A Gaussian Process model for UAV localization using millimetre wave radar

José A. Paredes[a,*], Fernando J. Álvarez[a], Miles Hansard[b], Khalid Z. Rajab[b,c]

[a]*Sensory System Research Group (GISS). University of Extremadura, Av. de Elvas s/n, 06006 Badajoz, Spain.*
[b]*School of EECS, Queen Mary University of London, Mile End Road, London E1 4NS, UK.*
[c]*NodeNs Medical Ltd., 10 Bloomsbury Way, London WC1A 2SL, UK.*

Abstract

The detection and positioning of unmanned aerial vehicles has become essential for both automation and surveillance tasks, in recent years. The design of accurate drone localization systems is challenging, especially in cluttered environments, where the target may be partially or even completely obscured. This paper proposes a precise detection and 3D localization system for drones, by means of a millimetre wave radar. Drone locations are estimated from spatial heatmaps of the received radar signals, which are obtained by applying the super-resolution MUSIC algorithm. These estimates are improved by analysis of the micro-Doppler effect, generated by the rotating propellers, which aids detection in poor visibility conditions. A novel Gaussian Process Regression model is developed, in order to compensate for systematic biases in the radar data. The complete system produces accurate estimates of the target range and direction, and is shown to outperform direct spectral analysis methods.

Keywords: positioning system, drone, radar, millimetre wave

*Corresponding author
Email addresses: japaredesm@unex.es (José A. Paredes), fafranco@unex.es (Fernando J. Álvarez), miles.hansard@qmul.ac.uk (Miles Hansard), khalid@nodens.eu (Khalid Z. Rajab)

1. Introduction

The development of reliable drone detection systems has become important, owing to the routine use of UAVs (unmanned aerial vehicles) in daily life. In recent years, these devices have been used for surveillance (Boddhu et al., 2013), logistical (Kellermann et al., 2020), audiovisual (Harvard et al., 2020) and military (de Swarte et al., 2019) tasks. All of these applications require a positioning system of some kind, which provides input to the process for both detection and navigation purposes. However, most of the research on this topic is not primarily concerned with establishing the precise absolute position of the drone with respect to a previously known location, but rather with general detection and tracking. For instance, one of the greatest concerns so far has had to do with the security risks posed by these devices, as discussed in Samland et al. (2012). In particular, much research has focused on detecting and tracking potentially hostile devices, as described in Guvenc et al. (2017, 2018) where the authors present a system consisting of a ground radar and a monitoring drone fleet with the final objective of interdicting malicious UAVs.

Another well-known goal is autonomous guidance, given that the device may travel beyond the communication range of the operator (Kendoul, 2012). In this regard, the basic problem of global positioning has already been solved by the inclusion of a GNSS (Global Navigation Satellite System) signal receiver for the most advanced drones, as can be seen in the features of several current off-the-shelf devices, such as Phantom (DJI, 2020), or in works such as Tahar & Kamarudin (2016), in which some corrections are applied to the data delivered by the GNSS chip, based on a prior statistical study of the noise and bias in these measurements.

Localization is still problematic in GNSS-denied environments, and in local contexts where an accurate position could be mandatory. The applications developed for these purposes can be tackled either in a dynamic way, as presented above, where other drones carrying different sensors provide assistance to the central system, or directly from the ground, i.e. from a known fixed location,

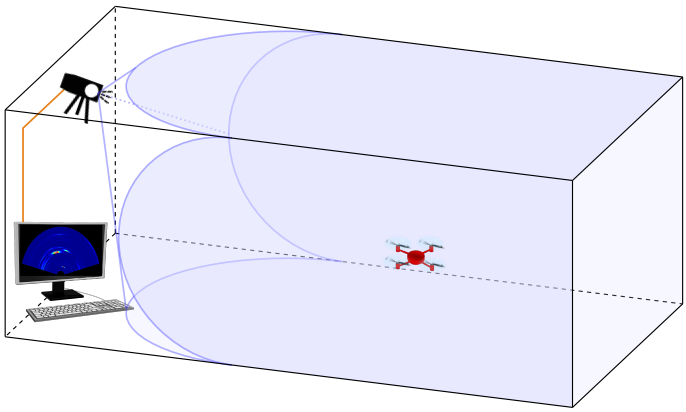

Figure 1: Overview of the proposed system. The radar is positioned in order to cover as much flying space as possible. In this example, it is placed near the ceiling and slightly tilted towards the ground, in order to increase the RCS (radar cross-section) of the drone.

in which the sensor is installed.

The second strategy requires neither an extensive infrastructure, nor the calibration of dynamic reference systems. Therefore, this paper proposes an accurate detection and 3D localization system for drones, designed for cluttered environments, using only a single commodity mmWave (millimetre wave) radar, in a fixed location. The general setup is illustrated in figure 1. The device uses short-wavelength electromagnetic signals, which enables detection through a variety of materials, even in optically challenging environmental conditions, such as darkness, fog, or smoke. These sensors are lightweight, and can easily be mounted inside a plastic car bumper, a security camera housing, or a robotic or drone housing (Ferguson et al., 2018). In general, mmWave technology permits the use of very small antennas, enabling integration within the sensing unit, and greater system miniaturization. These characteristics make mmWave technology appropriate for use in constrained environments, such as warehouses or urban delivery areas (Alwateer & Loke, 2020).

The contributions of this paper are as follows. A new way to enhance the detection of drones in mmWave radar data is presented in section 3.1. Section 3.2 develops new univariate and multivariate GP (Gaussian process) regres-

sion models, for accurate range and direction estimation. Section 3.3 proposes a spline-based dimension reduction method, to be used in conjunction with the GP regression. A combined mmWave radar and camera system, which allows ground-truth optical measurements to be recorded, is described in section 4.1. Our basic localization algorithms are evaluated in section 4.2. The more sophisticated GP regression models are evaluated in section 4.3. We report the final configuration of an effective UAV system, based on our experiments.

The organization of this paper is as follows. The relevant drone detection and localization systems are reviewed in section 2. Section Appendix A describes the main characteristics of the mmWave radar system, and the initial signal processing tasks. The proposed positioning algorithm is developed in section 3. The complete system is evaluated, and the results are analysed, in section 4.

2. Drone localization methods

A wide range of different technologies have been applied for detection and localization of drones. This section presents a brief review of the topic, including acoustic, optical, and radar modalities.

2.1. Acoustic methods

The first technology to be considered is that of acoustic analysis. Several works have specialized in detection of the characteristic propeller 'buzz' (Benyamin & Goldman, 2014; Busset et al., 2015). However, the applicability of this approach is limited by the cost of a complex microphone matrix, and by its high susceptibility to adverse weather conditions (e.g. wind or rain). Other solutions involve the fusion of different sensors, as proposed in Paredes et al. (2017), where an ultrasonic device estimates 2D position and a ToF (time of flight) camera provides the flying altitude.

2.2. Optical methods

Optical systems use one or more cameras to detect drones, based on methods from computer vision. This approach is one of the oldest, as described in Amidi

et al. (1999), in which a camera looking at the ground is mounted in a drone. The position is estimated by detecting known visual markers. Somewhat later, new marker-based systems emerged, for use with MAVs (micro aerial vehicles) (Masselli & Zell, 2012; Santana et al., 2014; Jiménez Lugo & Zell, 2014; Boudjit & Larbes, 2015; Vidal et al., 2017). A more recent work presents an accurate detection and tracking system for MAVs by means of a single ToF camera (Paredes et al., 2020). Nonetheless, while camera-based systems produce high resolution images across a wide FOV (field-of-view), detection relies on measuring reflected light. This makes them very susceptible to occlusions (by objects crossing the line of sight) and to scattering (by airborne particles such as fog or smoke).

2.3. Radar methods

Radar systems have been used to avoid the drawbacks of acoustic and optical technologies. The main difference from the former is in the signal that is used: RF (radio frequency) waves. Due to its wavelength and beam spread, a radar can return multiple readings from the same transmission, and generate a spatial representation of the environment. By capturing the reflected signal, a radar system can determine range, velocity and direction. An example can be found in Multerer et al. (2017), in which the authors propose the use of a 3D-type radar to track drones. To differentiate between dangerous and non-dangerous targets, they have their movements evaluated in real time and, if the algorithm determines it, a directional antenna generates a WiFi signal that jams the control of the device.

As mmWave-hardware has improved, more complex tasks have become feasible for these devices. The operation of a 2D detection and tracking system for UAVs, using mmWave, is explored in de Haag et al. (2016), where a comparison with a LiDAR (light detection and ranging) system can be found. The authors are able to detect drones at a distance up to 25m, although clutter at less than 5m prevents a proper detection. Besides, when a UAV is intended to be detected at low grazing angles, the multipath effect due to land clutter could mask the direct reflections from the device, as explained in Ezuma et al. (2019).

A detailed model of this effect is presented in this work, where the simulations show how the optimal detection performance of mmWave radar depends on the RCS of the targeted MAV, the radar properties, and the properties of the land clutter. Other works, like Dogru et al. (2019); Dogru & Marques (2020), focus on pursuing one drone with another one. The authors make use of a 2D radar, mounted in the follower, in order to detect the target drone. The missing axis information is extracted through geometric calculations, by estimating the first drone position.

In order to detect and identify UAVs via a radar device, it is often useful to take into account the micro-Doppler effect. A radar target will exhibit a Doppler frequency shift associated with its bulk motion (i.e. the velocity associate with a drone's flight path), but will also experience micro-motion dynamics – such as mechanical vibrations or rotations – which induce micro-Doppler modulations on the returned signal (Chen et al., 2006). The unique micro-Doppler signatures of a radar target can provide characteristic information about its properties and actions. This has been used to improve drone detection, such as in Caris et al. (2016) where two possible approaches for perimeter surveillance are presented. They cover distances from 10m to several hundreds meters, and they assert that the velocity (modulus and direction) can be estimated by analysing the micro-Doppler effect. This effect has been also applied for classification of drones (Fioranelli et al., 2015; Tahmoush, 2015; Jian et al., 2017; Nanzer & Chen, 2017; Sun et al., 2018), because the Doppler signature can be considered unique for every target, and to discriminate them from other potential flying targets, like birds, owing to the propeller rotation signal (Rahman & Robertson, 2018; Ezuma et al., 2019).

None of these works have addressed a precise drone detection and positioning in cluttered environments, where the clutter (including other non-drone targets) can make the detection unreliable or even impossible. These considerations are important, owing to the increasing use of drones in complex environments, such us warehouses and urban delivery areas. Hence, this paper proposes an accurate 3D localization system for UAVs in these scenarios, based on a novel approach

that applies Gaussian Process Regression (GPR) to the radar information, after improving target detection via the micro-Doppler effect generated by the propellers. Wang et al. (2018) took a related approach to *direction of arrival* estimation, from radar data, using Support Vector Regression (SVR). Our perspective is similar, but we obtain full 3D localization, and we use GPR rather than SVR. We argue that the GPR method has significant advantages, because it provides a probabilistic model of the data, including confidence intervals for prediction.

3. Proposed localization algorithm

The proposed mmWave radar localization system uses the following data, which is provided by the device: a 3D range-frame array $\mathbf{Q}$, in conjunction with a 2D range-Doppler array $\mathbf{Q}_D$. A non-negative 3D array $\mathbf{P}$, representing the signal intensity at any point $[r, \theta, \phi]$, can be obtained directly from the magnitude $|\mathrm{FFT}(\mathbf{Q})|$. In this work, however, we use the MUSIC (multiple signal classification) algorithm, which involves the appropriate steering matrix $\mathbf{A}$, for the antenna layout (see section 4.1). Hence the intensity array is computed as

$$\mathbf{P} \leftarrow |\mathrm{MUSIC}(\mathbf{Q}, \mathbf{A})| \,. \tag{1}$$

The basic principles of radar signal processing (including FFT and MUSIC methods) are summarized in Appendix A.

If one of the angular coordinates is fixed, then the resulting 2D slice through $\mathbf{P}$ corresponds to plane through the centre of the antenna, as indicated in figure 2. In the case of the Doppler array, it suffices to compute the absolute value, in order to obtain a range-velocity heatmap:

$$\mathbf{P}_D \leftarrow |\mathbf{Q}_D| \tag{2}$$

The next step in the proposed localization algorithm is to perform pointwise

background subtraction

$$\mathbf{P}' = \mathbf{P} - \mathrm{E}[\mathbf{P}]_{\mathcal{S}} \tag{3}$$

$$\mathbf{P}'_D = \mathbf{P}_D - \mathrm{E}[\mathbf{P}_D]_{\mathcal{S}} \tag{4}$$

where $\mathbf{P}'$ and $\mathbf{P}'_D$ are the signals relative to the corresponding temporal averages, denoted by $\mathrm{E}[\cdot]_{\mathcal{S}}$. The latter are taken over a reference time interval $\mathcal{S} = [s_0, s_1]$, during which the target is absent, and the scene is static. This step reduces the effects of environmental clutter, by removing the background component from the signal.

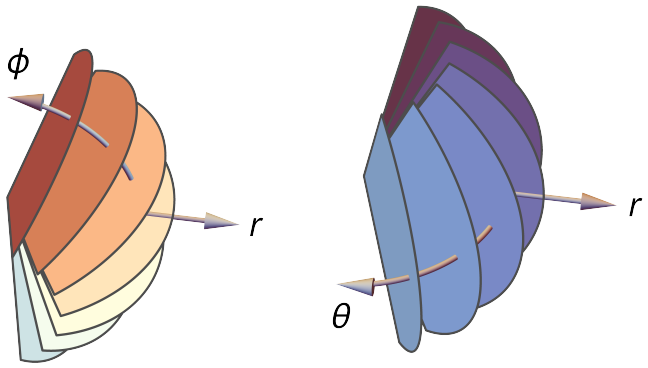

Figure 2: Left: elevation planes through the original data array $P(r, \theta, \phi)$. A single azimuth map $P'_A(r, \theta)$ is obtained by maximizing over ϕ, for each pair of coordinates (r, θ), as indicated (5). Right: analogous azimuthal planes. A single elevation map $P'_E(r, \phi)$ is obtained by maximizing over θ, for each pair of coordinates (r, ϕ), as indicated (6). The angular ranges correspond to the antenna array limits of $\pm 60°$, as specified by the radiation pattern's main beam (while the coarse angular sampling is for illustration only).

For clarity of exposition, we now regard the 3D array $\mathbf{P}'$ and the 2D array $\mathbf{P}'_D$, from (3) and (4), as non-negative scalar functions $P'(r, \theta, \phi)$ and $P'_D(r, v)$, respectively. This notation also allows us to consider, for example, bilinear interpolation of the data at non-integer coordinates (although we do not require it here). The 2D azimuth and elevation maps $P'_A(r, \theta)$ and $P'_E(r, \phi)$, in the analogous notation, can be obtained by taking the maximum over one angular

dimension, in each case:

$$P'_A(r,\theta) = \max_{\phi} P'(r,\theta,\phi), \quad \phi \in [\phi_{\min}, \phi_{\max}], \tag{5}$$

$$P'_E(r,\phi) = \max_{\theta} P'(r,\theta,\phi), \quad \theta \in [\theta_{\min}, \theta_{\max}]. \tag{6}$$

The resulting 2D maps are separately normalized, by the maximum value, in each case. A similar procedure is used to obtain a 1D function of the scene range:

$$P'(r) = \begin{cases} \max_{\theta} P'_A(r,\theta), & \theta \in [\theta_{\min}, \theta_{\max}], \\ \max_{\phi} P'_E(r,\phi), & \phi \in [\phi_{\min}, \phi_{\max}]. \end{cases} \tag{7}$$

Note that the two possibilities are equivalent: in both cases, $P'(r)$ is the maximum over the same spherical wedge of directions (the only difference is the ordering of the angular max operations).

3.1. Direct maximum estimation

The actual range of the target can be tentatively estimated from the maximum value of $P'(r)$. However, the detection can be strengthened by considering the $\mathbf{P}_D$ heatmap. When, in addition to the constant Doppler frequency shift induced by the bulk motion of a radar target, the target or any structure on the target undergoes micro-motion dynamics, such as mechanical vibrations or rotations, the micro-motion dynamics induce Doppler modulations on the returned signal, known as micro-Doppler effects (Chen et al., 2006). nb. In the case in question here, the range profile can be also given by:

$$P'_D(r) = \text{mean}\, P'_D(r,v), \quad v \in [v_{\min}, v_{\max}]. \tag{8}$$

Then, the range coordinate r_+ is obtained as:

$$r_+ = \begin{cases} \arg\max_{r} P'_D(r) & \text{if } \max P'_D(r) > I, \\ \arg\max_{r} P'(r) & \text{otherwise}, \end{cases} \tag{9}$$

where I is a threshold just to verify that this particular Doppler effect is strong enough for the peak to be considered as a possible detection. If the range

dimension is reduced as follows

$$P'_D(v) = \max_r P'_D(r, v), \quad r \in [r_{\min}, r_{\max}], \tag{10}$$

then the radial velocity v_+ can be estimated according to

$$v_+ = \arg\max_v P'(v) \tag{11}$$

assuming that the target velocity is that of the main intensity peak in this dimension.

Recall, from equation (9), that r_+ is the estimated range of the target. Two 1D angular profiles can now be taken at this range, in order to estimate the direction of the target. Indeed the maximum values of these profiles can be used as simple estimates $[\theta_+, \phi_+]$ of the true target direction $[\Theta, \Phi]$:

$$\theta_+ = \arg\max_\theta P'_A(r_+, \theta) \tag{12}$$

$$\phi_+ = \arg\max_\phi P'_E(r_+, \phi) \tag{13}$$

However, in order to achieve higher accuracy, a GP regression will be applied to the profiles defined in equations (5) and (6). The resulting system learns to understand the radar data, based on a set of training examples.

3.2. Gaussian Process Regression

The previous section described a simple target localization method, based on direct peak extraction, in the background-subtracted response signals. In practice, the direct range estimate is relatively accurate, as can be deduced from equation (A.2), where a resolution of 3.75cm is achieved for a typical bandwidth of 4GHz. The angular maxima, however, are much less reliable, as extracted from equation (A.6), where can be seen how that relationship is non-linear, so that the greater the angle, the worse the resolution. We therefore propose an alternative localization method, in which the entire angular profiles, at the estimated range, are used as inputs to a supervised learning algorithm.

Specifically, we obtain azimuth and elevation estimates, by analogy with equations (9) (12) and (13), from three GP regressions, each taking the concatenated 1D profiles $\left[P'(r),\, P'_A(r_+, \theta),\, P'_E(r_+, \phi)\right]$ as inputs. In general, the GP

kernel representation may be more effective if the dimensionality of the problem is reduced somewhat. This is possible here, because the data have limited resolution in range (A.2) and azimuth/elevation (A.6). In particular, it will be shown that the azimuth and elevation profiles $[P'_A(r_+, \theta),\ P'_E(r_+, \phi)]$ can be compactly represented by their coefficients $[\boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ with respect to a fixed spatial basis. We use a standard B-spline representation, as described in section 3.3. This leads to the following GP regression estimates, by analogy with (12) and (13):

$$\theta_\star \leftarrow \mathrm{GPR}_\theta(r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}) \tag{14}$$

$$\phi_\star \leftarrow \mathrm{GPR}_\phi(r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}). \tag{15}$$

We hypothesize that the GPR estimates $[\theta_\star, \phi_\star]$ are significantly better than the direct maximum estimates $[\theta_+, \phi_+]$ in (12) and (13), for two reasons. Firstly, the relationship between the target direction and the angular data is allowed to be more complex: for example, the peak may be asymmetric, or multimodal. Secondly, the estimator is allowed to be directionally variable, in both dimensions: for example, it is possible to compensate for attenuation, as the true direction $[\Theta, \Phi]$ approaches the angular limits of the sensor.

The GPR method has the following advantages, in this context. Firstly, it produces sensible estimates outside the range of the available data. Secondly, the model has a clear probabilistic interpretation, including confidence intervals for prediction. Thirdly, the parameterization is physically sensible, based on assumed smoothness of the signal distortions. Finally, the estimation process is exact, and very efficient in this application. We now consider the computational form of the estimates (14) and (15), in the standard GP framework (Rasmussen & Williams, 2006).

Let $\boldsymbol{\xi} = [r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ be the representation of a radar reading, from which we wish to obtain a direction estimate $[\theta_\star, \phi_\star]$, as in (14) and (15). A suitable GP model can be estimated from $i = 1, \ldots L$ training locations of the target, with known azimuths and elevations Θ_i and Φ_i. The corresponding RF measurements are concatenated in the L training vectors $\boldsymbol{\xi}_i$. The affinity of two such readings

can be represented by the exponentiated quadratic kernel:

$$k(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = \gamma^2 \exp\left(-\tfrac{1}{2}\, d^2(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j)\right) \quad \text{where}$$
$$d^2(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = \left[\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right]^T \begin{bmatrix} \boldsymbol{\gamma}_r^2 & & \\ & \boldsymbol{\gamma}_\theta^2 & \\ & & \boldsymbol{\gamma}_\phi^2 \end{bmatrix}^{-1} \left[\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\right] \tag{16}$$

is the Mahalanobis distance between $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_j$. The characteristic scale of the range representation is set by $\boldsymbol{\gamma}_r^2 = [\gamma_r^2, \ldots, \gamma_r^2]$, and similarly for $\boldsymbol{\gamma}_\theta^2$ and $\boldsymbol{\gamma}_\phi^2$, where the lengths of the constant vectors correspond to the partitioning of $\boldsymbol{\xi}$ into range and angular components. The overall scale factor γ^2 represents the variance of the estimated function. The kernel parameters can be estimated automatically from the data, using standard fitting routines (Rasmussen & Williams, 2006). It is also possible to constrain the model, e.g. by fixing $\gamma_r = \gamma_\theta = \gamma_\phi$, or $\gamma_\theta = \gamma_\phi$; these simplifications will be evaluated in section 4.

All applications of kernel (16) to the training data from L training locations can be encoded by an $L \times L$ symmetric matrix $\mathbf{K}$, with entries

$$K_{ij} = k(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j). \tag{17}$$

The regularized kernel matrix $\mathbf{L}$ is then defined, in the standard way (Rasmussen & Williams, 2006), where σ^2 is the noise variance:

$$\mathbf{L} = \mathbf{K} + \sigma^2 \mathbf{I}. \tag{18}$$

Let $\boldsymbol{\Theta}$ and $\boldsymbol{\Phi}$ contain the coordinates Θ_i and Φ_i, respectively, of the the L training locations, with corresponding RF representations $\boldsymbol{\xi}_i$. The optimal azimuth and elevation estimates (14) and (15), are obtained from the standard GPR formulas (Rasmussen & Williams, 2006):

$$\theta_\star(\boldsymbol{\xi}) = \mathbf{k}(\boldsymbol{\xi})\, \mathbf{L}^{-1} \boldsymbol{\Theta} \tag{19}$$

$$\phi_\star(\boldsymbol{\xi}) = \mathbf{k}(\boldsymbol{\xi})\, \mathbf{L}^{-1} \boldsymbol{\Phi} \tag{20}$$

where $\boldsymbol{\xi}$ contains the newly observed RF data, and the i-th entry of the $L \times 1$ vector $\mathbf{k}(\boldsymbol{\xi})$ is given by $k(\boldsymbol{\xi}_i, \boldsymbol{\xi})$. It is also possible to compute the conditional

probability of each estimate, given the training data (Rasmussen & Williams, 2006). For example,

$$\theta_\star \,|\, \boldsymbol{\Theta} \sim \mathcal{N}\Big(\mathbf{k}(\boldsymbol{\xi})\,\mathbf{L}^{-1}\boldsymbol{\Theta},\; k(\boldsymbol{\xi},\boldsymbol{\xi}) - \mathbf{k}(\boldsymbol{\xi})\,\mathbf{L}^{-1}\mathbf{k}^T(\boldsymbol{\xi})\Big) \tag{21}$$

where $\mathcal{N}(\mu,\sigma)$ is a 1D Normal distribution, and the analogous formula holds for the elevation estimate, $\phi_\star \,|\, \boldsymbol{\Phi}$.

3.3. B-spline representation

The angular estimates $[\theta_\star, \phi_\star]$ could be obtained by applying GP regression directly to the concatenated profiles $\big[P'(r),\, P'_A(r_+,\theta),\, P'_E(r_+,\phi)\big]$, as described in section 3.2. In practice, we use lower-dimensional azimuth and elevation representations $[\boldsymbol{\alpha}, \boldsymbol{\epsilon}]$, in equations (14) and (15), with $\boldsymbol{\xi} = [r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ being the complete representation, as indicated in section 3.1. This dimension reduction process improves the performance of the GPR method, as will be shown in section 4. In addition, there is a computational benefit when constructing the kernel matrix (16), although this is not a major concern in the present context.

We use a B-spline representation, which has the advantages of local support and numerical stability. Let $\mathbf{B}$ be the $M \times N$ matrix of N cubic B-spline basis functions, with optimally located knots (de Boor, 1978). The azimuth and elevation profiles are approximated by $\mathbf{B}\boldsymbol{\alpha}$ and $\mathbf{B}\boldsymbol{\epsilon}$, respectively, where the coefficient column-vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\epsilon}$ are estimated by standard least-squares methods. If the rows of $\mathbf{B}$ are denoted by $\mathbf{b}_i$, then the approximation can be written more explicitly as

$$P'_A(r_+,\theta_i) \approx \mathbf{b}_i\boldsymbol{\alpha} \tag{22}$$

$$P'_E(r_+,\phi_i) \approx \mathbf{b}_i\boldsymbol{\epsilon} \tag{23}$$

where the index $i = 1,\ldots,M$ ranges over the angular dimension of the corresponding RF data array. The appropriate dimensionality N of the spline basis $\mathbf{B}$ is an empirical question, which will be addressed in section 4. Figure 3 shows an example of dimension reduction for a typical angular profile; in this case the dimensionality is reduced from $M = 63$ to $N = 12$.

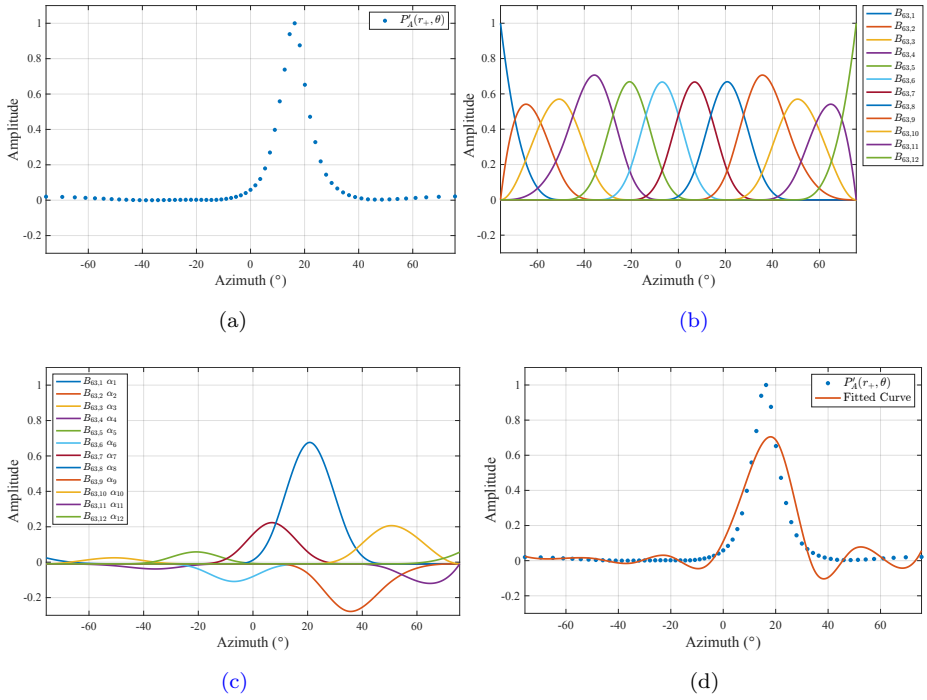

Figure 3: Representation of an example azimuth profile (a), in a B-spline basis (b). The basis functions (columns of $\mathbf{B}$) are individually multiplied by the corresponding coefficients $\boldsymbol{\alpha}$ in (c). In this particular example, the profile $P'_A(r_+, \theta)$ goes from 63 samples in (a) to the 12 dimensions of the $\mathbf{B}\boldsymbol{\alpha}$ representation shown in (d).

4. Localization experiments

This section presents a complete evaluation of the proposed algorithm. Section 4.1 describes the experimental setup, including the relevant radar configuration parameters. Next, section 4.2 shows how initial target detection can be performed, by background subtraction and direct maximum estimation. Finally, section 4.3 evaluates several variations of the proposed Gaussian Process approach.

4.1. Experimental setup

All experiments were performed using a mmWave radar 'system on a chip' (IWR6843ISK-ODS ES2.0, Texas-Instruments (2020)), a low power device op-

erating in the 60–64GHz frequency range, with a ±60° field of view, in both azimuth and elevation. The antenna module (figure 4) comprises three TXi *transmitters*, separated by a distance of λ, and four RXi *receivers*, separated by a distance of $\lambda/2$ (hence a total of 12 virtual antennas). The antenna pattern has relatively low directivity, with 3dB beamwidth of around 100° in the azimuthal plane and 70° in the elevation plane. Low directivity is beneficial for short-range environments, as this makes a wide field-of-view possible, and high gain is not necessary to overcome signal attenuation or reduced resolution at long distances. The performance of the device has been characterized in detail by the manufacturer, Texas-Instruments (2021). Note that the design allows both azimuth and elevation data to be obtained, given that the signals received at a subset of the antennas (marked with asterisks in figure 4) are phase-shifted by 180°. This information is used to form the steering matrix $\mathbf{A}$, as described in section 5.

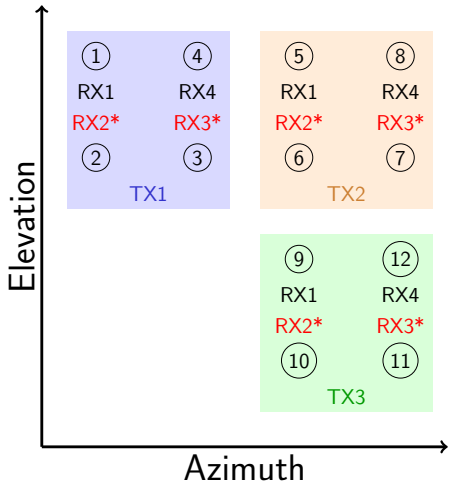

Figure 4: Antenna layout for the IWR6843ISK-ODS. The steering matrix $\mathbf{A}$ is obtained by considering this reception scheme.

The radar device is attached to a movable clamp, together with a digital camera, as shown in figure 5. The camera is used to provide estimates of the target ground truth $[\Theta, \Phi]$, using the standard Aruco markers (Romero-Ramírez

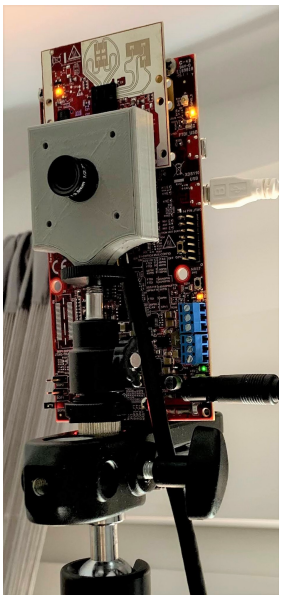

Figure 5: Experimental device. The antenna module is visible at the top, while the camera is fixed underneath.

et al., 2018; Garrido-Jurado et al., 2016). Two measures have been taken to improve RF reflection from the MAVs. Firstly, the sensor was mounted above the flight area, pointing down at the body and propellers. A custom aluminum trihedral retro-reflector was also attached to the drone, as can be seen in figure 6. The main role of the reflector is to increase the maximum detectable range, in order to enable a more complete evaluation of our proposed algorithms. We used a Parrot Mambo Minidrone, with a wingspan of 13.5cm, in our experiments, which were conducted in a relatively small room (approx. 6m×4m), with typical furniture, including shelves, desks, chairs, etc. This environment is relatively challenging, owing to the presence of clutter, and the proximity of reflective walls. The radar was appropriately configured, as summarized in table 1.

4.2. Preliminary analysis

This section contains the results of our preliminary experiments, which were designed to test the basic localization algorithms developed in section 3.1, using

Table 1: Main chirp parameters, adapted to the the present experimental setup.

Parameter	**Value**
Initial Frequency	60GHz
Bandwidth (B)	3.6GHz
Frequency Slope (B/T)	150MHz/µs
ADC Resolution	128 bits
ADC Sampling Frequency	6500kS/S
Chirps/frame	48

the setup described in section 4.1. The basic strategy is to estimate the ground truth position of the drone from Aruco markers, which are known to provide reliable estimates (Romero-Ramírez et al., 2018; Garrido-Jurado et al., 2016), assuming that the target is optically visible (which may not be the case, in subsequent RF applications). We emphasize that the camera system is only used to provide initial reference inputs, for training the Gaussian Process models. The resulting radar system can then operate independently, even in optically

Figure 6: Parrot Mambo Minidrone, with a trihedral retro-reflector for the RCS, for increased RF detectability.

opaque environments (or those in which cameras are forbidden, for privacy reasons).

We begin by validating the background subtraction method, described at the beginning of section 3. We find that this simple method performs well, as shown in figure 7. As can be seen, clutter and reflection would mislead the localization process. However, removing static parts of the signal greatly reduces this problem.

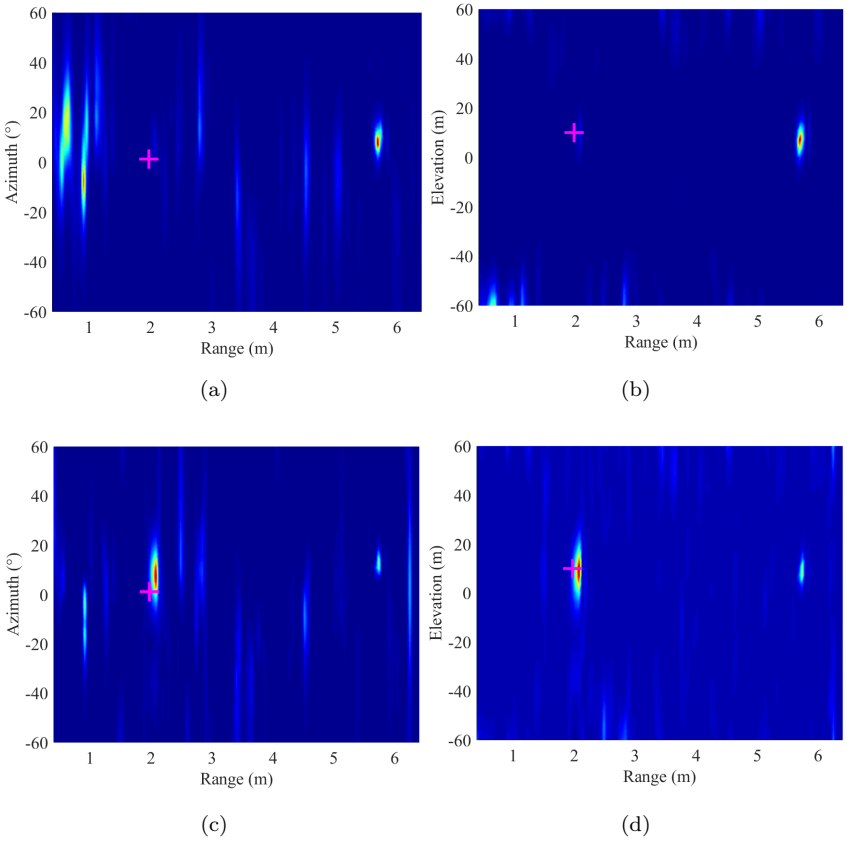

Figure 7: Background subtraction example for azimuth (a,c), and elevation (b,d) heatmaps. The red cross indicates the true position of the target. Environmental clutter produces false peaks in (a,b), which dominate the target. After background subtraction in (c,d), the dominant peak is close to the true position (estimated by the camera system). The residual biases are modelled by Gaussian Process regression.

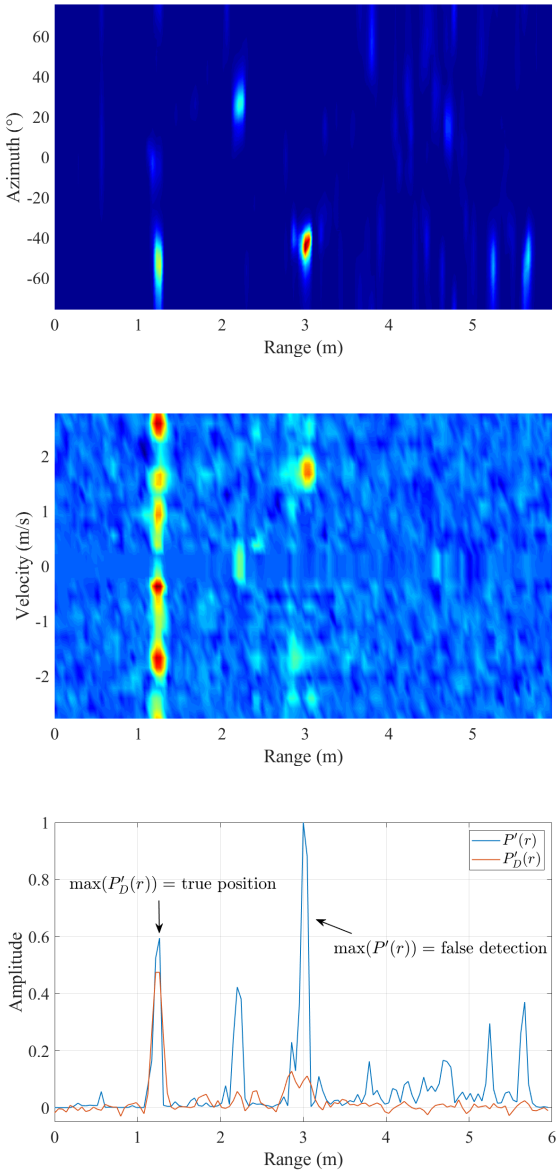

Figure 8: Example use of micro-Doppler information to assist target detection. The peak of the range profile extracted from the range-azimuth heatmap (top), around 3m, does not correspond to the true target range. However, the peak of the range-velocity data (middle) is located at the true range of the target (bottom).

As described in section 3.1, the second way to facilitate target detection is to consider the Doppler range profile, defined in equation (8). The corresponding estimate in this equation makes use of the drone bulk velocity, as well as the micro-Doppler effect generated by local motion of the propellers. In particular, the side-lobes also contribute to the main peak in $P'_D(r)$ from equation (8), which makes the detection more reliable. Figure 8 illustrates how this process performs. The peak of the range-azimuth heatmap (figure 8a) is around 3m, as shown by the blue profile in figure 8c. Meanwhile, the peak of the range-velocity profile, shown in red, is at the correct location of 1m.

We now evaluate the accuracy of the direct maximum approach, described in section 3.1, using the camera-based ground truth, as described above. A total of $L = 50$ different locations were used, with ranges between 1 and 4 m, azimuth angles between $-50°$ and $50°$, and elevation angles between $-20°$ and $20°$. These aerial sites cover most of the available flying space. All data were collected in a static environment, in order to make the ground truth acquisition with the camera as reliable as possible. The localization results are shown in figure 9, using the spherical $[r, \theta, \phi]$ coordinates that are appropriate for the radar device. It can be seen that the direct range estimates are in accordance with the ground truth, with 5.45 cm RMSE (root mean square error). The angular estimates are less accurate, and a positive bias is apparent in the elevation estimates. This systematic effect can be attributed to miscalibration of the device.

4.3. Regression models

Next we evaluate the accuracy of the GP regression model, developed in section 3.2. Recall that $(N \times 1)$ B-spline coefficient vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\epsilon}$ are computed for the azimuth and elevation profiles at range r_+. The complete RF representation $\boldsymbol{\xi} = [r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ is therefore $(1 + 2N)$ dimensional. We evaluate a range of possible choices, $N = 8$, 12, 20. All of the RF representations $\boldsymbol{\xi}_i$ are randomly divided in two groups: 70% training data and 30% test data. The training group is used to build the GP regression estimator, as described in section 3.2.

The results are shown in figure 10, for comparison with figure 9. The red

and blue points represent the estimates for training and test data respectively. Firstly, it can be seen that the training data is accurately represented, as would be expected. More importantly, the test estimates are more accurate than those in the direct maximum, as quantified by the RMSE values in table 2. These values were computed by averaging ten models, created with different random training/test groupings (the figures show example results, from a typical run). In detail, the range, azimuth and elevation errors of the direct maximum method are 5.45 cm, 7.60° and 11.46°, respectively. The proposed GP regres-

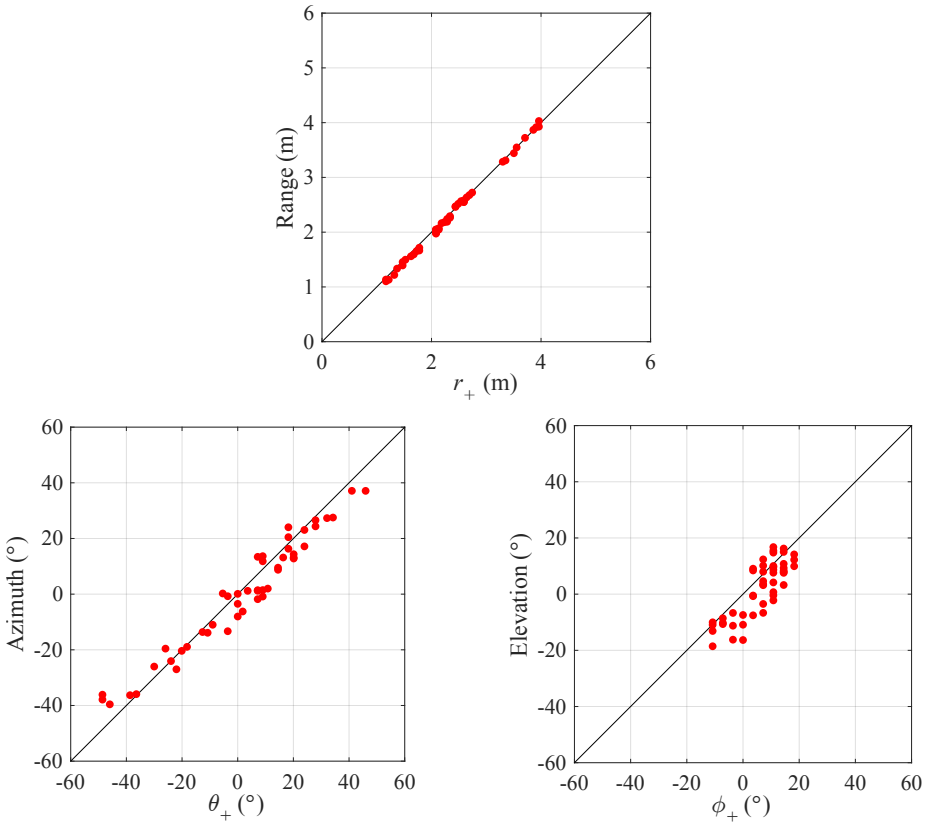

Figure 9: The horizontal axes represent each coordinate of the estimated position $[r_+, \theta_+, \phi_+]$, obtained from the maxima of the (background subtracted) RF profiles. The vertical axes represent the corresponding coordinates $[R, \Theta, \Phi]$, obtained by the camera-based reference method. Errors in the estimates are indicated by vertical deviations from the 45° lines, in each case.

sion, based on the B-spline representation, achieves substantially lower errors of 3.40 cm, 4.54° and 5.13°. It is clear from figure 11 that the GP regression method is able to compensate for the calibration biases that affect the locations of the maxima in figure 9. The RMSEs have been highlighted (bold) in table 2.

We performed additional experiments, to help understand the main results, all of which are shown in table 2. Apart from the direct maximum strategy already discussed, two kinds of GP regressions are evaluated. Firstly, a **uni-**

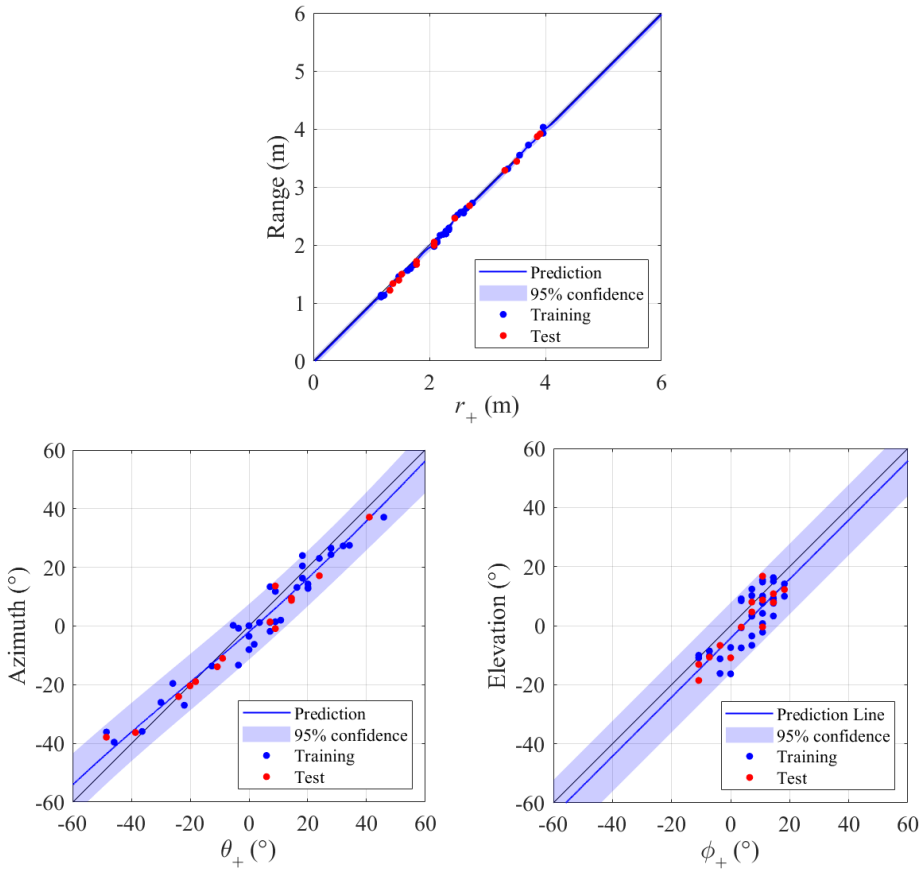

Figure 10: Gaussian process models of the data in figure 9, using independent inputs r_+, θ_+, and ϕ_+, respectively. Errors are represented by vertical deviations of the test points (red) from the estimated function (blue). Note that the azimuth estimate $\theta_\star$ (blue curve) is significantly better than the direct maximum estimate θ_+ (45° line). The shaded region is the 95% confidence envelope of the Gaussian Process model.

variate GPR in which the input for each model is taken separately. These inputs can be the maxima $[r_+]$, $[\theta_+]$, $[\phi_+]$, calculated by the direct strategy.
355 Secondly, a **multivariate** GPR is applied with different inputs. The first one is carried out with the separate profiles $[P'(r)]$, $[P'_A(r_+, \theta)]$, $[P'_E(r_+, \phi)]$, ex-

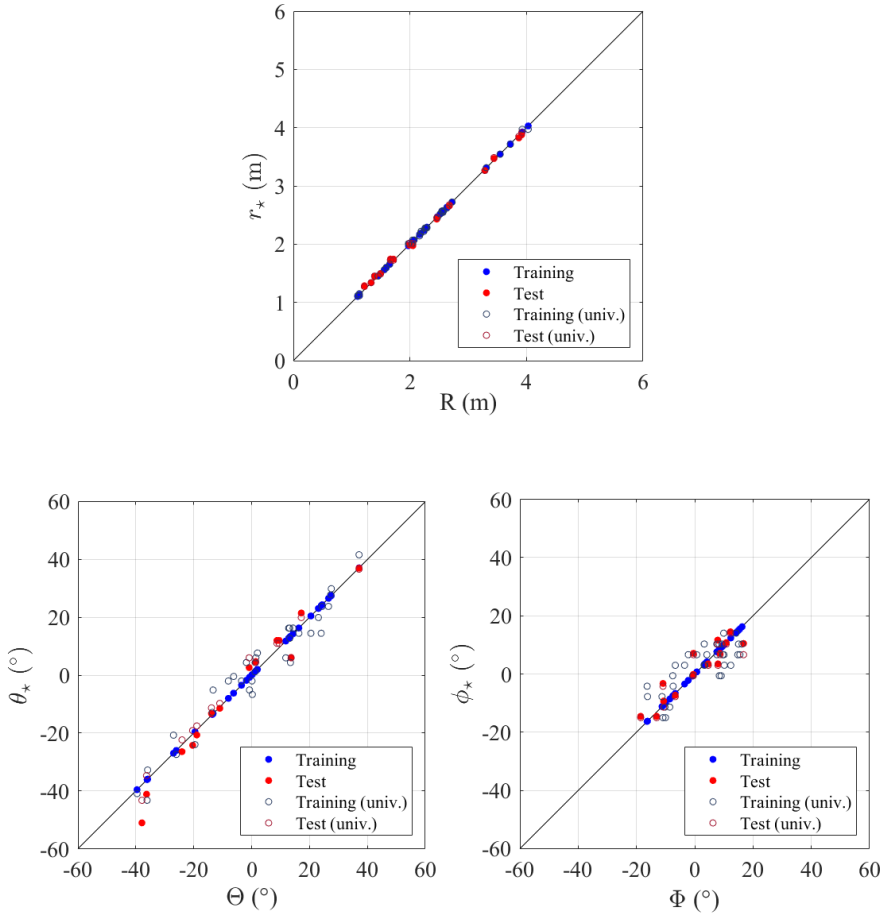

Figure 11: Results from the complete multivariate model (filled circles). Here the complete multidimensional representations $[r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ are used as inputs to each Gaussian Process Regression model. The univariate case results (unfilled circles), where the inputs are taken as $[r_+]$, $[\boldsymbol{\alpha}]$, $[\boldsymbol{\epsilon}]$ individually, are also shown. The horizontal axes show the reference coordinates $[R, \Theta, \Phi]$, in this figure. The vertical axes show the corresponding estimates, with vertical deviations being prediction errors.

tracted as proposed in section 3.1, or their reduced versions $[\boldsymbol{\rho}]$, $[\boldsymbol{\alpha}]$, $[\boldsymbol{\epsilon}]$. A general improvement can be observed for all of them, being noteworthy the one achieved for range. The second one is applied to the concatenated raw profiles, $\left[P'(r),\, P'_A(r_+,\theta),\, P'_E(r_+,\phi)\right]$, of dimension $[120, 63, 31]$, then the azimuth and elevation RMSEs rise to 5.99° and 6.02°. The method with this input does not provide a clear improvement over the last method. However, if the input consists in a concatenation of the reduced versions $[\boldsymbol{\rho}, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$ or $[r_+, \boldsymbol{\alpha}, \boldsymbol{\epsilon}]$, then a general RMSE improvement better than for univariate or separate multivariate GPR is found. This suggests that the GP kernel representation works more effectively in the lower-dimensional B-spline space.

Furthermore, we have completed the study by varying the freedom with which the GP routine can optimize the kernel parameters in equation (16). The overall variance γ^2 of the estimated function must be estimated in all cases. The length scales for range, azimuth and elevation can be constrained to a single number $\gamma_{r\theta\phi}$, a pair $[\gamma_r, \gamma_{\theta\phi}]$, or a triple $[\gamma_r, \gamma_\theta, \gamma_\phi]$. The performance of the GP regression was comparable, across these variations, as indicated in table 2. Overall, the most flexible model performed best, in conjunction with an intermediate spline dimensionality, of $N = 12$.

Table 2: RMSE for range, azimuth and elevation estimates obtained by applying direct maximum estimation, univariate GPR (only possible for separate inputs) and multivariate GPR for both separate and concatenated inputs, considering the whole profiles or their reduced versions.

	Direct Maxima	**Univariate GPR**	**Multivariate GPR**								
		$[r_+]\,[\theta_+]\,[\phi_+]$	$[P'_R]\,[P'_A]\,[P'_E]$	$[\rho]\,[\alpha]\,[\epsilon]$	$[P'_R, P'_A, P'_E]$	$[\rho, \alpha, \epsilon]$	$[r_+, \alpha, \epsilon]$				
							$N = 20$	$N = 12$			$N = 8$
								$[\gamma_r, \gamma_\theta, \gamma_\phi]$	$[\gamma_r, \gamma_{\theta\phi}]$	$\gamma_{r\theta\phi}$	
Range (cm)	5.45	**3.40**	4.61	4.06	4.29	4.06	4.08	4.19	4.20	3.92	4.46
Azimuth (°)	7.60	6.93	5.76	5.78	5.99	5.17	5.20	**4.54**	5.03	4.79	5.07
Elevation (°)	11.46	5.94	5.73	5.89	6.02	5.71	6.03	**5.13**	5.42	5.15	6.05

A summary of the results, plotted in Cartesian coordinates, can be seen in figure 12. In most cases, the GPR estimates (blue points) are closer to

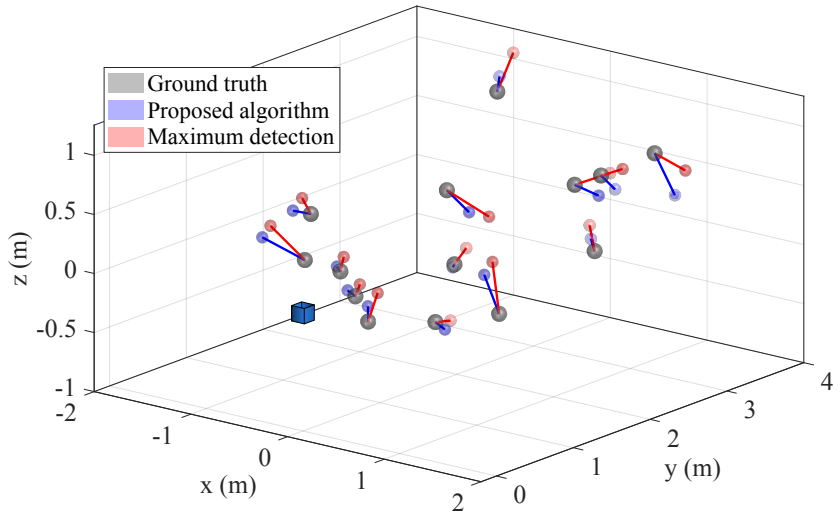

Figure 12: Summary of results, plotted in physical coordinates. The gray cubes represent the ground truth positions, with the diameter of the spheres representing the drone size. The results from the direct maximum strategy are shown in red, and those from the final proposed GPR method in blue. The line segments show the pattern of residual errors. The small cube indicates the location of the radar device, and the line segments show the pattern of residual errors.

the ground truth (gray points) than those extracted from the direct maximum method (red points), as expected. The overall RMSEs in space are 22.86 cm for the first method and 31.84 cm for the second.

5. Conclusions

This paper has presented a novel 3D localization system for UAVs, flying in cluttered environments, based on mmWave radar signals. This is a prototype for a low-cost and low-power system, which could function in poor visibility conditions, such as fog or smoke. The ability to avoid recording conventional images has additional privacy benefits, in urban settings, for example.

We have shown that conventional algorithms, including background subtraction and spectral analysis, can be used to estimate the range of a drone. The additional value of micro-Doppler signals, generated by rotors and propellers,

has also been demonstrated. A novel Gaussian Process regression method has been used to refine the range and the angular localization of the target. The importance of performing dimension reduction, before training the regression model, has also been established. The improvement in the RMSE is approximately 3° for azimuth, 6° for elevation, and 2cm in range, with respect to direct maximum methods. Most importantly, the Gaussian Process regression is able to compensate for any mis-calibration of the radar system, which is an important consideration when using commodity devices. Our experiments also indicate that this approach is robust to environmental clutter. Future work will investigate the performance of these methods in dynamic environments.

Author statement and acknowledgments

The authors' contributions to the paper have been as follows: José A. Paredes: conceptualization, methodology, investigation, and writing - original draft; Fernando J. Álvarez: resources, supervision and writing - review & editing; Miles Hansard: formal analysis, supervision and writing - review & editing; Khalid Rajab: software and resources, supervision and writing - review & editing.

This work has been partially supported by the UK EPSRC National Centre for Nuclear Robotics (NCNR) EP/R02572X/1; by the Spanish Government and the European Regional Development Fund (ERDF) through Project MICROCEBUS under Grant RTI2018-095168-B-C54; and by the Regional Government of Extremadura and ERDF - ESF, under Project GR18038 and through the Pre-Doctoral Scholarship under Grant 45/2016 Exp. PD16030.

Appendix A. Radar signal processing

The fundamental concept in radar systems is the emission of a signal, and its reception after being reflected from the surrounding objects. Specifically, an FMCW (frequency modulated continuous wave) radar is used in this work, in which a continuous chirp-like signal is employed to synthesize a wide bandwidth, and thus a high range resolution.

The radar working principles are depicted in figure A.13. First, a linear chirp-like wave is emitted with a defined period T. Once the echo has been received (delayed version of the emitted wave), the signal passes through an electronic mixer, together with the original version. The output from this mixer is a new signal whose instantaneous frequency comes from the subtraction of the emitted and received instantaneous frequencies, and has a lower frequency f_{IF} component that can be extracted by applying a low pass filtering. This low frequency f_{IF} is directly proportional to the distance r from the radar to the target:

$$r = \frac{cT}{2B} f_{\mathrm{IF}} \tag{A.1}$$

where c is the speed of light, and B the bandwidth swept by the emitted chirps. The range resolution is given by:

$$\Delta r = \frac{c}{2B} \tag{A.2}$$

A detailed analysis of an FMCW radar working principle can be read in Richards (2005).

Further target characteristics can be obtained with the emission of two or more time-separated signals. A small change Δd in the distance of an object results in a phase change $\Delta\varphi$ in the intermediate signal, which can be expressed as:

$$\Delta\varphi = 2\pi f \Delta\tau = 2\pi \frac{c}{\lambda} \frac{2\Delta d}{c} = 4\pi \frac{\Delta d}{\lambda} \tag{A.3}$$

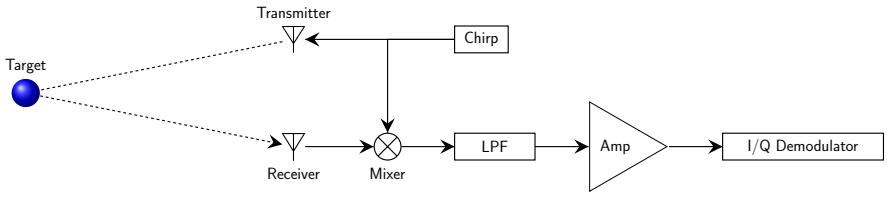

Figure A.13: Radar operation: First, a wave is transmitted, and once the reflection is received, it is mixed with the generated chirp waveform. The resulting signal is low-pass filtered (LPF), and amplified (Amp). Finally, the I/Q demodulator extracts the real and imaginary components from the IF signal, for input into an ADC.

where $\lambda = c/f$ is the wavelength. Thus, velocity v is obtained directly from the phase measurement as:

$$v = \frac{\Delta d}{T_c} = \frac{\lambda \Delta \varphi}{4\pi T_c} \tag{A.4}$$

Here T_c is the period between two consecutive chirps. Usually, several chirps are emitted in a certain time interval T_f, known as a frame, to have a more reliable velocity measurement. As will be shown below, this velocity will be used to strengthen the detection of targets in the proposed algorithm.

Finally, extracting azimuth and elevation is necessary in order to obtain the complete set of coordinate information. For this purpose, the device (described in section 4.1) is based on a MIMO (multiple-input multiple-output) approach (Fishler et al., 2004; Li & Stoica, 2007), which leverages improved processing capabilities to apply digital beamforming for optimizing angular estimation, without the need for expensive phased arrays. Thus, the AoA (angle of arrival) can be computed by making use of multiple receivers:

$$\theta \,(\text{or}\, \phi) = \sin^{-1}\left(\frac{\lambda \Delta \hat{\varphi}}{2\pi L}\right) \tag{A.5}$$

where L is the distance between the two receivers, assuming a planar wavefront, and $\Delta\hat{\varphi}$ is now the phase difference caused by the delay in the wave arriving at the different receivers. Note that this is a nonlinear dependency, hence the smaller the AoA (closer to 0°), the more accurate the estimates. Moreover, the angular resolution is dependent on the number of antennas K as follows:

$$\Delta\theta \geq \frac{\lambda}{KL\cos\theta} \tag{A.6}$$

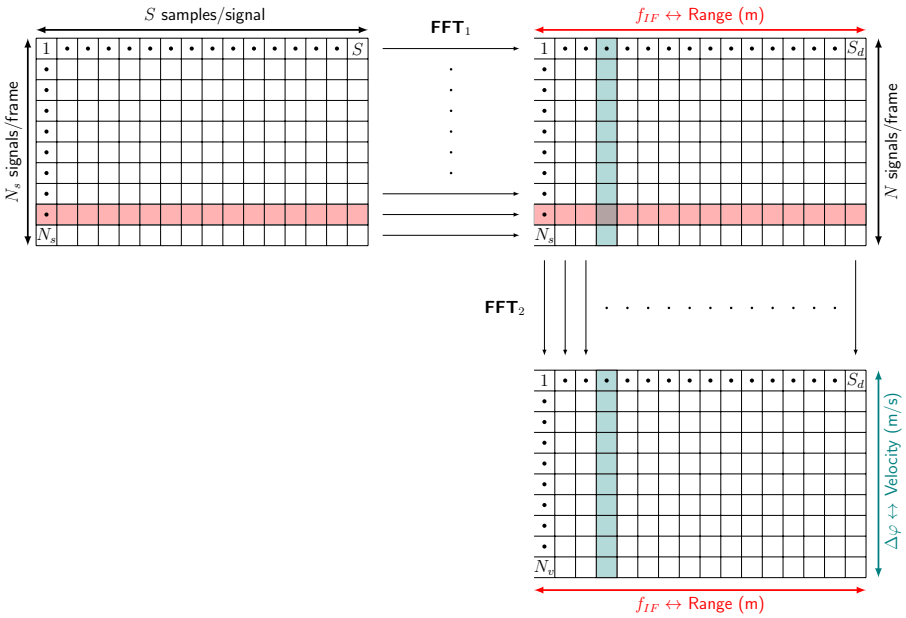

Figure A.14: Spectral analysis required to get the range-velocity heatmap. At the beginning, every antenna receives N_s signals in a determined period of time (frame). Once all received signals have been organized in an array (TOP LEFT), an FFT (fast Fourier transform) (FFT_1) is performed to obtain f_{IF}, or equally, the range r, according to equation (A.1). Finally, the elements of each column for the new array (TOP RIGHT) will have the same magnitude with linearly increasing phase $\Delta\varphi$, which can be obtained by applying a new FFT (FFT_2) that will provide velocity information (BOTTOM), as in equation (A.4).

In order to extract information about range, velocity, azimuth and elevation, this paper employs spectral analysis techniques, such as FFT (fast Fourier transform) and MUSIC (Schmidt, 1986), through which the response versus range or angle can be estimated, and then the peaks can be found to extract their values. These techniques are reviewed below.

Direct spectral analysis

Firstly, it is necessary to extract f_{IF} from equation (A.1), over the total amount of chirps emitted/received during each acquisition period (frame), which provides data about the object range. This operation, which involves an FFT

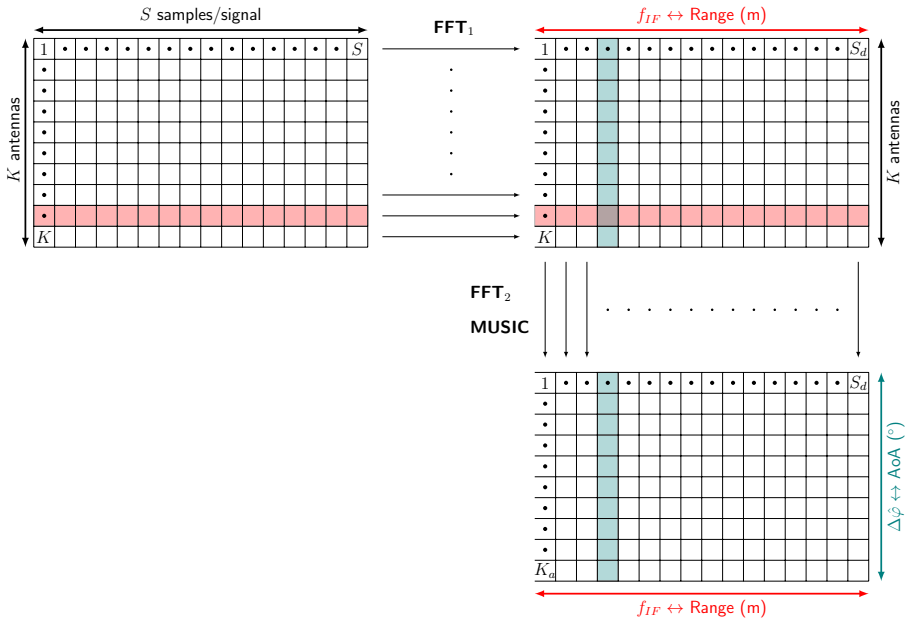

Figure A.15: Spectral analysis required to get range-AoA (azimuth/elevation) heatmaps. Analogously with the velocity computation, the received signals are arranged in an array, but this time the signals are those received at the same time in different antennas (TOP LEFT). First, an FFT (FFT_1) can be done to figure out f_{IF}, or equally, the range r - equation (A.1). Finally, the elements of each column for the new array (TOP RIGHT) will have the same magnitude with linearly increasing phase $\Delta\hat{\varphi}$, which can be obtained by applying a new FFT (FFT_2) that will provide angular information (BOTTOM) - equation (A.5).

(FFT_1 in figures A.14 and A.15), is performed by the device. A second analysis must then be performed, in order to apply equations (A.4) and (A.5) thereby obtaining $\Delta\varphi$ and $\Delta\hat{\varphi}$. After the first FFT, all elements of each column in the second matrix, both in figures A.14 and A.15, have the same magnitude, while the phases vary linearly. Hence, these phases can be obtained by performing a further spectral analysis, this time over each column. If this is calculated by taking into account the chirps arriving individually at each antenna, then a range-velocity heatmap is obtained (FFT_2 in figure A.14). On the other hand, if it is done by arranging the signals according to the pair of antennas that has emitted/received, using a steering vector/matrix (Chen et al., 2010),

a range-azimuth or a range-elevation heatmap is produced (FFT_2 or MUSIC in figure A.15).

Subspace spectral analysis

A drawback of the FFT approach, when used to extract AoA information, is that the side lobes can mask the main peak. For this reason, a different spectral analysis is applied here: the MUSIC algorithm (Schmidt, 1986), which estimates the frequency content of a signal by using an eigenspace method. The idea is based on decomposing the input signal covariance matrix into signal and noise components:

1. Consider K antennas receiving M-signals $\mathbf{x}_i$:

$$\mathbf{x}_i(t) = \sum_{m=1}^{M} \mathbf{s}_m(t) \mathrm{e}^{\mathrm{j}2\pi\tau_m(i-1)} + \boldsymbol{\psi}_i \quad \text{with } 0 < i \leq K \tag{A.7}$$

   where $\mathbf{s}_m$ is the m-th emitted signal out of M, $\tau_m = L\cos(\theta_m)/\lambda$ is the time shift for each antenna, and $\boldsymbol{\psi}_i$ corresponds to noise.

2. The former equation can be expressed in matrix form:

$$\begin{bmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \\ \vdots \\ \mathbf{x}_K(t) \end{bmatrix} = \begin{bmatrix} \mathbf{a}(\theta_1) & \mathbf{a}(\theta_2) & \cdots & \mathbf{a}(\theta_M) \end{bmatrix} \begin{bmatrix} \mathbf{s}_1(t) \\ \mathbf{s}_2(t) \\ \vdots \\ \mathbf{s}_M(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{\psi}_1(t) \\ \boldsymbol{\psi}_2(t) \\ \vdots \\ \boldsymbol{\psi}_K(t) \end{bmatrix} \tag{A.8}$$

   or, in general:

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \boldsymbol{\Psi} \tag{A.9}$$

3. The $\mathbf{a}(\theta_i)$ are the steering vectors:

$$\mathbf{a}(\theta_i) = \begin{bmatrix} 1 & \mathrm{e}^{\mathrm{j}2\pi\tau_i(\theta_i)} & \mathrm{e}^{\mathrm{j}2\pi 2\tau_i(\theta_i)} & \cdots & \mathrm{e}^{\mathrm{j}2\pi(K-1)\tau_i(\theta_i)} \end{bmatrix}^{\mathrm{T}} \tag{A.10}$$

The covariance matrix for $\mathbf{X}$ can be written as:

$$\mathbf{R} = \mathrm{E}\left[\mathbf{X}\mathbf{X}^{\mathrm{H}}\right] \tag{A.11}$$

where E is the operator denoting the expected value (mean) of its argument, and $\mathbf{X}^{\mathrm{H}}$ is the Hermitian transpose of $\mathbf{X}$. It follows that:

$$\begin{aligned}\mathbf{R} &= \mathrm{E}\big[(\mathbf{AS}+\mathbf{\Psi})(\mathbf{AS}+\mathbf{\Psi})^{\mathrm{H}}\big] = \mathbf{S}\mathrm{E}\big[\mathbf{AA}^{\mathrm{H}}\big]\mathbf{S}^{\mathrm{H}} + \mathrm{E}\big[\mathbf{\Psi\Psi}^{\mathrm{H}}\big] = \\ &= \mathbf{SR}_s\mathbf{S}^{\mathrm{H}} + \sigma^2\mathbf{I} = \mathbf{SR}_s\mathbf{S}^{\mathrm{H}} + \mathbf{R}_n\end{aligned} \tag{A.12}$$

where $\mathbf{R}_n$ is the noise correlation matrix. The signal correlation matrix $\mathbf{R}_s$ is diagonal and can be expressed as:

$$\mathbf{R}_s = \mathbf{AA}^{\mathrm{H}} = \begin{bmatrix} \mathrm{E}\big[|\mathbf{s}_1|^2\big] & 0 & \cdots & 0 \\ 0 & \mathrm{E}\big[|\mathbf{s}_2|^2\big] & \cdots & 0 \\ \vdots & \cdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathrm{E}\big[|\mathbf{s}_N|^2\big] \end{bmatrix} \tag{A.13}$$

In practice, $\mathbf{R}$ is unknown, so it must be estimated from the received signal as follows:

1. Estimate the (positive semi-definite) covariance matrix, based on the M received signals:
$$\hat{\mathbf{R}} = \frac{1}{M}\sum_{m=1}^{M} \mathbf{x}_m\ \mathbf{x}_m^{\mathrm{H}} \tag{A.14}$$
2. Order the eigenvalues from highest to lowest $(\lambda_1, ..., \lambda_M, ..., \lambda_N)$, and take the top M (with corresponding eigenvectors) as the signal subspace $\mathbf{E}_s$. The remaining $N - M$ components constitute the noise subspace $\mathbf{E}_n$
3. Since the steering vectors corresponding to signal components are orthogonal to the noise subspace eigenvectors, $\mathbf{AE}_n\mathbf{E}_n^{\mathrm{H}}\mathbf{A} = 0$ for the angles corresponding to AoAs, then the MUSIC spectrum is constructed by taking the following inverse:
$$\mathcal{P}(\theta) = \frac{1}{\mathbf{AE}_n\mathbf{E}_n^{\mathrm{H}}\mathbf{A}} \tag{A.15}$$
The AoAs of the multiple incident signals can be estimated by locating the peaks in this equation.

Conceptually, the MUSIC algorithm can achieve high frequency estimation accuracy, but it involves searching for the spectral peak over the full range frequency

domain, which is computationally demanding. This fact requires the code to be optimized in terms of execution time.

It is worth noting that we have considered only one moving target, and the presence of a second target complicates the extraction of Doppler characteristics. We note that the beamforming techniques (i.e. MUSIC) presented here can be applied to the MIMO radar outputs, to isolate the Doppler responses of multiple targets. Using equation (A.9) a beamformer can be represented as:

$$\mathbf{X}' = \mathbf{W}^{\mathrm{H}}\mathbf{X} = \mathbf{W}^{\mathrm{H}}\big(\mathbf{AS} + \mathbf{\Psi}\big) \tag{A.16}$$

where $\mathbf{W}$ represents the beamforming weight matrix, which is dependent on the AoA and can be calculated using the previously presented techniques as:

$$\mathbf{W} = \mathcal{P}(\theta) \cdot \hat{\mathbf{R}}^{-1} \cdot \mathbf{A} \tag{A.17}$$

By selection of the beamforming weights to correspond to the AoA of a specific target, and applying those weights to the Doppler spectrum, one may therefore isolate the Doppler response of one specific target. properly designing the beamformer weight matrix, the beam can be steered in the desired direction.

This is useful because, once the associated weight is figured out for range-angle information, it can be applied to the Doppler response to obtain the Doppler spectrum for that specific angle. In this paper, an only target is considered in order to make the process less demanding and focus on the extraction of coordinates.

References

Alwateer, M., & Loke, S. W. (2020). Emerging Drone Services: Challenges and Societal Issues. *IEEE Technol. Soc. Mag.*, *39*, 47–51. doi:10.1109/MTS.2020.3012325.

Amidi, O., Kanade, T., & Fujita, K. (1999). A visual odometer for autonomous helicopter flight. *Robotics and Autonomous Systems*, *28*, 185–193. doi:10.1016/S0921-8890(99)00016-0.

Benyamin, M., & Goldman, G. H. (2014). *Acoustic Detection and Tracking of a Class I UAS with a Small Tetrahedral Microphone Array*. Technical Report Army Research Laboratory.

Boddhu, S. K., McCartney, M., Ceccopieri, O., & Williams, R. L. (2013). A collaborative smartphone sensing platform for detecting and tracking hostile drones. In T. Pham, M. A. Kolodny, & K. L. Priddy (Eds.), *SPIE Defense, Security, and Sensing* (p. 874211). Baltimore, Maryland, Estados Unidos. doi:10.1117/12.2014530.

de Boor, C. (1978). *A Practical Guide to Splines*. Applied Mathematical Sciences. Nueva York: Springer-Verlag.

Boudjit, K., & Larbes, C. (2015). Detection and implementation autonomous target tracking with a quadrotor AR.drone:. In *12th International Conference on Informatics in Control, Automation and Robotics* (pp. 223–230). Colmar, Alsace, France. doi:10.5220/0005523102230230.

Busset, J., Perrodin, F., Wellig, P., Ott, B., Heutschi, K., Ruehl, T., & Nussbaumer, T. (2015). Detection and tracking of drones using advanced acoustic cameras. In *SPIE The International Society for Optics and Photonics (SPIE)* (pp. 1–8). Toulouse, France: Spie-Int Soc Optical Engineering volume 9647. doi:10.1117/12.2194309.

Caris, M., Stanko, S., Johannes, W., Sieger, S., & Pohl, N. (2016). Detection and tracking of Micro Aerial Vehicles with millimeter wave radar. In *Proceedings of the 2016 European Radar Conference (EuRAD)* (pp. 406–408). London, UK.

Chen, V., Fayin Li, Shen-Shyang Ho, & Wechsler, H. (2006). Micro-doppler effect in radar: Phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.*, *42*, 2–21. doi:10.1109/TAES.2006.1603402.

Chen, Z., Gokeda, G., & Yu, Y. (2010). *Introduction to Direction-of-Arrival Estimation*. Artech House Signal Processing Library. Boston: Artech House.

de Haag, M. U., Bartone, C. G., & Braasch, M. S. (2016). Flight-test evaluation of small form-factor LiDAR and radar sensors for sUAS detect-and-avoid applications. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)* (pp. 1–11). Sacramento, California, US. doi:10.1109/DASC.2016.7778108.

de Swarte, T., Boufous, O., & Escalle, P. (2019). Artificial intelligence, ethics and human values: The cases of military drones and companion robots. *Artif Life Robotics*, *24*, 291–296. doi:10.1007/s10015-019-00525-1.

DJI (2020). Phantom 4 Pro. URL: https://www.dji.com/phantom-4-pro-v2/specs.

Dogru, S., Baptista, R., & Marques, L. (2019). Tracking drones with drones using millimeter wave radar. In *Robot 2019: Fourth Iberian Robotics Conference* (pp. 392–402). Oporto, Portugal volume 1093. doi:10.1007/978-3-030-36150-1_32.

Dogru, S., & Marques, L. (2020). Pursuing drones with drones using millimeter wave radar. *IEEE Robot. Autom. Lett.*, *5*, 4156–4163. doi:10.1109/LRA.2020.2990605.

Ezuma, M., Ozdemir, O., Anjinappa, C. K., Gulzar, W. A., & Guvenc, I. (2019). Micro-UAV detection with a low-grazing angle millimeter wave radar. In *IEEE RWS: IEEE Radio Wireless Symposium*. Orlando, Florida, US. arXiv:1902.05483.

Ferguson, R., Chevrier, M., & Rankin, A. (2018). *mmWave Radar: Enabling Greater Intelligent Autonomy at the Edge*. Technical Report Texas Instruments Incorporated.

Fioranelli, F., Ritchie, M., Griffiths, H., & Borrion, H. (2015). Classification of loaded/unloaded micro-drones using multistatic radar. *Electron. Lett.*, *51*, 1813–1815. doi:10.1049/el.2015.3038.

Fishler, E., Haimovich, A., Blum, R., Chizhik, D., Cimini, L., & Valenzuela, R. (2004). MIMO radar: An idea whose time has come. In *Proceedings of the 2004 IEEE Radar Conference (IEEE Cat. No.04CH37509)* (pp. 71–78). Philadelphia, PA, US. doi:10.1109/NRC.2004.1316398.

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Medina-Carnicer, R. (2016). Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition*, *51*, 481–491. doi:10.1016/j.patcog.2015.09.023.

Guvenc, I., Koohifar, F., Singh, S., Sichitiu, M. L., & Matolak, D. (2018). Detection, tracking, and interdiction for amateur drones. *IEEE Commun. Mag.*, *56*, 75–81. doi:10.1109/MCOM.2018.1700455.

Guvenc, I., Ozdemir, O., Yapici, Y., Mehrpouyan, H., & Matolak, D. (2017). Detection, localization, and tracking of unauthorized UAS and jammers. In *2017 IEEE/AIAA 36th Digital Avionics Systems Conference (DASC)* (pp. 1–10). St. Petersburg, FL, US: IEEE. doi:10.1109/DASC.2017.8102043.

Harvard, J., Hyvönen, M., & Wadbring, I. (2020). Journalism from Above: Drones and the Media in Critical Perspective. *MaC*, *8*, 60–63. doi:10.17645/mac.v8i3.3442.

Jian, M., Lu, Z., & Chen, V. C. (2017). Experimental study on radar micro-Doppler signatures of unmanned aerial vehicles. In *2017 IEEE Radar Conference (RadarConf)* (pp. 0854–0857). Washington, DC, US. doi:10.1109/RADAR.2017.7944322.

Jiménez Lugo, J., & Zell, A. (2014). Framework for autonomous on-board navigation with the AR.Drone. *J Intell Robot Syst*, *73*, 401–412. doi:10.1007/s10846-013-9969-5.

Kellermann, R., Biehle, T., & Fischer, L. (2020). Drones for parcel and passenger transportation: A literature review. *Transportation Research Interdisciplinary Perspectives*, *4*, 100088. doi:10.1016/j.trip.2019.100088.

Kendoul, F. (2012). Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. *J. Field Robot.*, *29*, 315–378. doi:10.1002/rob.20414.

Li, J., & Stoica, P. (2007). MIMO Radar with Colocated Antennas. *IEEE Signal Process. Mag.*, *24*, 106–114. doi:10.1109/MSP.2007.904812.

Masselli, A., & Zell, A. (2012). A novel marker based tracking method for position and attitude control of MAVs. In *2012 International Micro Air Vehicle Conference and Flight Competition* (pp. 1–6). Lion, France.

Multerer, T., Ganis, A., Prechtel, U., Miralles, E., Meusling, A., Mietzner, J., Vossiek, M., Loghi, M., & Ziegler, V. (2017). Low-cost jamming system against small drones using a 3D MIMO radar based tracking. In *2017 European Radar Conference (EURAD)* (pp. 299–302). Nuremberg, Germany: IEEE. doi:10.23919/EURAD.2017.8249206.

Nanzer, J. A., & Chen, V. C. (2017). Microwave interferometric and Doppler radar measurements of a UAV. In *2017 IEEE Radar Conference (RadarConf)* (pp. 1628–1633). doi:10.1109/RADAR.2017.7944468.

Paredes, J. A., Álvarez, F. J., Aguilera, T., & Aranda, F. J. (2020). Precise drone location and tracking by adaptive matched filtering from a top-view ToF camera. *Expert Systems with Applications*, *141*, 112989. doi:10.1016/j.eswa.2019.112989.

Paredes, J. A., Álvarez, F. J., Aguilera, T., & Villadangos, J. M. (2017). 3D Indoor Positioning of UAVs with Spread Spectrum Ultrasound and Time-of-Flight Cameras. *Sensors*, *18*, 89. doi:10.3390/s18010089.

Rahman, S., & Robertson, D. A. (2018). Radar micro-Doppler signatures of drones and birds at K-band and W-band. *Sci Rep*, *8*, 17396. doi:10.1038/s41598-018-35880-9.

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. Cambridge, Mass: MIT Press.

Richards, M. A. (2005). *Fundamentals of Radar Signal Processing*. McGraw Hill Professional.

Romero-Ramírez, F. J., Muñoz-Salinas, R., & Medina-Carnicer, R. (2018). Speeded up detection of squared fiducial markers. *Image and Vision Computing*, *76*, 38–47. doi:10.1016/j.imavis.2018.05.004.

Samland, F., Fruth, J., Hildebrandt, M., Hoppe, T., & Dittmann, J. (2012). AR.Drone: Security threat analysis and exemplary attack to track persons. In *IS&T/SPIE Electronic Imaging*. California, CA, US. doi:10.1117/12.902990.

Santana, L. V., Brandao, A. S., Sarcinelli-Filho, M., & Carelli, R. (2014). A trajectory tracking and 3D positioning controller for the AR.Drone quadrotor. In *2014 International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 756–767). Florida, US: IEEE. doi:10.1109/ICUAS.2014.6842321.

Schmidt, R. (1986). Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propag.*, *34*, 276–280. doi:10.1109/TAP.1986.1143830.

Sun, Y., Fu, H., Abeywickrama, S., Jayasinghe, L., Yuen, C., & Chen, J. (2018). Drone Classification and Localization Using Micro-Doppler Signature with Low-Frequency Signal. In *2018 IEEE International Conference on Communication Systems (ICCS)* (pp. 413–417). Chengdu, China: IEEE. doi:10.1109/ICCS.2018.8689237.

Tahar, K. N., & Kamarudin, S. S. (2016). UAV onboard GPS in positioning determination. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, *XLI-B1*, 1037–1042. doi:10.5194/isprs-archives-XLI-B1-1037-2016.

Tahmoush, D. (2015). Review of micro-Doppler signatures. *Sonar Navig. IET Radar*, *9*, 1140–1146. doi:10.1049/iet-rsn.2015.0118.

Texas-Instruments (2020). IWR6843ISK-ODS IWR6843 intelligent mmWave overhead detection sensor (ODS) antenna plug-in module. URL: http://www.ti.com/tool/IWR6843ISK-ODS.

Texas-Instruments (2021). 60GHz mmWave Sensor EVMs. URL: https://www.ti.com/lit/ug/swru546d/swru546d.pdf?ts=1618770130459&ref_url=https%253A%252F%252Fwww.ti.com%252Fproduct%252FIWR6843AOP.

Vidal, V. F., Honorio, L. M., Santos, M. F., Silva, M. F., Cerqueira, A. S., & Oliveira, E. J. (2017). UAV vision aided positioning system for location and landing. In *2017 18th International Carpathian Control Conference (ICCC)* (pp. 228–233). Sinaia, Romania: IEEE. doi:10.1109/CarpathianCC.2017.7970402.

Wang, R., Wen, B., & Huang, W. (2018). A Support Vector Regression-Based Method for Target Direction of Arrival Estimation From HF Radar Data. *IEEE Geosci. Remote Sensing Lett.*, *15*, 674–678. doi:10.1109/LGRS.2018.2807405.