

Multi-target multi-sensor localization and tracking using passive antennas and optical sensors on UAVs

Marek Schikora^{a,b} and Daniel Bender^a and Wolfgang Koch^a and Daniel Cremers^b

^aFraunhofer FKIE, Neuenahrer Straße 20, Wachtberg, Germany

^bTechnical University of Munich, Boltzmannstraße 3, Garching, Germany

ABSTRACT

In this work we focus on the task to localize and track multiple non-cooperative targets by a passive antenna array and an optical sensor. Both sensor systems are mounted on a UAV and obtain bearing measurements from the targets, where the number of targets is unknown. To solve the localization and tracking problem, the imprecise but unique bearing data collected from the antenna array has to be correlated with the precise but ambiguous bearing data gained from the optical system. We perform this by a Monte Carlo realization of a multi-sensor probability hypothesis density (PHD) filter.

Keywords: Bearing data, multi-target tracking, localization, passive system, UAV, sensor fusion, PHD filter

1. INTRODUCTION

Unmanned aerial vehicles (UAVs) have an operational benefit in all dangerous and/or dull tasks. Especially small UAVs are often used for surveillance. Often an optical system is used to investigate an interesting area. In recent time antenna arrays have been constructed to localize and track communication emitters with a small UAV. Keeping in mind that small UAVs have a limited payload capability the performance of the localization results, achieved with such antenna arrays is not always sufficient. The aim in this paper is to improve the localization and tracking results of those array systems by fusing the bearing data computed by them with bearing data computed from optical systems. Both sensor types measure 2D bearing data in 3D coordinates. The challenging task is now to invert the measurement step of every sensor in order to generate 3D coordinates from 2D data. In the field of computer vision this task is usually done through a set of multiple cameras, which are observing the same scene. This approach leads to stereo systems and 3D-reconstruction algorithms.¹⁻³ The drawback of this approach is that it is computationally very expensive to calculate 3D coordinates. In a task where only the 3D information of some objects is essential, like tracking and surveillance, the computed additional information is redundant. Using antenna arrays or sonar systems one traditionally only gets bearing data to process.

In this work we focus on the task of passive localization and tracking of multiple non-cooperative emitting objects. To solve this task we derive a general sensor model and scenario. The localization, tracking and fusion of the sensor data is done by a sequential Monte Carlo implementation of a probability hypothesis density (PHD) filter. We use this iterative approach to avoid the time-consuming data association, as was used by Hue et al.⁴ for a single moving sensor. In our approach we obtain unique but imprecise data from the antenna array system (detection of emitting sources) and ambiguous but precise data from the optical sensor. In this paper we present how to fuse both kinds of data to get a unique and precise localization and tracking result.

Further author information: (Send correspondence to M. Schikora)

M. Schikora: E-mail: marek.schikora@fkie.fraunhofer.de, Telephone: +49 228 9435-816

D. Bender: E-mail: daniel.bender@fkie.fraunhofer.de

W. Koch: E-mail: wolfgang.koch@fkie.fraunhofer.de

D. Cremers: E-mail: cremers@tum.de

2. SENSOR AND TARGET MODEL

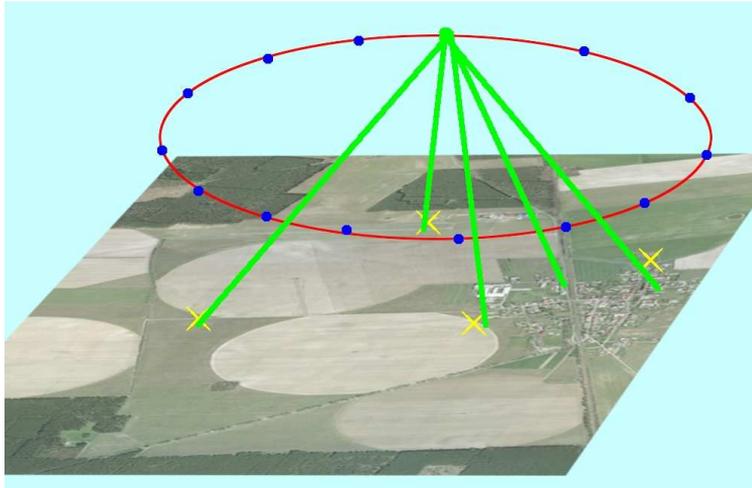


Figure 1. Typical scenario used in our simulations and real world experiments. A possible observer path is illustrated by the red ellipse, whereby the blue dots represent the discrete measurement points. The yellow crosses are objects of interest in this scene, which will be localized and tracked from the algorithm by bearing data. The latter is represented by green rays which point in the direction resulting from processing the sensor outputs.

The scenario considered in this work is the following: An observer (UAV) with a camera and a antenna array performs a (non-linear) flight maneuver over a region. Possible objects of interest are detected by some sensor processing and the corresponding bearing angles are calculated. Additionally, measurements from false alarms and clutter are produced by this process and have to be handled by the proposed algorithm to localize and track multiple objects. A typical scenario can be seen in Figure 1. The proposed algorithm is general enough to be easily adapted to another sensor-object geometry. In general we assume the following time scheme: First the antenna system computes bearing data. This data is used to steer the optical system. In the image possible targets are detected and corresponding bearing data is produced. Subsequently a new measurement from the antenna system is computed. This alternating order is kept for the whole scenario. For every new measurement the corresponding sensor model is chosen and the data is processed in the multi-sensor PHD filter.

We assume that the objects move with respect to a linear constant velocity model, so that the object state is a six-dimensional vector

$$\mathbf{x} = (x, y, z, v_x, v_y, v_z)^T \in \mathbb{R}^6, \quad (1)$$

with three entries for the position and three entries for the velocities in a Cartesian coordinate system. A local measurement is described by a two-dimensional vector

$$\mathbf{z}_{\text{obs}} = (\alpha_{\text{obs}}, \varepsilon_{\text{obs}})^T \in \mathbb{R}^2, \quad (2)$$

whereby $\alpha \in (-\pi, +\pi]$ is azimuth and $\varepsilon \in [0, \pi/2]$ is elevation in the local coordinate system of the observer.

2.1 Generating bearing measurements from an antenna array

Possible objects of interest are detected by some sensor processing. Using an antenna array, a lot of direction finding algorithms exists,⁵⁻⁷ which produce local bearing data in the coordinate system of the sensor. In order to localize and track an object over the time we need to transform local measurements in a global coordinate system. This is done by the following three steps:

1. Transform to polar coordinates:

$$u = \sin(\alpha_{\text{obs}}) \sin(\varepsilon_{\text{obs}}) \quad (3)$$

$$v = \cos(\alpha_{\text{obs}}) \sin(\varepsilon_{\text{obs}}). \quad (4)$$

2. Rotate the bearing ray $\boldsymbol{\rho}_{\text{obs}}$ according to the attitude of the observer

$$\boldsymbol{\rho}_{\text{obs}} = (v, u, \sqrt{1 - u^2 - v^2})^T \quad (5)$$

$$\boldsymbol{\rho}_{\text{global}} = \mathbf{R}^{-1}(\phi, \theta, \psi)\boldsymbol{\rho}_{\text{obs}}, \quad (6)$$

with ϕ, θ and ψ denoting the heading, pitch and roll of the observer and $\mathbf{R}(\phi, \theta, \psi)$ denotes a suitable rotation matrix for this angles.

3. Calculate the measurements in a global coordinate system as follows:

$$\alpha_{\text{global}} = \arctan\left(\frac{\rho_{\text{global}}(2)}{\rho_{\text{global}}(1)}\right) \quad (7)$$

$$\varepsilon_{\text{global}} = \arctan\left(\frac{\sqrt{\rho_{\text{global}}^2(1) + \rho_{\text{global}}^2(2)}}{\rho_{\text{global}}(3)}\right), \quad (8)$$

leading to

$$\mathbf{z} = (\alpha_{\text{global}}, \varepsilon_{\text{global}}). \quad (9)$$

2.2 Generating bearing measurements from an optical system

In a camera system possible targets are detected by applying an object detection strategy on the gained images.^{8,9} These objects are represented by a central pixel \mathbf{q} or rather by its homogeneous counterpart \mathbf{q}' that is used to calculate the global bearing angles \mathbf{z} :

1. Determine an arbitrary 3D point $\mathbf{Q} = (x_{\mathbf{Q}}, y_{\mathbf{Q}}, z_{\mathbf{Q}})^T$ on the projection ray corresponding to the pixel $\mathbf{q} = (x_{\mathbf{q}}, y_{\mathbf{q}})^T$ or rather $\mathbf{q}' = (x'_{\mathbf{q}}, y'_{\mathbf{q}}, 1)^T$ by decomposing the projection matrix \mathbf{P} as follows:

$$\lambda \mathbf{q}' = \mathbf{P} \begin{bmatrix} \mathbf{Q} \\ 1 \end{bmatrix} = [\mathbf{A} \mid \mathbf{a}] \begin{bmatrix} \mathbf{Q} \\ 1 \end{bmatrix} = \mathbf{A}\mathbf{Q} + \mathbf{a}, \quad (10)$$

with λ denoting the freely selectable factor of the homogeneous representation of \mathbf{q} . This leads to the projection ray:

$$\mathbf{Q} = -\mathbf{A}^{-1}\mathbf{a} + \lambda \mathbf{A}^{-1}\mathbf{q}', \quad (11)$$

whereby $-\mathbf{A}^{-1}\mathbf{a}$ represents the projection center \mathbf{C} of the camera. Setting the height of the point \mathbf{Q} to valid value, e.g. in the presented aerial case $z_{\mathbf{Q}} = 0$, leads to a value for λ and subsequent to the complete point \mathbf{Q} .

2. Calculate the global bearing angles \mathbf{z} between the point $\mathbf{Q} = (x_{\mathbf{Q}}, y_{\mathbf{Q}}, z_{\mathbf{Q}})^T$ and the projection center $\mathbf{C} = (x_{\mathbf{C}}, y_{\mathbf{C}}, z_{\mathbf{C}})^T$ of the camera:

$$\alpha_{\text{global}} = \arctan\left(\frac{x_{\mathbf{Q}} - x_{\mathbf{C}}}{y_{\mathbf{Q}} - y_{\mathbf{C}}}\right), \quad (12)$$

$$\varepsilon_{\text{global}} = \arctan\left(\frac{z_{\mathbf{Q}} - z_{\mathbf{C}}}{\sqrt{(x_{\mathbf{Q}} - x_{\mathbf{C}})^2 + (y_{\mathbf{Q}} - y_{\mathbf{C}})^2}}\right) + \frac{\pi}{2}. \quad (13)$$

For global measurements gained from the antenna system and the optical system we write $\mathbf{z}_{\text{ant}} := \mathbf{z}$ and $\mathbf{z}_{\text{cam}} := \mathbf{z}$, respectively.

3. RANDOM FINITE SET MODEL

In a single-object system, the state and measurement at time k are represented as two random vectors of possibly different dimensions. These vectors evolve in time, but maintain their initial dimension. However, this is not the case in a multi-object system. Here the multi-object state and multi-object measurement are two collections of individual objects and measurements. The number of these may change over time and lead to another dimension of the multi-object state and multi-object measurement. Furthermore, there exist no ordering for the elements of the multi-object state and measurement. Using the theory proposed by Mahler,¹⁰ the multi-object state and measurement are naturally represented as finite subsets X_k and Z_k defined as follows:

Let $N(k)$ be the number of objects at time-stamp k , which are located at $\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,N(k)}$ in the single-object state space E_S , e.g. \mathbb{R}^d then,

$$X_k = \{\mathbf{x}_{k,1}, \dots, \mathbf{x}_{k,N(k)}\} \in \mathcal{F}(E_S) \quad (14)$$

is the multi-object state, where $\mathcal{F}(E_S)$ denotes the collection of all finite subsets of the space E_S . Analogous to this, we define the multi-object measurement at the time-step k

$$Z_k = \{\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,M(k)}\} \in \mathcal{F}(E_O), \quad (15)$$

assuming $M(k)$ measurements $\mathbf{z}_{k,1}, \dots, \mathbf{z}_{k,M(k)}$ in the single-object space E_O that correspond to real targets and clutter. The sets X_k and Z_k are also called random finite sets, for which the first moment, or probability hypothesis density, is the analog of the expectation of a random vector. The integral value of the PHD over a given region in state space leads to the expected number of objects within this region. We define $D(\mathbf{x}_k|Z^k)$ as the PHD associated with the multi-object posterior $p(X_k|Z^k)$ at a time-step k , with Z^k denoting the accumulated measurement from the time-steps 1 to k . The PHD filter consists of two steps: prediction and update.

The prediction can be realized through the following equation:

$$D(\mathbf{x}_k|Z^{k-1}) = b(\mathbf{x}_k) + \int [p_s(\mathbf{x}_{k-1})p(\mathbf{x}_k|\mathbf{x}_{k-1}) + b(\mathbf{x}_k|\mathbf{x}_{k-1})] D(\mathbf{x}_{k-1}|Z^{k-1})d\mathbf{x}_{k-1}, \quad (16)$$

where $b(\mathbf{x}_k)$ denotes the intensity function of spontaneous birth of new objects, $b(\mathbf{x}_k|\mathbf{x}_{k-1})$ describes the intensity function of the random finite set of objects spawned from the previous state \mathbf{x}_{k-1} . $p_s(\mathbf{x}_{k-1})$ is the probability that the object still exists at the time-step k given its previous state \mathbf{x}_{k-1} , and $p(\mathbf{x}_k|\mathbf{x}_{k-1})$ is the transition probability density of the individual objects. The update equation can be written as

$$D(\mathbf{x}_k|Z^k) \cong F(Z_k|\mathbf{x}_k)D(\mathbf{x}_k|Z^{k-1}), \quad (17)$$

$$F(Z_k|\mathbf{x}_k) = 1 - p_D(\mathbf{x}_k) + \sum_{\mathbf{z} \in Z_k} \frac{p_D(\mathbf{x}_k)p(\mathbf{z}|\mathbf{x}_k)}{\lambda c(\mathbf{z}) + \int p_D(\mathbf{x}_k)p(\mathbf{z}|\mathbf{x}_k)D(\mathbf{x}_k|Z^{k-1})d\mathbf{x}_k}, \quad (18)$$

with $p_D(\mathbf{x}_k)$ denoting the probability of the detection of the state \mathbf{x}_k . Furthermore, $p(\mathbf{z}|\mathbf{x}_k)$ is the measurement likelihood, $c(\mathbf{z})$ the probability distribution for every clutter point and λ is the average number of clutter points per scan.

4. TRACKING AND LOCALIZATION

We implemented the theory described in the last section with a sequential Monte Carlo method, also known as particle filter. The first attempt to implement this technique for a tracking system, using the random finite set theory, was presented by Vo et al.¹¹ This was only done for a simulated scenario and not optimized for multi-sensor tracking and localization using bearing data. Following the definition in Equation (1) the state of an individual object will be represented by $\mathbf{x}_k \in \mathbb{R}^6$. Each measurement $\mathbf{z}_{k,t} \in \mathbb{R}^2$ is represented analogous, c.f. (9), with $t \in \{\text{ant, cam}\}$. For the sake of simplicity, we assume that the object motion model of each target is linear with a constant velocity. With this, the object state prediction can be written as:

$$\mathbf{x}_k = \Phi(k, k-1)\mathbf{x}_{k-1} + \mathbf{s}_k, \quad (19)$$

where \mathbf{s}_k denotes the zero-mean Gaussian white process noise and

$$\Phi(k, k') = \begin{pmatrix} \mathbf{I}_3 & (k - k')\mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{pmatrix} \quad (20)$$

the transition matrix from time-step k' to k . Since the measurements are two-dimensional, the likelihood function is given by:

$$p(\mathbf{z}|\mathbf{x}) = \frac{1}{2\pi|\Sigma_t|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\mathbf{z} - f(\mathbf{x}))^T \Sigma_t (\mathbf{z} - f(\mathbf{x}))\right), \quad (21)$$

with the covariance matrix of the measurement noise of the corresponding sensor system Σ_t and the noise-free measurement equation

$$f(\mathbf{x}) = \begin{pmatrix} \arctan\left(\frac{\mathbf{x}(1) - \mathbf{x}_{\text{obs}}(1)}{\mathbf{x}(2) - \mathbf{x}_{\text{obs}}(2)}\right) \\ \frac{\pi}{2} + \arctan\left(\frac{\mathbf{x}(3) - \mathbf{x}_{\text{obs}}(3)}{\sqrt{(\mathbf{x}(1) - \mathbf{x}_{\text{obs}}(1))^2 + (\mathbf{x}(2) - \mathbf{x}_{\text{obs}}(2))^2}}\right) \end{pmatrix}. \quad (22)$$

For the imprecise bearing data gained from the antenna array we, use the covariance matrix Σ_{ant} and a very small clutter rate model, represented through λ_{ant} and $c_{\text{ant}}(\mathbf{z})$. For the optical system we use a covariance matrix Σ_{cam} designed for small errors and a high clutter model represented through λ_{cam} and $c_{\text{cam}}(\mathbf{z})$.

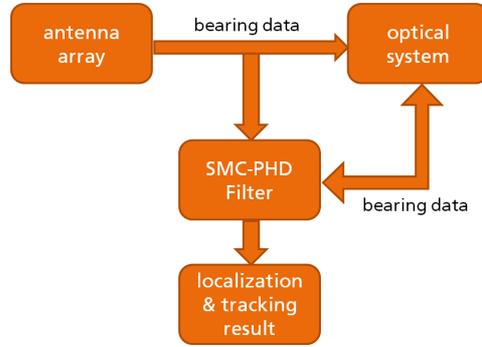


Figure 2. Workflow of the proposed algorithm. Bearing data from the antenna array is processed by the SMC-PHD filter. Additionally the actual localization result and this bearing data is used to steer an optical system. The optical bearing data is fused within the filter to get a precise and unique result.

Using the particle filter we can model the birth process $b(\mathbf{x}_k)$ as a uniformly distributed set of new particles with small weights. Let L_k denote the number of particles at a time-step k , and let J_k denote the number of the newly created particles at this time-step. The sequential Monte Carlo (SMC) PHD filter can be summarized as follows:

At a given time-step $k \geq 0$, let $\{\mathbf{x}_k^i, w_k^i\}_{i=1}^{L_k}$ represent a particle-based approximation of the PHD. Then four steps are performed: prediction, detection, update and resample.

1. **Predict** the state of each particle pursuant to a given dynamic model. For $i = 1, \dots, L_{k-1}$, transform each particle to $\tilde{\mathbf{x}}_k^i$ according to (19) and keep the weights unchanged. Uniformly sample a new set of particles in the image and set

$$\tilde{w}_k^i = \frac{1}{J_k}, \text{ for } i = L_{k-1} + 1, \dots, L_{k-1} + J_k. \quad (23)$$

2. **Detect** possible objects in the scene. This step depends on the used sensor system. The resulting bearing measurements are formulated analogous to (9).

3. **Update** the state of each particle by a new set of measurements. For each measurement $\mathbf{z} \in Z_k$ (c.f. (9)) compute

$$C_k(\mathbf{z}) = \sum_{i=1}^{L_{k-1}+J_k} p_D(\mathbf{x}_k^i) p(\mathbf{z}|\mathbf{x}_k^i) \tilde{w}_k^i. \quad (24)$$

For all particles $i = 1, \dots, L_{k-1} + J_k$ update their weights

$$\hat{w}_k^i = \left(1 - p_D(\tilde{\mathbf{x}}_k^i) + \sum_{\mathbf{z} \in Z_K} \frac{p_D(\mathbf{x}_k^i) p(\mathbf{z}|\mathbf{x}_k^i)}{\lambda_t c_t(\mathbf{z}) + C_k(\mathbf{z})} \right) \tilde{w}_k^i. \quad (25)$$

4. **Resample** the particle set. Firstly the target number has to be estimated as follows:

$$\hat{N}_k = \sum_{i=1}^{L_{k-1}+J_k} \hat{w}_k^i. \quad (26)$$

Initialize the cumulative probability with $c_1 = 0$ and set

$$c_i = c_{i-1} + \frac{\hat{w}_k^i}{\hat{N}_k}, \text{ for } i = 2, \dots, L_{k-1} + J_k. \quad (27)$$

Draw a uniformly distributed starting point a_1 from the interval $[0, L_k^{-1}]$.

For $j = 1, \dots, L_k$,

$$a_j = a_1 + L_k^{-1} \cdot (j - 1) \quad (28)$$

while $a_j > c_i, i = i + 1$. End while.

$$\mathbf{x}_k^j = \tilde{\mathbf{x}}_k^i \quad (29)$$

$$w_k^j = L_k^{-1} \quad (30)$$

Rescale the weights by \hat{N}_k to get a new particle set $\left\{ \mathbf{x}_k^i, \hat{N}_k/L_k \right\}_{i=1}^{L_k}$.

After each time-step k , we generate a particle cloud $\left\{ \mathbf{x}_k^i, \hat{N}_k/L_k \right\}_{i=1}^{L_k}$, which represents the PHD. To estimate the correct object states from this cloud we have to perform a clustering. The SMC-PHD filter estimates the number of objects for every time-step, so it is possible to use a clustering technique, which requires the number of clusters, e.g. k -means clustering.¹² However the estimated object number from the PHD has a high variance.¹³ To deal with this problem, we use in our experiments the adaptive resonance theory (ART) clustering,¹⁴ which estimates the number of clusters automatically, with a distance parameter as predefined user input. With this kind of clustering we are robust against estimation errors in the number of objects. In fact we only use a subset

$$S \subset \left\{ \mathbf{x}_k^i, w_k^i \right\}_{i=1}^{L_k}, \quad (31)$$

with

$$\left(\mathbf{x}_k^i, w_k^i \right) \in S \text{ if } w_k^i \geq \tau. \quad (32)$$

5. RESULTS

In this section, we present localization and tracking results generated by our algorithm. As already presented in our previous work¹⁵ the PHD-filter is very efficient for bearings only tracking, so we are only interested in showing the benefit of sensor data fusion with this approach. We will demonstrate that using both sensors in comparison to use only an antenna array can improve the accuracy strongly. This will be done with simulated data. Following this, we will show real-world results. For all experiments we used 5000 particles in the SMC-PHD filter. All other parameters of the proposed strategy are sensor-dependent and can be estimated, if they are not known.

5.1 Simulated data

Similar to Figure 1 we build up a scenario where an observer performs a half-circle flight over an region of interest. The observer is equipped with a antenna array, specialized for direction finding of communication emitters, and an optical system, which can be steered. The aim is to localize and track non-corporative emitters. Both sensors are modeled with realistic error models and a individual clutter and detection rate. An emitting target is placed in the scene. Additional false targets, which can only be detected by the optical system, are also installed. The quality of the proposed method is measured as localization accuracy. Figure 3 displays localization results for 1000 Monte Carlo simulations for this scenario. Remarkable is the big performance benefit using the proposed method to fuse both sensor types. We achieve an improvement in accuracy by a factor of about six.

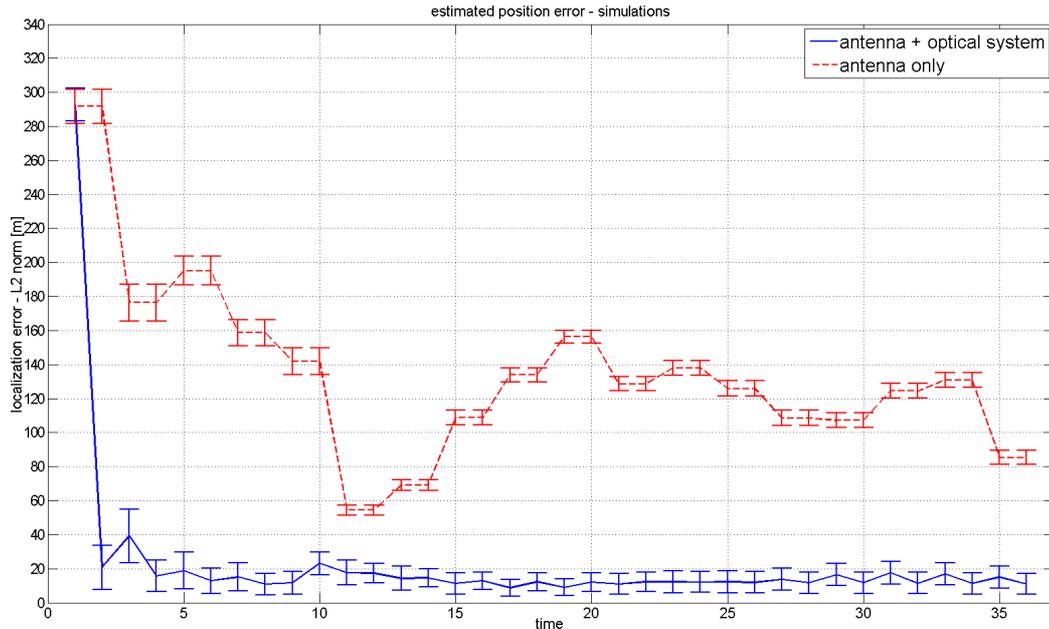


Figure 3. Estimated position error over time for 1000 Monte Carlo runs. Curves show mean value and standard derivation for every time-step. Both are generated by a SMC-PHD filter.

5.2 Real data

Additionally to simulated data we test our algorithm with real data gained from an optical system and a three-element antenna array designed for UAVs. In this scenario we have three stationary targets, which are emitting signals and one that is inertially moving. Figure 4 illustrates the tracking and localization result for those targets. As easily can be recognized the proposed method works well, even for real data. All four targets are localized and tracked correctly. For every time-step the estimated target number is four. The challenge in this data lies in its non-Gaussian error distribution and additional grating lobes effects (i.e. ambiguous array data).

6. CONCLUSION

In this paper we demonstrated an approach, that is able to localize and track an unknown number of targets, given bearing data from heterogeneous sensors mounted on a UAV. In order to fuse this data a sequential Monte Carlo implementation of a probability hypothesis density filter is used. This filter is able to handle the non-linear measurements equation and the data association problem in this multi-target scenario. With a Monte Carlo simulation we could show that the fusion of optical bearing measurements and bearing data from antenna array can improve the localization accuracy by a factor of six in comparison to results achieved using only an antenna array. In addition to this statement we presented real world results, which were generated by an antenna array system and an optical system mounted on a UAV. We were able to localize and track all targets correctly in

the presence of clutter and high bearing errors. These experiments make clear that even with strongly corrupted input data very good results can be achieved using the proposed strategy to fuse bearing data from heterogeneous sensors.

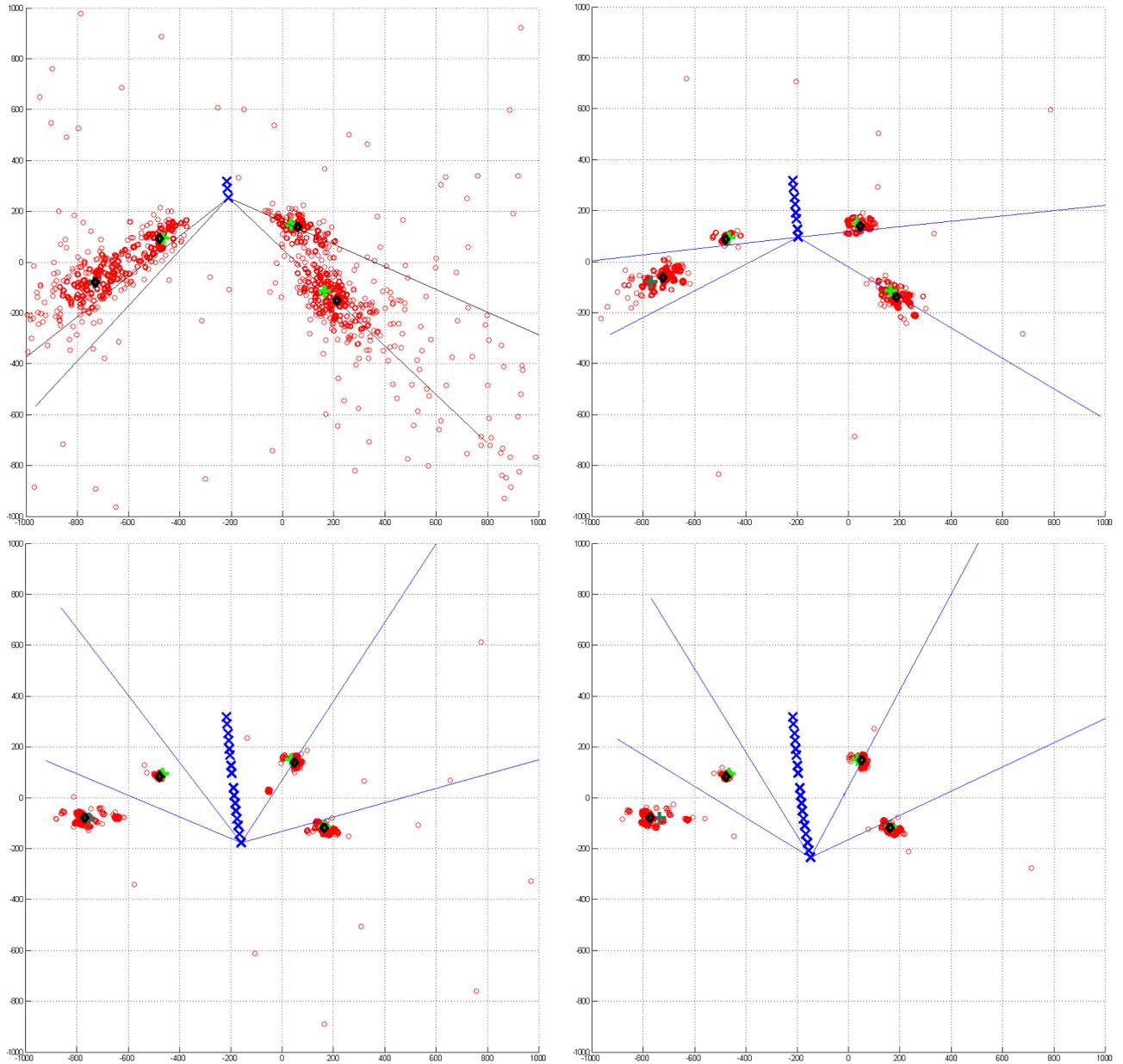


Figure 4. Particle evolution and tracking results for real data. Blue x displays the current observer position. Green + is the true target position for this time-step. The dark green target moves inertially. The black diamonds represent the current tracking result. The red circles illustrate the particle cloud. For a better perspicuity this Figure only shows a projection on the xy -plane of the six-dimensional estimation problem.

REFERENCES

- [1] Hartley, R. I. and Zisserman, A., [*Multiple View Geometry in Computer Vision*], Cambridge University Press, ISBN: 0521540518, second ed. (2004).
- [2] Klaus, A., Sormann, M., and Karner, K., “Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure,” in [*ICPR*], (2006).
- [3] Pock, T., Schoenemann, T., Graber, G., Bischof, H., and Cremers, D., “A convex formulation of continuous multi-label problems,” in [*European Conference on Computer Vision (ECCV)*], (October 2008).
- [4] Hue, C., Cadre, J.-P. L., and Pérez, P., “Sequential monte carlo methods for multiple target tracking and data fusion,” *IEEE Trans. Signal Processing* **50**(2), 309–325 (2002).
- [5] Godara, L. C., “Application of antenna arrays to mobile communications, Part II: Beam-forming and direction-of-arrival considerations,” *Proc. IEEE* **85**, 1195–1245 (Aug. 1997).
- [6] Schell, S. V. and Gardner, W. A., “High-resolution direction finding,” in [*Handbook of Statistics, Signal Processing and its Applications*], Bose, N. K. and Rao, C. R., eds., **10**, ch. 18, 755–817 (1993).
- [7] Schmidt, R. O., “Multiple emitter location and signal parameter estimation,” *Proc. RADC Spectrum Estimation Workshop, Griffith AFB*, 243–258 (1979).
- [8] Leibe, B., Leonardis, A., and Schiele, B., “Robust object detection with interleaved categorization and segmentation,” *International Journal of Computer Vision Special Issue on Learning for Recognition and Recognition for Learning* **77**(1–3), 259–289 (2008).
- [9] Schikora, M., “Global optimal multiple object detection using the fusion of shape and color information,” in [*Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*], (August 2009).
- [10] Mahler, R., “Multitarget bayes filtering via first-order multitargets moments,” *IEEE Trans. Aerosp. Electron. Syst.* **39**(4), 1152–1178 (2003).
- [11] Vo, B.-N., Singh, S., and Doucet, A., “Sequential Monte Carlo methods for multi-target filtering with random finite sets,” *IEEE Trans. Aerosp. Electron. Syst.* **41**(4), 1224–1245 (2005).
- [12] Lloyd, S., “Least squares quantization in pcm,” *IEEE Trans. Inform. Theory* **28**(2), 129–137 (1982).
- [13] Mahler, R., “PHD filters of higher order in target number,” *IEEE Trans. Aerosp. Electron. Syst.* **43**(4), 1523–1543 (2007).
- [14] Carpenter, G. and Grossberg, S., “ART 2: Self-organizing stable category recognition codes for analog input patterns,” *Applied Optics* **26**(23), 4919–4930 (1987).
- [15] Schikora, M., Bender, D., Cremers, D., and Koch, W., “Passive multi-object localization and tracking using bearing data,” in [*13th International Conference on Information Fusion (FUSION)*], (July 2010).