This is the post peer-review accepted manuscript of:

F. Valmori, A. Giorgetti, M. Mazzotti, E. Paolini and M. Chiani, "Indoor detection and tracking of human targets with UWB radar sensor networks," IEEE International Conference on Ubiquitous Wireless Broadband (ICUWB), Nanjing, 2016, pp. 1-4.

The published version is available online at:

https://doi.org/10.1109/ICUWB.2016.7790435

Indoor Detection and Tracking of Human Targets with UWB Radar Sensor Networks

Filippo Valmori, Andrea Giorgetti, Matteo Mazzotti, Enrico Paolini, and Marco Chiani

Abstract—In this paper, the whole signal processing chain for an ultra-wideband (UWB) radar sensor network (RSN) is presented, starting from real measurements collected by sensor devices, and ending with the estimation of the target trajectory. The RSN is composed by one transmitter and six receivers, and monitors an indoor area of about $70\,\mathrm{m}^2$ performing tracking of a human target. The proposed processing chain consists of: clutter removal, a novel detection scheme, one-dimensional clustering, trilateration, two-dimensional clustering, particle-based probability hypothesis density filter tracking, and data association. Numerical results show the remarkable performance of the system, resulting in a root mean square localization error of 36 cm, a value smaller than the target size. The presented experimental study show that it is possible to accurately track human targets using a UWB RSN in a densely cluttered environment.

I. Introduction

RADAR SENSOR NETWORKS (RSNs) for localization and tracking of human targets and objects in indoor environments have gained an increasing interest, especially for anti-intrusion applications. In the context of Industry 4.0, for example, indoor localization in the manufacturing sector will enable the application of ideas such as geofencing, to check whether people or vehicles are located in specific areas, such as in the proximity of machines. RSNs can also support the optimization of the logistics processes.

These networks often employ the ultra-wideband (UWB) technology, especially its impulse-radio version, characterized by the transmission of short pulses (with duration in the order of nanoseconds) [1]. The extremely large bandwidth of UWB systems possesses interesting features such as robustness to jamming, walls and objects penetration, low probability of interception, excellent multipath resolution, and accurate localization. The RSNs considered in this paper are UWB multistatic radar systems aimed at detecting and tracking non-cooperative targets (e.g., human subjects) moving inside a surveillance area.

Non-collaborative localization through multistatic UWB radars is the subject of several works, e.g., [1], [2]. However, in this field just a few works focus on experimentation in real environments and propose the analysis of the whole processing chain. To the authors knowledge, only [3] provides a proof-of-concept study to show that it is possible to accurately track

multiple targets using a UWB RSN. Most of the works focus on specific aspects of the radar system. For example, in the context of tracking, there are several works proposing non-linear Bayesian filtering techniques such as particle filters, and analyze their feasibility in terms of computational complexity [4]–[6].

In this paper, we propose a complete signal processing chain for UWB-based RSNs which consists of several processing steps, performed in part at the single sensor, alleviating the computational burden due the amount of data exchanged among sensors, and in part at the fusion centre (FC). The performance of the overall system is validated using waveforms collected in a real indoor environment with a human being as target, and thus constitute a very important benchmark to assess the capability of UWB technology to perform non-collaborative localization in the presence of dense clutter.

II. System Overview

The RSN is a mutistatic radar composed of one transmitter (TX) and at least three receivers (RXs), often placed along the perimeter of the monitored area to improve system efficiency and coverage [1]. During every scan the transmitter emits very short (on the order of a nanosecond) pulses; the RXs (or sensors) estimate the time-of-arrival (ToA) of direct and target-reflected paths, perform proper signal processing algorithms and finally forward the data to a FC which performs target detection, localization and tracking. Each RX, as it estimates the ToA t_0 of the target-reflected path, can locate the target on an ellipse having the transmitter and itself as foci and $c\,t_0/2$ as semi-major axis (where c is the speed of light). With at least three receivers it is possible to estimate the target position within the monitored area through the intersection of their ellipses. The proposed signal processing chain consists of tasks performed at the sensor level (clutter removal, constant false alarm rate (CFAR) detection, and 1D-clustering), and at the FC (trilateration, 2D-clustering, and tracking).

III. Single Receiver Processing

A. Clutter Removal

The first processing step consists of clutter removal, to cancel the undesired echoes from static objects. For this step, we adopt the empty room (ER) technique, which consists of subtracting from the currently received waveform a waveform collected when there are no targets within the monitored area (the ER scan). In particular, after extracting the envelopes, the ER and the current scan are up-sampled (by a factor of 10), aligned based on their cross-correlation, and subtracted.

The resulting waveform is expected to reveal possible target echoes, together with a clutter residue. Finally, a half-wave rectifier is applied to avoid negative values caused by a *ghost*-type problem, specifically a target obscuring a clutter echo which is present in the ER scan.

B. CFAR Detection with adaptive threshold

In the proposed processing chain, the next step is aimed at turning each soft value obtained after clutter removal into a binary value via comparison with a suitable threshold. In this paper an innovative CFAR algorithm, named double threshold and buffer (DTB), is proposed for adaptive threshold setting. DTB-CFAR, summarized in Algorithm 1, consists of calculating the threshold, η, with which the current sample under test (SUT) is compared, as

$$\eta = \alpha P_{\mathrm{L}} + \beta P_{\mathrm{G}} \tag{1}$$

where P_{L} and P_{G} are defined as the *local* and the *global* power respectively, whereas α and β are scaling factors. In particular, P_{G} is the average power of the whole scan, whereas P_{L} is the average power of a circular buffer filled with the samples already processed during the same scan that were found to be below threshold (this way we reduce the misdetection rate in case of multiple-target scenarios as we keep memory of the decisions previously taken during the scan). If the SUT turns out to be below its threshold, before being set to '0', it is saved in the buffer replacing the oldest sample. Otherwise the sample is set to '1' and the buffer is not updated.

Keeping the length of the buffer short, it is possible to get a local estimate of the noise and residual clutter power level for the near-SUT samples. However, since target presence does not typically manifest through an instantaneous transient, the overall threshold (due to P_{L}) increases very quickly when a waveform peak occurs, causing misdetections. To avoid such phenomenon, if the increment of the SUT with respect to the previous sample is larger than a predefined value, `ThrStepUp`, for at least `ThrCounter` consecutive samples, then P_{L} is not updated. Thus, since residual clutter and target presence is characterized by similar waveform peaks with the only difference that the latter ones are generally significantly larger than the former ones, the overall threshold is expected to increase quickly enough to avoid false alarms but also be frozen soon enough to properly detect the target. Later, P_{L} is unfrozen when a waveform decrease occurs. To further decrease misdetections, because of the asymmetric target waveform, DTB-CFAR is applied to the scan twice in opposite directions. Then, the two outputs are merged through an element-wise OR operator to form the final vector.

We have experimentally verified that this detection algorithm far improves processing results compared to other conventional approaches (e.g., cell-averaging CFAR) in case of multiple human targets in densely cluttered environments.

C. 1D-Clustering and weight-based thresholding

After CFAR detection, scans are generally characterized by bursts of '1's related to the same target or (to a lesser extent) to residual clutter.

Algorithm 1 DTB-CFAR

$\text{SCANin} \leftarrow flip\_and\_concatenate(\text{SCANin})$
$\text{P}_{\text{G}} \leftarrow get\_averagepower(\text{SCANin})\,; \text{ Freeze} \leftarrow 0$
for $\text{i} = 1, \ldots, length(\text{SCANin})$ **do**
   if $\text{Freeze} = 0$ **then**
      $\text{P}_{\text{L}} \leftarrow \text{P}_{\text{L\_updated}}$
   else
      $\text{P}_{\text{L}} \leftarrow \text{P}_{\text{L\_frozen}}$
   end if
   $\eta \leftarrow \alpha \cdot \text{P}_{\text{L}} + \beta \cdot \text{P}_{\text{G}}$
   if $\text{SCANin[i]}^2 > \eta$ **then**
      $\text{SCANout[i]} = 1$
   else
      $\text{BUFFER[i\_oldest]} \leftarrow \text{SCANin[i]}$
      $\text{P}_{\text{L\_updated}} \leftarrow get\_averagepower(\text{BUFFER})$
      $update(\text{i\_oldest})\,; \text{ SCANout[i]} \leftarrow 0$
   end if
   if $\text{SCANin[i]}^2 > \text{SCANin[i}-1]^2 + \text{ThrStepUp}$ **then**
      $\text{Counter} \leftarrow \text{Counter} + 1$
   else if $\text{SCANin[i]}^2 + \frac{\text{ThrStepUp}}{2} < \text{SCANin[i}-1]^2$ **then**
      $\text{Counter} \leftarrow 0$
   end if
   if $\text{Counter} > \text{Thr\_Counter} \,\&\, \text{Freeze} = 0$ **then**
      $\text{Freeze} \leftarrow 1\,; \text{ P}_{\text{L\_frozen}} \leftarrow \text{P}_{\text{L\_updated}}$
   else if $\text{Freeze} = 1 \,\&\, \text{P}_{\text{L\_updated}} < \text{P}_{\text{L\_frozen}}$ **then**
      $\text{Freeze} \leftarrow 0\,; \text{ Counter} \leftarrow 0$
   end if
end for

In fact, close '1's, corresponding to comparable times of arrival, are most likely due to the same pulse which has a duration of nanoseconds (i.e., several consecutive samples), while isolated '1's are due to noise and residual clutter. Therefore, to decrease false alarms and to limit the computational complexity, a mono-dimensional (1D) clustering is applied to the binary scans. Since the number of final clusters (i.e., number of targets) is not known a-priori, a hierarchical 1D-clustering approach is considered. At the beginning, each '1' is associated with a stand-alone cluster. Then, iteratively until the minimum distance[1] between two clusters is less than a pre-set distance-threshold, those clusters are merged. At the end, each survived cluster is replaced by a single '1' placed in the median position. Furthermore, a weight equal to the number of clusterized ones is associated with every survived cluster. As the clusters associated with targets typically exhibit higher weights than those related to clutter or noise, it is useful to set also a weight-based threshold so that the lightest survived clusters (having a weight less than this threshold), are neglected and the corresponding output sample is reset to '0' to further reduce false alarms.

---

[1]We choose the minimum distance among two clusters as the distance between their centroids (i.e., the median positions among the '1's which compose those clusters at the current iteration).

IV. Fusion Centre Processing

The binary vectors resulting from 1D-clustering performed by each sensor, are sent to the FC to be further processed.

A. Trilateration and measures reliability

The trilateration is aimed at collecting and processing the post-clustering data incoming from all sensors to locate targets in a 2D space. To do so, we adopted a least-squares (LS) approach applied to all possible triplets of '1's, each associated with a different RX [7]. When the number of RXs, N_R is larger than 3, even in the ideal situation where noise and residual clutter are absent (i.e., each sensor output is composed of an all-'0' vector apart from a '1' for each target in its actual ToA), this step causes false alarms arising when three times of arrival not associated with the same target are combined. For example, when there are N_t targets within the monitored area, the number of false alarms due to LS trilateration is

$$N_{\text{fa}} = \frac{N_R!}{3!(N_R - 3)!}\left(N_t^3 - N_t\right). \quad (2)$$

Therefore, criteria are needed to understand whether the several 2D points obtained from trilateration (referred to as the *measures*) correspond to real targets.

Firstly, all points falling outside the monitored area are discarded. Moreover, a weight (calculated as the sum of the post-clustering weights of the three ones combined to get that measure) is associated with each measure; since points related to real targets are expected to have higher weights than measures contributed by noise and clutter, applying a threshold on the combined weights allows to reduce the false alarms. Finally, a further filtering based on least square error (LSE) estimation has been introduced. The LSE represents an unreliability factor associated with each measure. In fact, when three '1's related to different targets (or to false alarms) are combined, ellipses (related to the TX-RXs pairs) do not intersects in one point, even in the absence of clutter and noise, resulting in a large LSE. Hence, employing an additional threshold on the LSE, it is possible to further filter out the most unreliable measures.

B. 2D-Clustering

The measures generated by the previous steps tend to form clouds of points inside the monitored area (especially if the RSN has more than three receivers), so a bi-dimensional (2D) clustering is used to merge together the observations referring to the same target. The approach is still hierarchical, iterative, centroid-based and employs again both distance and weight thresholds, as 1D-clustering in Section III-C, but considering Euclidean distances over 2D-space. The weight-based threshold is particularly effective because the residual false alarms after trilateration are randomly placed within the monitored area in a sparse manner.

C. PHD tracking and data association

The final step of the processing chain is tracking, composed by the cascade of a particle-based probability hypothesis density (PHD) filter and a data association (DA) block.

The PHD filter combines the post-clustering measures and the past states of the targets in order to calculate their most likely position at the current scan [8]. The PHD filter formulation involves nonlinear equations not admitting closed-form solutions, regardless the kinematic model of the targets, so a particle filter implementation is adopted to discretize the PHD and limit the computational complexity [6].

At each scan the PHD filter generates a cloud of N_p particles around each new received measure and also around the most recent points of the existing tracks. The particle positions for the new measures are randomly generated according to a circularly symmetric bivariate Gaussian distribution with variance σ_p^2. The velocities associated with these particles are generated randomly according to a uniform distribution $\mathcal{U}[-v_{\max}, v_{\max}]$ (independently for both X and Y components), where $v_{\max}$ is the maximum speed of targets. The particles associated to recent points of the existing tracks are generated following the chosen targets dynamic model (here linear Gaussian). Moreover, a constant weight is initially associated with every particle. Then, the weights are updated to let the most important particles, which best represent the new received measures, to be heavier. This way a discrete representation of the PHD can be obtained and targets states information can be extracted through its peaks. The PHD peaks estimation is managed by an expectation-maximization (EM) algorithm, which iteratively approximates the PHD through a linear combination of Gaussian functions. Finally the particles are resampled using a Monte Carlo approach in order to discard the particles with lowest probability and keep their number constant during scans.

After the tracking stage, the data association block combines the PHD peaks with the existing tracks through the evaluation of suitable association costs and manages other possible cases (such as the initialization or death of tracks after a fixed number of scans) which occur when the number of input measures does not correspond to the number of current tracks. The implemented algorithm is based on a *0-scan* approach and, in order to find the best measure-track associations, all possible permutations are tested [9]. In particular, the cost of a single measure-to-track association at the current scan is calculated combining the information about position and velocity of the existing tracks (as estimated at the previous scan) with the current measures. So if the minimum cost hypothesis (composed by the sum of all its single associations) is below an input cost-threshold, those combinations are applied; otherwise each measure-track association is singularly checked comparing its cost to a gating-threshold to be confirmed.

V. Experimental Results

The measurement campaign has been carried out for a single target scenario in an $11.5\,\text{m} \times 6\,\text{m}$ room, with a RSN composed by one TX and 6 RXs. The sensor nodes are Time Domain's *PulsOn 410*, used to generate and collect UWB waveforms. The sampling rate is $16.4\,\text{GHz}$, the scan duration is of $100\,\text{ms}$, the overall number of scans collected along the trajectory is 61, and the human target (of approximately $90\,\text{kg}$ weight and
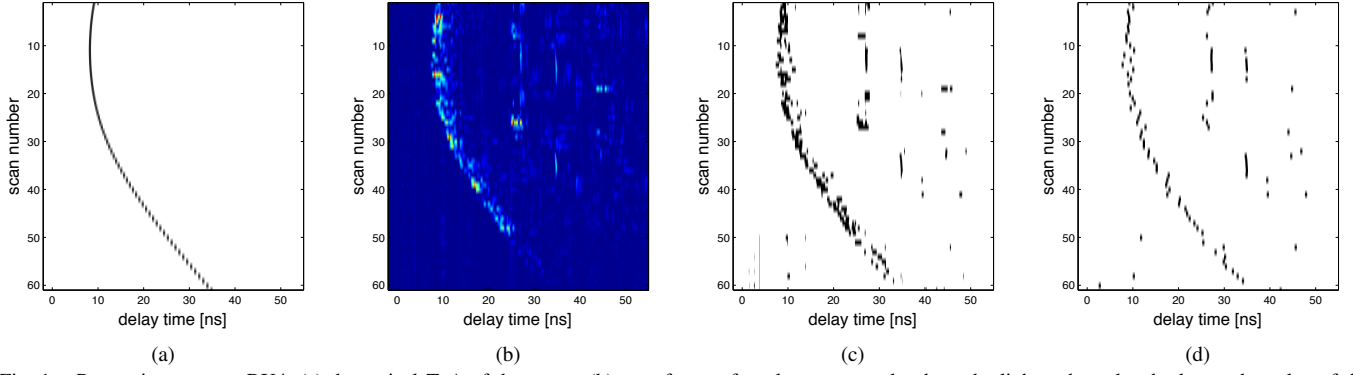
Fig. 1. Processing steps at RX4: (a) theoretical ToA of the target; (b) waveforms after clutter removal, where the lighter the color the larger the value of the corresponding sample; (c) binary waveforms after DTB-CFAR detection; (d) binary waveforms after 1D-clustering.
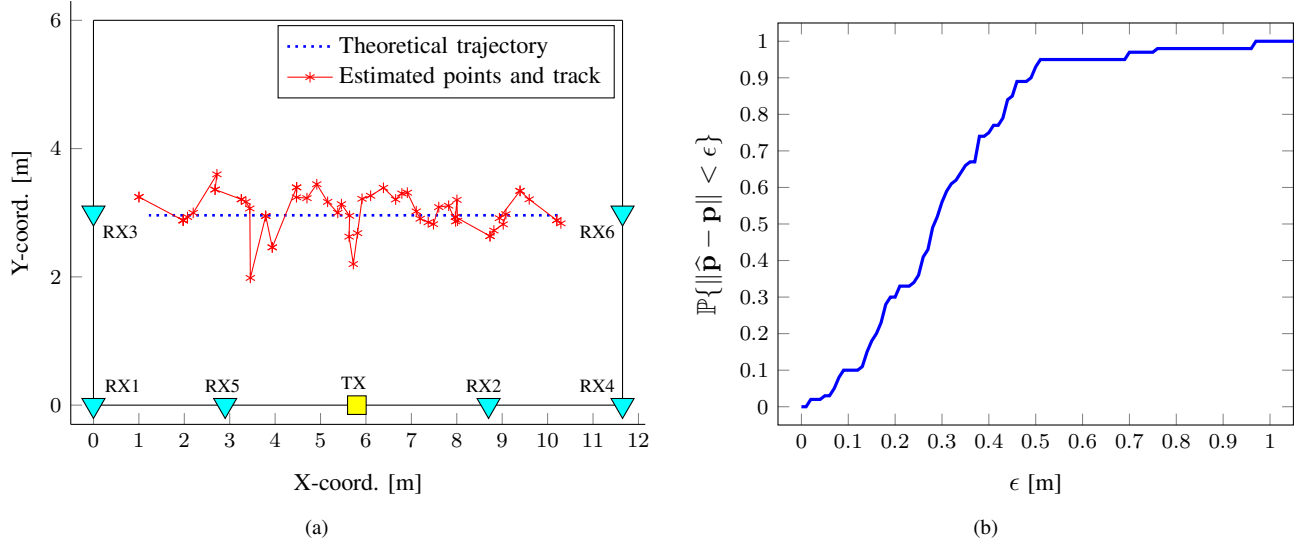

Fig. 2. Experimental results: (a) RSN configuration and the estimated track; (b) CDF of the localization error $\|\widehat{\mathbf{p}} - \mathbf{p}\|$.

185 cm height) moves with uniform rectilinear motion from right to left along his trajectory, with velocity $v_{\text{target}} = 1.5 \, \text{m/s}$.

Fig. 1 shows the processing results up to the 1D-clustering phase for the receiver RX4. In every subfigure the X-axis reports the difference (in nanoseconds) between the target-reflected path ToA and the direct path ToA as a function of the scan number on the Y-axis. Fig. 2(a), depicts the RSN configuration and the final points and track estimated after the whole processing at the FC. Defining the localization error as the distance between the true and the measured position of the target (respectively, $\mathbf{p}$ and $\widehat{\mathbf{p}}$), the points of the estimated track are characterized by a root mean square error (RMSE) of 36 cm.[2] A more detailed behavior of the localization error is represented by the cumulative density function (CDF) shown in Fig. 2(b).

REFERENCES

[1] E. Paolini, A. Giorgetti, M. Chiani, R. Minutolo, and M. Montanari, "Localization capability of cooperative anti-intruder radar systems," *EURASIP J. Advances Signal Process.*, vol. 2008, pp. 1–14, 2008.

[2] D. Hong, C. Chen, P. Shirui, L. Xin, and Z. Linhua, "Multistatic ultra-wideband localization for NLOS environments," in *Proc. Int. Conf. Intelligent Sys. Design Eng. App.*, Jan. 2012, pp. 380–384.

[3] B. Gulmezoglu, M. B. Guldogan, and S. Gezici, "Multiperson tracking with a network of ultrawideband radar sensors based on Gaussian mixture PHD filters," *IEEE Sensors J.*, vol. 15, no. 4, pp. 2227–2237, April 2015.

[4] T. Kai, W. Shiyou, C. Jie, L. Xiaojun, and F. Guangyou, "An efficient and low-complexity through wall moving target tracking algorithm by UWB radar," in *Proc. Int. Conf. Ground Pen. Radar*, Jun. 2012, pp. 966–971.

[5] M. Chiani, A. Giorgetti, M. Mazzotti, R. Minutolo, and E. Paolini, "Target detection metrics and tracking for UWB radar sensor networks," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Sep. 2009, pp. 469–474.

[6] B. Sobhani, E. Paolini, A. Giorgetti, M. Mazzotti, and M. Chiani, "Target tracking for UWB multistatic radar sensor networks," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 125–136, 2014.

[7] M. Švecová, D. Kocur, and R. Zetik, "Object localization using round trip propagation time measurements," in *Proc. Int. Conf. Radioelektronika*, April 2008, pp. 1–4.

[8] O. Erdinc, P. Willett, and Y. Bar-Shalom, "Probability hypothesis density filter for multitarget multisensor tracking," in *Proc. Int. Conf. on Inf. Fusion*, vol. 1, July 2005, pp. 1–8.

[9] K. Panta, B. n. Vo, and S. Singh, "Novel data association schemes for the probability hypothesis density filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 2, pp. 556–570, April 2007.

[2]It is worth noting that such error is much less than the human target size, hence it is significantly small.