in the network area, if s_ℓ is the RSS of the link ℓ at a given time instant, the objective of DF-L is the estimation of \mathbf{x} from the RSS set $\mathbf{s} = [s_1, \dots, s_L]^T$. As shown in Fig. 5, RSS is here approximated as a Gaussian variable in the logarithmic domain (i.e., as log-normal shadowing [11]):

$$s_\ell = \begin{cases} \mu_{0,\ell} + w_{0,\ell}, & \mathbf{x} \notin \mathcal{X}_\ell \\ \mu_{1,\ell}(\mathbf{x}) + w_{1,\ell}, & \mathbf{x} \in \mathcal{X}_\ell \end{cases} \quad (6)$$

for target outside ($\mathbf{x} \notin \mathcal{X}_\ell$) or inside ($\mathbf{x} \in \mathcal{X}_\ell$) the sensitivity area of the link ℓ , respectively. In the first case (empty scenario), the RSS has a deterministic mean $\mu_{0,\ell}$, that accounts for path-loss and other static effects due to fixed obstructions or scattering objects, while the random term $w_{0,\ell} \sim \mathcal{N}(0, \sigma_{0,\ell}^2)$ accounts for RSS measurement errors due to the hardware (i.e., for noisy RSSI) as well as for small power fluctuations due to variations in the surrounding environment. The mean RSS for $\mathbf{x} \in \mathcal{X}_\ell$ is $\mu_{1,\ell}(\mathbf{x}) = \mu_{0,\ell} + \Delta\mu_\ell(\mathbf{x})$ (with $\Delta\mu_\ell(\mathbf{x}) \leq 0$), while the random shadowing is modeled as $w_{1,\ell} \sim \mathcal{N}(0, \sigma_{1,\ell}^2(\mathbf{x}))$ with $\sigma_{1,\ell}(\mathbf{x}) = \sigma_{0,\ell} + \Delta\sigma_\ell(\mathbf{x})$ (and $\Delta\sigma_\ell(\mathbf{x}) \geq 0$). An experimental evidence for model (6) is depicted in Fig. 5 (top): a measurement campaign was conducted to evaluate experimental RSS distributions to be compared with model (6) for target moving in the surrounding of location \mathbf{x} , both outside (blue) and inside (red) the sensitivity area of two selected links. Although better fits can be provided using other parametric distributions (Weibull, Nakagami [29]), the approximation is reasonable to design estimation methods and infer performance bounds of

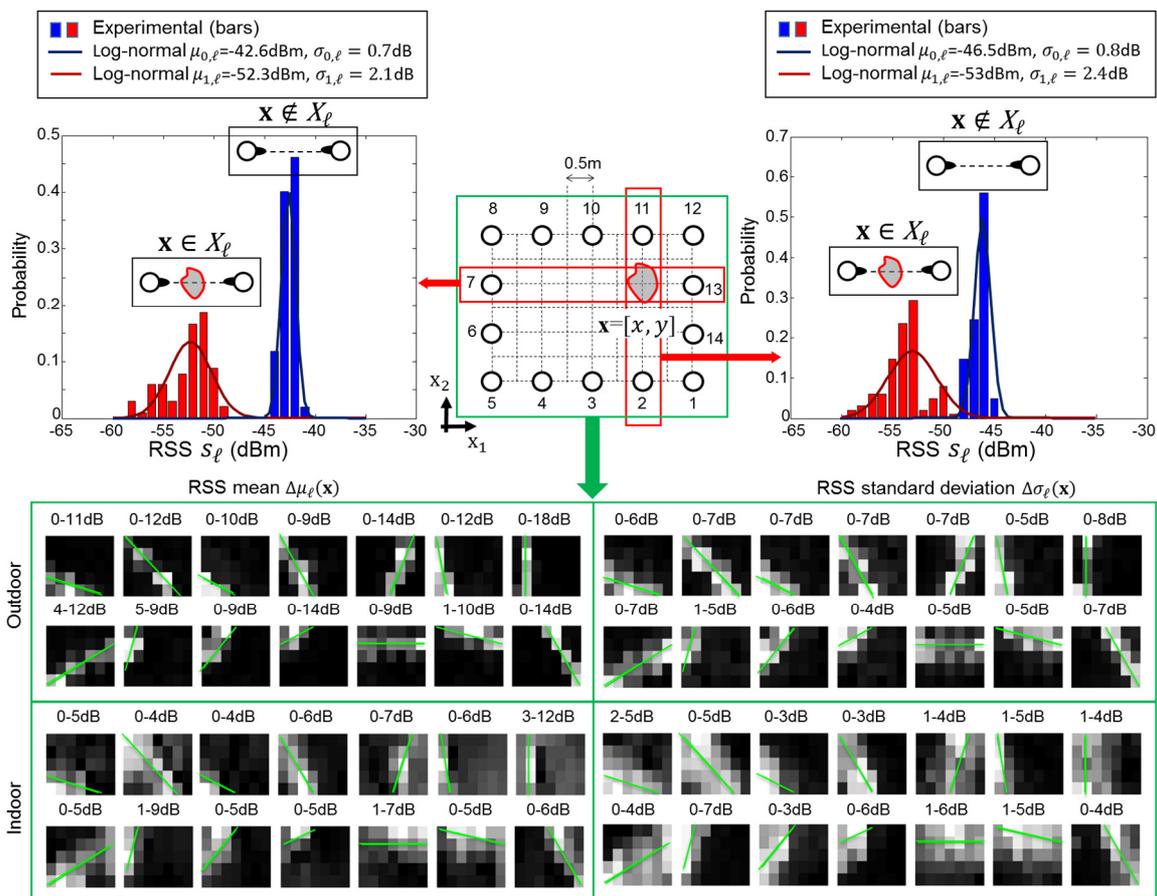


Figure 5: Log-normal RSS modeling for two links (top-left and top-right) and for target inside (red) and outside (blue) the sensitivity area \mathcal{X}_ℓ for link ℓ . Bottom: experimental RSS mean and standard deviation maps vs. the target location \mathbf{x} for the highlighted network layout (center) and some links in indoor and outdoor scenarios.

practical interest (see tutorial “Accuracy bounds for DF-L”). As shown in several experimental studies [6][11][20], both average $\Delta\mu_\ell(\mathbf{x})$ and standard deviation $\Delta\sigma_\ell(\mathbf{x})$ maps of the RSS perturbations (i.e., *perturbation maps*) are strongly related to the target position relative to the link. This can be further appreciated in Fig. 5, where 2D experimental RSS profiles are shown at bottom for a subset of links. Based on the RSS measurements collected by a network of 14 nodes deployed in indoor and outdoor scenarios, the maps are evaluated for a human body moving over the positions \mathbf{x} within a 4×3 m area. The sensitivity area \mathcal{X}_ℓ for each link ℓ is centered on the LOS path (highlighted in green) and it is larger in indoor than in outdoor due to multipath effects.

For position estimation, knowledge of the reference parameters $\{\mu_{0,\ell}, \sigma_{0,\ell}\}$ (for the human-free case) is required for all links $\ell = 1, \dots, L$, together with the information about the perturbation maps $\{\Delta\mu_\ell(\mathbf{x}), \Delta\sigma_\ell(\mathbf{x})\}$ for all position values $\mathbf{x} \in \mathcal{X}_\ell$. While $\{\mu_{0,\ell}, \sigma_{0,\ell}\}$ can be easily pre-calibrated when no target is moving in the network area, evaluation of profiles $\{\Delta\mu_\ell(\mathbf{x}), \Delta\sigma_\ell(\mathbf{x})\}$ is more critical as it requires extensive fingerprinting campaigns [9][11][19] or ray-tracing simulations [14].

Analytical modeling, when viable, has to be preferred as it allows to simplify the calibration to few model parameters and/or to evaluate pre-deployment performance. In [19] and [20], a simple single-parameter model is considered where $\Delta\mu_\ell(\mathbf{x})$, and $\Delta\sigma_\ell(\mathbf{x})$ are assumed to be constant and inversely proportional to the square root of the link distance for $\mathbf{x} \in \mathcal{X}_\ell$ with \mathcal{X}_ℓ modelled as an ellipsoid with foci at the two nodes. Parametric models for shadowing effects can be also found in [2], while diffraction-based models are considered in [16]. Once the perturbation maps are available, the target position \mathbf{x} can be estimated from model (6) using inference methods. The weighted least squares (WLS) criterion is $\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{s} - \boldsymbol{\mu}(\mathbf{x})\|_{\mathbf{C}^{-1}(\mathbf{x})}$, with $\|\mathbf{s} - \boldsymbol{\mu}(\mathbf{x})\|_{\mathbf{C}^{-1}(\mathbf{x})} = [\mathbf{s} - \boldsymbol{\mu}(\mathbf{x})]^T \mathbf{C}^{-1}(\mathbf{x}) [\mathbf{s} - \boldsymbol{\mu}(\mathbf{x})]$, $\boldsymbol{\mu}(\mathbf{x}) = [\mu_1(\mathbf{x}) \cdots \mu_L(\mathbf{x})]^T$, and covariance $\mathbf{C}(\mathbf{x}) = \text{diag}[\sigma_1^2(\mathbf{x}), \dots, \sigma_L^2(\mathbf{x})]$ as weighting factor. Assuming the RSS fluctuations as independent over the links, the maximum likelihood (ML) criterion also applies as $\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} \mathcal{L}(\mathbf{s}|\mathbf{x})$, with the log-likelihood function $\mathcal{L}(\mathbf{s}|\mathbf{x}) = -\ln(\det[\mathbf{C}(\mathbf{x})]) - \|\mathbf{s} - \boldsymbol{\mu}(\mathbf{x})\|_{\mathbf{C}^{-1}(\mathbf{x})}^2$ and $\det[\cdot]$ the determinant operator. The information provided by instantaneous measurements \mathbf{s} can also be augmented with prior information about the target motion, using sequential Bayesian filtering such as Kalman (KF), grid-based (GF) or a particle (PF) filtering [11].

Another DF-L approach is the radio tomographic imaging (RTI) [19], where the DF-L problem is formulated as the estimation of a *motion image* of the area, capturing any variation with respect to the human-free scenario observed during the calibration phase. In RTI, the area is divided into M voxels, $m = 1, \dots, M$, the image to be estimated is $\mathbf{v} = [v_1 \cdots v_M]^T$ where $0 \leq v_m \leq 1$ measures the target occupancy (i.e., in terms of ‘‘probability’’ metric) of voxel m . For sparse motion, RSS is approximated as the sum of the contributions generated by all occupied voxels:

$$s_\ell = \sum_{m=1}^M \Delta\mu_\ell(m) v_m + \mu_{0,\ell} + w_\ell \quad (7)$$

where $\Delta\mu_\ell(m)$ is now the attenuation contribution due to target in voxel m and $w_\ell \sim \mathcal{N}(0, \sigma_\ell^2)$ the RSS fluctuation. Considering all links, it is $\mathbf{s} = \Delta\boldsymbol{\mu} \cdot \mathbf{v} + \boldsymbol{\mu}_0 + \mathbf{w}$, with matrix $\Delta\boldsymbol{\mu} \in \mathbb{R}^{L \times M} = [\Delta\mu_\ell(m)]$ that collects the perturbations for all links and voxels, $\boldsymbol{\mu}_0 = [\mu_{0,1} \cdots \mu_{0,L}]^T$ the human-

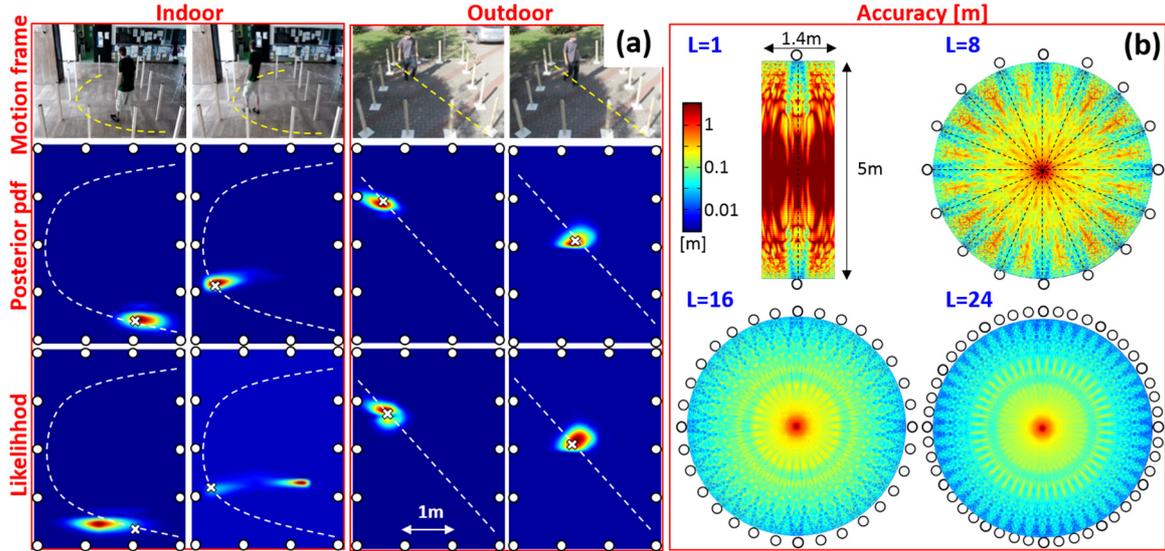


Figure 6.a: DF-L: belief images of the target location for two frames of the user trajectories in 4mx3m indoor and outdoor scenarios. Figure 6.b: CRB limit to the DF-L accuracy considering $L=1, 8, 16, 24$ RF links of 5m length.

free reference vector, and $\mathbf{w} = [w_1 \cdots w_L]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ the shadowing term. LS estimation has been proposed for imaging solution, as $\hat{\mathbf{v}} = (\Delta\boldsymbol{\mu}^T \Delta\boldsymbol{\mu} + \boldsymbol{\Gamma}_{x_1}^T \boldsymbol{\Gamma}_{x_1} + \boldsymbol{\Gamma}_{x_2}^T \boldsymbol{\Gamma}_{x_2})^{-1} \Delta\boldsymbol{\mu}^T \mathbf{s}$ with $\boldsymbol{\Gamma}_{x_i}$ accounting for regularization along each direction [19]. The target position corresponds to the voxel associated with the maximum image value: $\hat{m} = \underset{m}{\operatorname{argmax}} \hat{v}_m$. In [20], RTI has been also applied to RSS variances modelled as linearly increasing with the number of occupied voxels $\sigma_\ell^2 = \sigma_\ell^2(\mathbf{v}) = \sigma_{0,\ell}^2 + \sum_{m=1}^M \Delta\sigma_\ell^2(m) v_m$: the same LS approach is then adopted for image evaluation.

Examples of DF-L results based on both mean and variance measurements are given in Fig. 6.a, for indoor (left) and outdoor (right) scenarios, where a human target moves according to the highlighted trajectory. The snapshot likelihood $\mathcal{L}(\mathbf{s}|\mathbf{x})$ and the posterior pdf $p(\mathbf{x}|\mathbf{s})$ are evaluated by GF Bayesian filtering and serve as location belief images. It can be seen that filtering is especially useful in indoor environments as it allows to solve ambiguities due to multipath.

TUTORIAL: ACCURACY BOUNDS FOR DF-L

The Cramer Rao Bound (CRB) provides a useful benchmark for assessing DF-L performances. The covariance matrix for any unbiased estimator $\hat{\mathbf{x}}$ of the target position \mathbf{x} is lower bounded as $\operatorname{Cov}(\hat{\mathbf{x}}) = \mathbb{E}[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T] \geq \mathbf{F}^{-1}(\mathbf{x})$, where $\mathbf{F}(\mathbf{x})$ is the 2×2 Fisher information matrix (FIM). According to the Gaussian model $\mathbf{s} \sim \mathcal{N}(\boldsymbol{\mu}(\mathbf{x}), \mathbf{Q}(\mathbf{x}))$ in (6), the FIM term (i, j) with $i, j \in \{1, 2\}$, is

$$F_{i,j} = [\mathbf{F}(\mathbf{x})]_{i,j} = \sum_{\ell=1}^L \frac{1}{\sigma_\ell^2(\mathbf{x})} \left[\frac{\partial \mu_\ell(\mathbf{x})}{\partial x_i} \frac{\partial \mu_\ell(\mathbf{x})}{\partial x_j} + 2 \frac{\partial \sigma_\ell(\mathbf{x})}{\partial x_i} \frac{\partial \sigma_\ell(\mathbf{x})}{\partial x_j} \right] \quad (8)$$

where the gradient functions $\frac{\partial \mu_\ell(\mathbf{x})}{\partial x_i} = \frac{\partial \Delta \mu_\ell(\mathbf{x})}{\partial x_i}$ and $\frac{\partial \sigma_\ell(\mathbf{x})}{\partial x_i} = \frac{\partial \Delta \sigma_\ell(\mathbf{x})}{\partial x_i}$, $i \in \{1,2\}$, embody the information on the target location provided by the perturbation maps of attenuation $\Delta \mu_\ell(\mathbf{x})$, and standard deviation $\Delta \sigma_\ell(\mathbf{x})$, respectively [11]. According to the diffraction model (5), and the related analytical maps $\Delta \mu_\ell(\mathbf{x}) = \mathbb{E}_\varphi [\Delta E_{\text{dB}}^2(\mathbf{x}; \ell; \varphi)]$ and $\Delta \sigma_\ell(\mathbf{x}) = \sqrt{\mathbb{E}_\varphi \left[\left(\Delta E_{\text{dB}}^2(\mathbf{x}; \ell; \varphi) - \Delta \mu_\ell(\mathbf{x}) \right)^2 \right]}$ from the tutorial “Diffraction-based modeling of human body shadowing”, by computing all derivative terms of (8), it is possible to obtain the $\mathbf{F}(\mathbf{x})$ matrix and, finally, the CRB for the complete L -link network. To demonstrate the effects of multiple links on the localization accuracy, the maps of Fig 6.b show the lower bound to the root mean square error of the location estimate $\text{RMSE} = \sqrt{\mathbb{E}[\|\hat{\mathbf{x}} - \mathbf{x}\|^2]} \geq \sqrt{\text{tr}[\mathbf{F}^{-1}(\mathbf{x})]}$ for $L = 1, 8, 16, 24$ links. As expected, the localization accuracy is higher for a target near the terminals and along the LOS paths. For targets located in-between, the reduced sensitivity could be counter-balanced by increasing the number of links [11].

6. HUMAN ACTIVITY AND GESTURE RECOGNITION (DF-AR)

Focusing on device-free activity recognition (DF-AR), active and passive systems (Fig. 2) can be further differentiated into systems exploiting base-band signal processing (e.g. using SDR devices) or RSS-related metrics for the prediction of surrounding activities. With respect to DF-L, DF-AR systems typically require a higher sampling frequency. Typical recommendations for optimal sampling frequencies in activity recognition are above 6 Hz but higher sampling potentially fosters good recognition accuracy [8][25] (see Fig. 7). Methods such as RTI and fingerprinting are too slow and thus not employed. Instead, either systems conditioned on characteristic signal patterns or machine learning techniques are frequently applied. Apart from RSS, movement-indicating features/profiles in frequency-domain (for instance Doppler shift) are exploited.

The main classes considered for DF-AR are the detection of basic whole-body activities, whole/half body gestures and human breathing detection. The achievable recognition accuracy for DF-AR is limited by the system class (active or passive), the CQI (baseband CSI or RSS), the radio technology, the sampling rate (6+ Hz) and time/frequency domain features. For basic activities such as

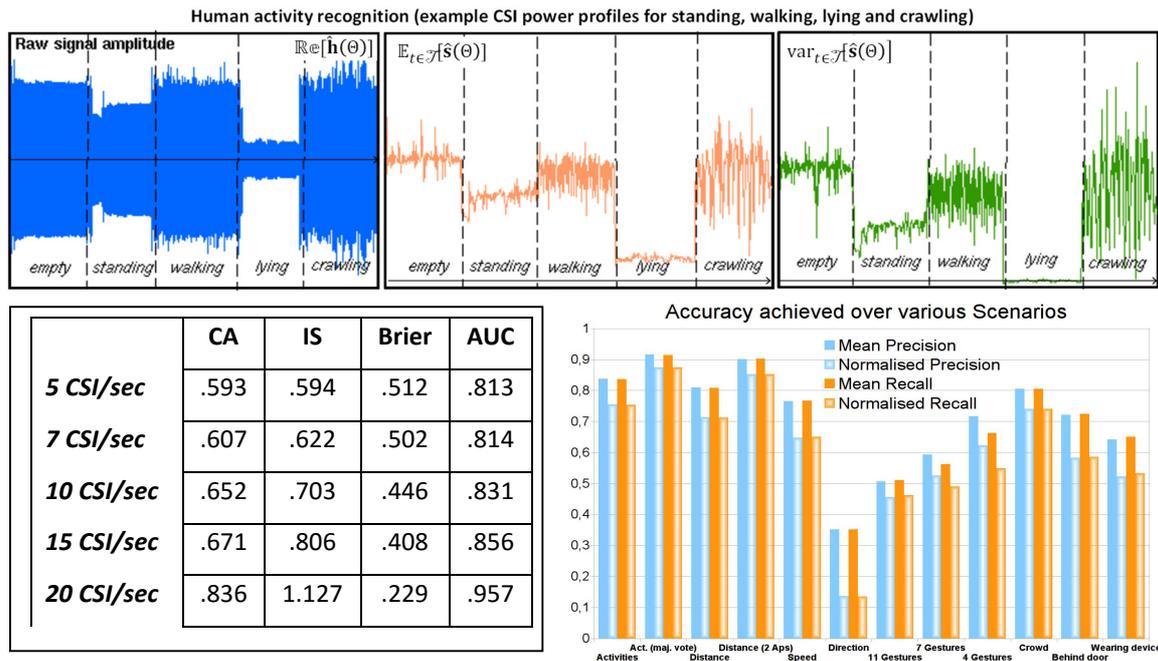


Figure 7: Impact of human body movements/activities: CSI power footprints for human standing, walking, lying and crawling with time domain features. Left bottom: Impact of different CSI sampling rates (CSI/sec) on the performance of a k-NN classifier [8]. Bottom right: classification accuracy from RSSI for movement in proximity, distance to receiver, walking speed, gestures (standing, walking, lying and crawling) and crowd size. In this example, RSSI from environmental 802.11 access points was captured by a single receiver (a smartphone) [25].

walking (at different speeds), crawling or standing, the recognition accuracies reached from CSI-based systems match those achievable with body-worn accelerometers [8]. Highly accurate, fine-grained part- and whole-body motions can be recognized via frequency-domain features like Doppler shifts from OFDM sub-channels (micro-Doppler fluctuation) [28]. However, RSS-based systems, are further limited in their achievable accuracy and recognized classes as depicted in Fig. 7: they thus qualify for applications in ambient assisted living but not e.g. for highly reliable systems.

BASIC WHOLE-BODY ACTIVITIES. Machine-learning approaches can be applied to extract information about the environmental situation from RSS fluctuations. Radio devices at frequencies between 900 MHz up to 5 GHz with arbitrary node deployments are typically utilized. Simple time-domain features are employed, such as root mean square (RMS), average magnitude squared (AMS), signal-to-noise ratio (SNR), signal amplitude, signal peaks and the number of large delta in successive peaks. The E-eyes system [12] combines WiFi 2.4 GHz links from different devices (e.g., access points, thermostats, laptops) and collects fine-grained CSI measurements as location-activity profiles. In [8], recognition accuracy has been improved by exploiting frequency domain features. Furthermore, the authors have compared DF-AR performance with accelerometer-based

recognition, showing comparable accuracy, in the order of 90-95% for indoor scenarios. Signals from nearby broadcasting FM radio stations qualify also for the detection of activities [26].

GESTURES. Simultaneous detection of gestures from multiple individuals can be obtained by utilizing multi-antenna nodes and leveraging micro Doppler fluctuations. Utilizing a MIMO-OFDM receiver, the WiSee system [28] can distinguish 9 pre-defined gestures from different people simultaneously with accuracy of 94%. In a related system (WiVi [27]), a single antenna receiver is used, while a preamble transmission stage is designed to isolate the time-varying reflections induced by the human body and null direct and wall-reflected disturbance. The system tracks the direction of the moving object using inverse synthetic aperture radar (ISAR): consecutive CSI measurements are collected over time to emulate an antenna array at the receiver. Gestures can be also detected by monitoring RSS at link-layer. In [25], simple classes of hand gestures can be recognized using off-the-shelf smart-phone devices (observing 10+ RSS samples per second). Although the achievable accuracy is lower than for the active systems previously discussed, a clear distinction of up to 11 hand gestures performed in proximity of the phone is possible.

RECOGNITION OF BREATHING. Contact-free RF respiration monitoring systems can detect the breathing rate of a single individual. Detection of breathing can be based on monitoring the RSS fluctuations from a pre-existing IEEE 802.15.4 network surrounding the subject [22]. Using ML estimation, an error of 0.3 breaths per minute is shown as achievable. The nodes transmitted every 240 ms at 2.48 GHz. Prediction is taken after a 10 second to 60 second measurement period. The UbiBreathe system [13] uses off-the-shelf WiFi devices and provides a reasonable breathing estimation accuracy even using a single point-to-point link. Other systems are based on microwave Doppler radars and ultra-wideband (UWB) (see [13] and references therein): they provide high accuracy but limited range and require an ad-hoc design and PHY layer optimization.

TUTORIAL: CQI BASED HUMAN ACTIVITY RECOGNITION

Base-band CSI based. CSI processing enables accurate activity recognition thanks to fine-grained frequency- and high time-resolution. Standard machine learning techniques (e.g., k-nearest neighbor, decision trees, support vector machines) can be applied to time-domain CSI features such as

mean $\mathbb{E}_{t \in \mathcal{T}}[\hat{\mathbf{S}}]$, variance $\text{var}_{t \in \mathcal{T}}[\hat{\mathbf{S}}]$ (Fig. 7), or CSI power footprint $\hat{\mathbf{S}} \in \mathbb{R}^{T \times 1} = \left[|\hat{h}_t(\Theta)|^2 \right]_{t \in \mathcal{T}}$. Elementary activities such as phone calls, opened/closed doors or human body standing, walking at different speeds [8], lying and crawling can be distinguished. These activities (or combination of) have to be trained separately by machine learning methods beforehand. For recognition of fine-grained activities or gestures, frequency-domain features, e.g., micro Doppler fluctuations, are required [28]. However, Doppler shifts caused by human motion are several magnitudes smaller than the signal bandwidth. Focusing on MIMO OFDM modulations, analysis of such fluctuations is possible after transforming the received symbols into narrowband pulses. Then, patterns from falling/rising signal-edges of Doppler fluctuations can be exploited for gesture and activity recognition.

RSS based. RSS-based passive systems measure noisy link-layer RSSI. In such settings, human activity can be detected using simple time-domain features (see Fig. 3a), including de-noising [13] and dynamic time-warping [12]. Although RSS samples are often bursty, a weak distinction between simple human activity classes is feasible [25] with about 10 RSS observations/sec while best accuracy is achieved with 40-80 RSS/sec where a precision and recall rate in the order of 90% for simple activities and 70% for gestures is possible. Further improvement is achieved by filtering noise, and focusing on the falling and rising edges of the composition of the signal features.

7. MONITORING OF HUMAN FALL: A CASE STUDY

The adoption of a device-free wireless fall detection technology is highly attractive in the context of assisted living as a fallen person might not be able to activate a personal emergency response system, if not forgetting how to use it. Today's commercially available products use already a broad range of active devices (e.g., necklaces with emergency buttons, fall sensors in mobile phones, etc.). However, these devices are often too difficult for elderly people to operate and are useless in emergencies [30]. Body-worn sensing devices also require cooperation from the monitored subjects and might hinder daily activities. Systems based on cameras, video or acoustic sources are also effective but penalized by privacy concerns. However, the proliferation of in-home wireless connected devices as part of the internet of things paradigm is acting as a boost to the development of

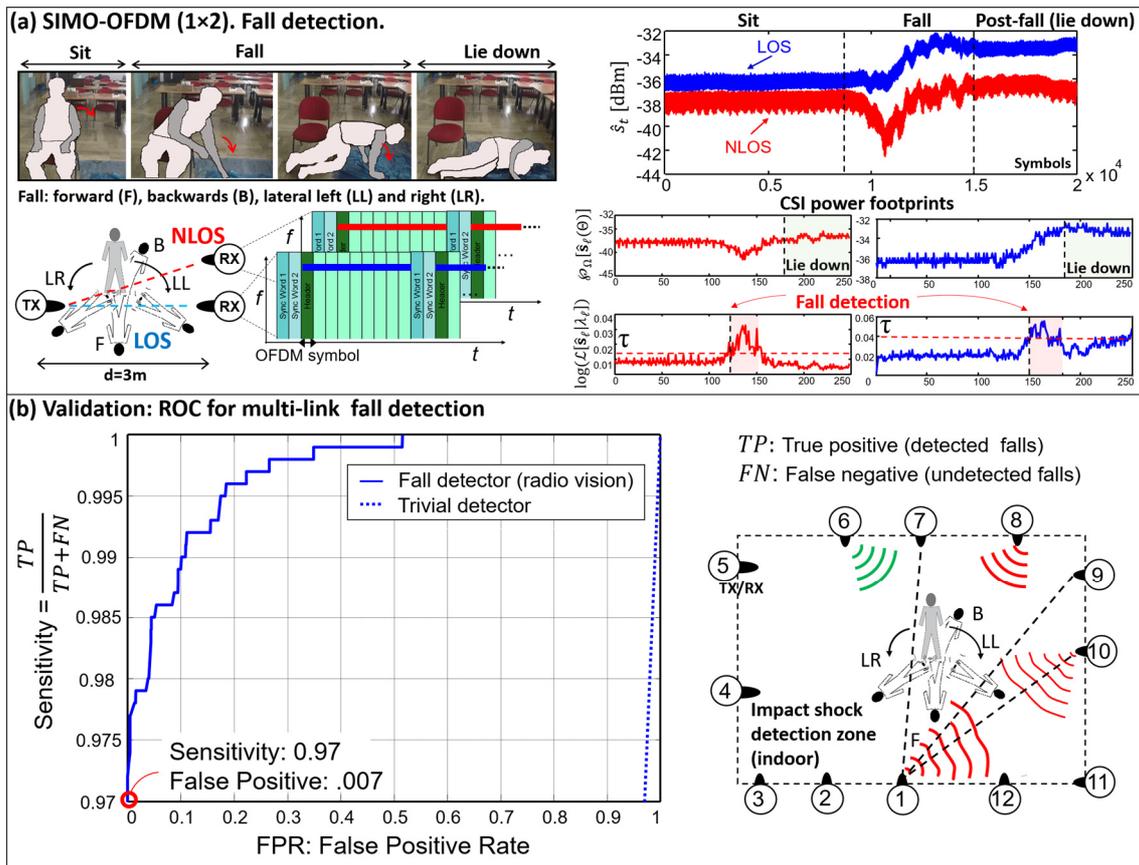


Figure 8.a: Hidden Markov model-based fall detection using SDR devices. SIMO-OFDM, (2 RX antennas, single RS subcarrier). **Figure 8.b:** impact shock detector based on CQI data (ROC curve).

new radio-based vision technologies. The possibility of monitoring human well-being by leveraging pre-existing indoor network infrastructures is becoming attractive in several applications.

Here, we highlight an experimental case study specifically focused on real-time processing of CQI for detection of the impact shock during body fall. The in-lab system of Fig. 8.a consists of a deployment of wireless devices exchanging data over 2.6 GHz bands using an OFDM radio front-end. A single antenna OFDM transmitter communicates with a receiver employing two antennas (with spacing of 24 cm). The receiver extracts and processes two, possibly incomplete, CSI power footprints \hat{s}_ℓ from the corresponding links (LOS and NLOS) $\ell \in \{\ell_1, \ell_2\}$. Body falling is monitored over a pre-defined position (\mathbf{x}): localization can be obtained by DF-L methods.

The observed sequences $\hat{\mathbf{s}}_\ell$ are modelled by a hidden Markov model (HMM) [23][24] with state space $\mathbf{Q}_\ell(\theta) \in \mathbb{R}^{Q \times 1} = [q_j]_{j=1}^Q$ containing Q selected values from the CQI profiles $\Delta \mathbf{s}_\ell(\theta)$ learned during training for falling-state estimation. HMM parameters, $\lambda_\ell(\theta) = [\mathbf{A}_\ell, \mathbf{B}_\ell, \boldsymbol{\pi}_\ell]$, include probabilities of state transition $[\mathbf{A}_\ell]_{i,j} = p[q_t = q_i | q_{t-1} = q_j]$, observation $[\mathbf{B}_\ell]_{i,j} =$

$p[\hat{s}_{\ell,t} = s_i | q_t = q_j]$, and initial state $[\boldsymbol{\pi}_\ell]_i = p[q_0 = q_i]$. The HMM parameters are learned by expectation maximization algorithm (e.g., Baum-Welch algorithm) and trained separately for each link [23]. Other methods [24] can be adopted to leverage space-time profiles correlation over co-located links. Decision about human fall is based on the model likelihood

$$\mathcal{L}[\hat{\mathbf{s}}_\ell | \lambda_\ell(\Theta)] = \sum_{\forall \mathbf{q} \in \mathbf{Q}_\ell(\Theta)} p[\hat{\mathbf{s}}_\ell, \mathbf{q} | \lambda_\ell(\Theta)] \quad (9)$$

with state sequence $\mathbf{q} = [q_1, \dots, q_T]$ and joint probability $p[\hat{\mathbf{s}}_\ell, \mathbf{q} | \lambda_\ell(\Theta)] = \prod_{t=1}^T p[\hat{s}_{\ell,t} | q_t] \cdot p[q_t | q_{t-1}]$. Functions $\mathcal{L}[\hat{\mathbf{s}}_\ell | \lambda_\ell]$ are continuously evaluated for both links (Fig. 8.a on the right).

Fall detection can be based on hard decision with respect to pre-calibrated threshold τ such that $\mathcal{L}[\hat{\mathbf{s}}_\ell | \lambda_\ell(\Theta)] / \mathcal{L}[\hat{\mathbf{s}}_\ell | \lambda_\ell(\emptyset)] > \tau$. Likelihood $\mathcal{L}[\hat{\mathbf{s}}_\ell | \lambda_\ell(\emptyset)]$ is obtained for HMM $\lambda_\ell(\emptyset)$ that considers arbitrary (but safe) body movements in the same position. After the impact shock is detected, a simple change detector can be applied to the observed CQI sequences for tracking any post-fall event, and in turn detect possibly long lie conditions, corresponding to negligible RF fluctuations.

IMPACT SHOCK DETECTION. In the complete case study highlighted in Fig. 8.b, human fall detector is now based on an optimized subset of pre-existing links [23] deployed around the subject of interest and selected during a calibration procedure (non-informative links are purged). Decision about the fall/non-fall event is based on majority voting over the optimized link subset.

Analysis of detector performance is crucial: undetected falls might have a dramatic impact, on the other hand, an excessive number of false activations might cause the detector to be perceived as useless. Validation of detector accuracy is thus illustrated in Fig. 8.b where the receiver operating characteristic (ROC) curve relates sensitivity versus false positive rate. A sensitivity of 0.97 and false positive rate of 0.007 compares well with performances of existing device-based systems [30].

8. CONCLUDING REMARKS AND FUTURE DIRECTIONS

This article focused on device-free radio vision systems acting as a flexible sensing tool and addressing key challenges in assisted living applications. The goal of this emerging research field is to develop models and processing methodologies for exploiting the inherent (but currently unused) sensing capabilities of the multitude of available wireless communication links, opening also to

investigate new radio technologies and unexplored bands. Future research on radio vision systems is expected to combine the use of localized RF signal inspection with large-scale and big-data processing. Running real-time analytics from massive volumes of RF data will pose new signal processing relevant problems, as well as the re-design of conventional statistical learning tools applied to unprecedented high-dimensional data structures.

REFERENCES

- [1] N. D. Lane, et al., A survey of mobile phone sensing. *IEEE Comm. Magazine*, pp. 140-150, Sept. 2010.
- [2] R.M. Buehrer, C.R. Anderson, R.K. Martin, N. Patwari, N., M.G. Rabbat, Introduction to the Special Issue on Non-Cooperative Localization Networks. *IEEE J. Sel. Topics in Signal Proc.*, Vol. 8, No. 1, Feb. 2014.
- [3] A. Laya, W. Kun, A. A. Widaa, J. Alonso-Zarate, J. Markendahl, L. Alonso, Device-to-device communications and small cells: enabling spectrum reuse for dense networks. *IEEE Wireless Communications*, Vol. 21, No. 4, pp. 98-105, Aug. 2014.
- [4] J. Lloret, A. Canovas, S. Sendra, L. Parra, A smart communication architecture for Ambient Assisted Living. *IEEE Comm. Magazine*, pp. 26-33, Jan. 2015.
- [5] M. Seifeldin, A. Saeed, et al., Nuzzer, A Large-Scale Device-Free Passive Localization System for Wireless Environments. *IEEE Trans. Mobile Computing*, Vol. 12, No. 7, pp.1321-1334, July 2013.
- [6] N. Patwari, J. Wilson, RF sensor networks for Device-Free localization: measurements, models and algorithms. *Proc. of the IEEE*, Vol. 98, No. 11, Nov. 2010.
- [7] K. Slavakis, G. B. Giannakis, G. Mateos, Modeling and Optimization for Big Data Analytics. *IEEE Signal Processing Magazine*, Sep. 2014.
- [8] S. Sigg, M. Scholz, et al., RF-Sensing of Activities from Non-Cooperative Subjects in Device-Free Recognition Systems Using Ambient and Local Signals. *IEEE Trans. on Mobile Computing*, Vol.13, No.4, pp. 907-920, Apr. 2014.
- [9] A. Saeed, A.E. Kosba, M. Youssef, Ichnaea, A Low-Overhead Robust WLAN Device-Free Passive Localization System. *IEEE J. Sel. Topics in Signal Proc.*, Vol.8, No.1, pp. 5-15, Feb. 2014.
- [10] K. Youngwook L. Hao, Human Activity Classification Based on Micro-Doppler Signatures Using a Support Vector Machine. *IEEE Trans. on Geosc. and Remote Sens.*, Vol. 47, No. 5, pp. 1328-1337, May 2009.
- [11] S. Savazzi, M. Nicoli, F. Carminati, M. Riva, A Bayesian approach to Device-Free Localization: modeling and experimental assessment. *IEEE J. Sel. Topics in Signal Proc.*, Vol. 8, No. 1, pp. 16-29, Feb. 2014.
- [12] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, H. Liu, E-eyes: Device-free Location-oriented Activity Identification Using Fine-grained WiFi Signatures. *Proc. of the 20th annual international conference on Mobile computing and networking*, pp. 617-628, 2014.
- [13] H Abdelnasser, M Youssef, K Harras, UbiBreathe: A Ubiquitous non-Invasive WiFi-based Breathing Estimator. *Proc. of Mobicom, Paris, France, 2015*.
- [14] M. Scholz, L. Kohout, M. Horne, M. Budde, M. Beigl, M. A. Youssef, Device-free radio-based low overhead identification of subject classes. *Proc. 2nd Workshop on Physical Analytics WPA'15*, pp. 1-6, 2015.
- [15] D.B. Smith, D. Miniutti, T.A. Lamahewa, L.W. Hanlen, Propagation Models for Body-Area Networks: A Survey and New Outlook. *IEEE Antennas and Propagation Mag.*, Vol. 55. No. 5, pp. 97-117, Oct. 2013.
- [16] V. Rampa, S. Savazzi, M. Nicoli, M. D'Amico, Physical modeling and performance bounds for device-free localization systems. *IEEE Signal Processing Letters*, vol. 22 no. 11, pp.1864-1868, November 2015.
- [17] K. Woyach, D. Puccinelli, M. Haenggi, Sensorless Sensing in Wireless Networks: Implementation and Measurements. *Proc. of 4th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks*, pp.1-8, 3-6 Apr. 2006.
- [18] M. Youssef, M. Mah, and A. Agrawala, Challenges: Device-free Passive Localization for Wireless Environments. *Proc. of the 13th annual ACM MobiCom 2007*.
- [19] J. Wilson, N. Patwari, Radio tomographic imaging with wireless networks. *IEEE Trans. on Mobile Comp.*, Vol. 9, No. 5, pp. 621-632, 2010.
- [20] J. Wilson and N. Patwari, See-Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks. *IEEE Transactions on Mobile Computing*, Vol. 10, No. 5, May 2011.

- [21] [Y. Mostofi, Cooperative wireless-based obstacle/object mapping and see-through capabilities in robotic networks. IEEE Trans. On Mobile Computing, Vol. 12, No. 5, May 2013.](#)
- [22] [N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, M. Bocca, Breathfinding: a wireless network that monitors and locates breathing in a home. IEEE J. Sel. Topics in Signal Proc., Vol. 8, No.1, pp. 30-42, Feb. 2014.](#)
- [23] [S. Kianoush, S. Savazzi, F. Vicentini, V. Rampa, M. Giussani, Leveraging RF signals for human sensing: fall detection and localization in human-machine shared workspaces. Proc. of IEEE International Conference on Industrial Informatics \(INDIN 2015\), Cambridge, U.K., July 2015.](#)
- [24] [B. Mager, N. Patwari, M. Bocca, Fall detection using RF sensor networks. Proc. of IEEE 24th Int. Symp. on Personal Indoor and Mobile Radio Communications \(PIMRC\), pp.3472,3476, 8-11 Sept. 2013.](#)
- [25] [S. Sigg, S. U. Blanke, G. Troster, The telepathic phone: frictionless activity recognition from WiFi-RSSI. Proc. of IEEE Int. Conf. on Pervasive Computing, pp. 148-155, 24-28 Mar. 2014.](#)
- [26] [S. Shi, S. Sigg, and Y. Ji, Monitoring of Attention from Ambient FM-radio Signals. IEEE Pervasive Computing, Special Issue Managing Attention in Pervasive Environments, Jan-March 2014.](#)
- [27] [F. Adib and D. Katabi. See through walls with Wi-Fi! Proc. of ACM SIGCOMM, 2013.](#)
- [28] [Q. Pu, et al, Whole-home gesture recognition using wireless signals. Proc. of ACM MobiCom, Miami, USA, Sept. 2013.](#)
- [29] [H. Hashemi, M. McGuire, et al., Measurements and modeling of temporal variations of the indoor radio propagation channel, IEEE Trans. on Vehicular Technology, vol.43, no.3, pp.733-737, Aug 1994.](#)
- [30] [Igual, Raul, Medrano, Carlos and Plaza, Inmaculada, Challenges, issues and trends in fall detection systems. Biomed. Eng. Online, Vol. 12, No. 66, pp. 1-66, 2013.](#)

AUTHORS

Stefano Savazzi (stefano.savazzi@ieiit.cnr.it) is Researcher at the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT) of the National Research Council of Italy (CNR). He received the Ph.D. degree (with honors) in information technology from Politecnico di Milano, in 2008. He was researcher at Uppsala University in 2005, University of California San Diego (UCSD) in 2008 and Forschungszentrum Telekommunikation Wien (FTW) in 2010. His main research interests include cooperative and cognitive wireless and sensor networks, device-free methods for RF vision and localization.

Stephan Sigg (stephan.sigg@cs.uni-goettingen.de) is with the Computer Networks group of Georg-August-University of Göttingen. Before, he was a researcher at TU Braunschweig and an academic guest in the Wearable computer Lab at ETH-Zurich, in the Nodes Laboratory at University of Helsinki and at the National Institute of Informatics (NII) in the information systems architecture research division. He obtained his PhD from University of Kassel. His research interest include the design, analysis and optimization of algorithms for ubiquitous systems, in particular for device-free passive activity recognition.

Monica Nicoli (monica.nicoli@polimi.it) is Assistant Professor at Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano. She received the PhD degree in communication engineering from Politecnico di Milano, in 2002. She was visiting researcher at Uppsala University in 2001. Her research interests are in the area of signal processing, with emphasis on wireless communications, distributed and cooperative systems, radio localization and intelligent transportation systems.

Vittorio Rampa (vittorio.rampa@ieiit.cnr.it) is Senior Researcher at the Institute of Electronics, Computer and Telecommunication Engineering (IEIIT) of the National Research Council of Italy (CNR). Since 1999 he has been Adjunct Professor at the Politecnico di Milano where he taught courses on Software Radio and Radio localization Systems. His main research interests are focused on signal processing algorithms and architectures for wireless communications, virtual instrumentation techniques for test and verification of wireless systems, radio localization algorithms and architectures for wireless sensor networks.

Sanaz Kianoush (sanaz@kianoush@ieiit.cnr.it) received her PhD Degree in Electronic Engineering from the University of Pavia, Italy in 2014 and currently postdoctoral researcher at Institute of Electronics, Computer and Telecommunication Engineering (IEIIT) of the National Research Council of Italy (CNR). Her research interests include Localization in Wireless Sensor and Cognitive Radio Networks, and Context-aware Activity Recognition.

Umberto Spagnolini (umberto.spagnolini@polimi.it) is Professor of Politecnico di Milano, and his interests are in statistical signal processing for communication systems and remote sensing. He is author of 250+ peer-reviewed papers and some patents, his areas of experience include channel estimation and space-time processing for wireless communication systems, cooperative and distributed systems, parameter estimation/tracking and wavefield interpolation for UWB radar, oil exploration and remote sensing. More info: <http://home.deib.polimi.it/spagnoli>