Simultaneous Navigation and Synthetic Aperture Radar Focusing

Zoran Sjanic and Fredrik Gustafsson Fellow, IEEE

Abstract—Synthetic Aperture Radar (SAR) equipment is a radar imaging system that can be used to create high resolution images of a scene by utilising the movement of a flying platform. Knowledge of the platform’s trajectory is essential to get good and focused images. An emerging application field is real-time SAR imaging using small and cheap platforms, where estimation errors in navigation systems imply unfocused images. This contribution investigates a joint estimation of the trajectory and SAR image. Starting with a nominal trajectory, we successively improve the image, by optimizing a focus measure and updating the trajectory accordingly. The method is illustrated using simulations using typical navigation performance of an unmanned aerial vehicle. One real data set is used to show feasibility, where the result indicates that in particular the azimuth position error is decreased as the image focus is iteratively improved.

Keywords: Optimisation, navigation, Synthetic Aperture Radar, auto-focusing

I. INTRODUCTION

A general method for creating high resolution radar images from low resolution radar data, or real aperture images, is to use relative motion between radar antenna and the imaged scene and integrate all the partial real aperture images taken along the flown trajectory [1]. Traditionally, this operation is performed in the frequency domain using FFT-like methods, [2]–[10], due to their computational efficiency. The common denominator of these methods is that they assume that the aircraft’s (or antenna’s) flown path is linear, i.e. without cross-track velocity and acceleration and with constant along-track velocity, and that is generally not the case in practice. If the trajectory is not linear the integration will result in an unfocused image. It is possible to partly correct for the deviation from the nonlinear trajectory but then the methods become computationally inefficient. Another approach is to perform integration in time domain by means of solving the back-projection integral [11].

Even in this process it is assumed that the radar antenna’s flown path is linear with constant altitude and heading, but the method can be extended to non-linear tracks as well. However exact inversion is not guaranteed. More details about this method will be provided in Section II. The main disadvantage of this method is the large amount of operations required to create an image, where the complexity is proportional to $O(NKM)$ for $K \times M$ pixels image using an aperture with $N$ positions. Here, $O(\cdot)$ denotes “in the order of”. However, by means of coordinate transformation, an approximation to exact back-projection can be performed, which is called Fast Factorised Back-projection, see [12]. The complexity of this algorithm is proportional to $KM \log N$ operations, which for large $N$ implies an important saving. With this faster algorithm it should be possible to create images in real time, possibly in dedicated hardware. Since back-projection algorithms are dependent on exact knowledge of the antenna’s position in order to get focused images, the image focus can be measured and used for estimation of the trajectory. An example of this is depicted in Figure 1, where 10 point targets are imaged. In Figure 1a, a linear path is simulated, which results in a perfectly focused image. In the other three images the variation in cross-track position was added as $A \sin(2\pi nk/N)$ where $A = \{0.5, 1, 1.5\} [\text{m}], \; n = 1.5, \; k = 1 : N$ and the images are created with an assumption that the path was linear. The range to the scene is 2182 m. This gives unfocused images as depicted. Traditional methods for auto-focusing are mostly open-loop type methods where either SAR images or raw radar data are used, [13]–[29]. The significant common denominator for all these methods is that the image is created with assumptions on linear flight trajectory and focusing is
done afterwards in an open-loop way discarding eventual flight path information. This is a consequence of the off-line image generating process where the trajectory is no longer interesting. In the setup where SAR images are generated on-line an idea is to use information from the image focus and navigation system, like measured accelerations, together and in a sensor fusion framework try to obtain the best solution to both image focusing and navigation simultaneously. This approach can be particularly useful in the Unmanned Aerial Vehicles (UAV) application where navigation system has lower quality. This implies that the estimated trajectory has larger error giving unfocused SAR images. In the view of this approach for focusing of SAR images and estimating the trajectory, the problem is related to Inverse Synthetic Aperture Radar (ISAR) where the radar is stationary and the task is to image a moving target [30]–[32]. Even in this case, the target’s motion is arbitrary and, in addition, usually no other measurements of the motion are available, but the focusing problem is in principle the same [33]–[35].

Also, another closely related and well known problem is Simultaneous Localisation and Mapping (SLAM), [36], [37], where a map of the unknown environment is estimated jointly with the platform’s position. The SLAM problem has been well studied during recent years and many different solution methods have been proposed. One method that has been quite successful is to solve the SLAM problem in a sensor fusion framework. In the SAR application, the map of the environment from SLAM, is the unknown scene that is imaged and can be seen as the two dimensional map of point reflectors. The problem of positioning the platform is the same in SLAM. However, the main difference is that we consider a non-parametric SAR image rather than a parametric map of point reflectors, that would be a too restrictive assumption in SAR imaging. That is, though there are many conceptual similarities of joint navigation and mapping, the state of the art algorithms cannot be applied here.

This contribution applies a sensor fusion framework, where the SAR image together with a focus measure is interpreted as a “sensor” that contains information about the position of the platform. The image creating and auto-focusing methods described above can be illustrated as in Figure 2. The method based on sensor fusion can be implemented in a centralised or decentralised manner. In this work we focus on the decentralised manner.

The outline is as follows. Section II summarizes notation and makes a high-level mathematical formulation of the approach. Section III introduces the navigation framework and system and measurement models used. Section IV describes the image focus measures that will be used in the auto-focus procedure. In Section V an optimisation framework and methods are introduced and their usage is explained. Numerical examples for the simulated images are covered in Section VI and for the real SAR data in Section VII. Finally, conclusions and future work are discussed in Section VIII.

II. Notation and Problem Formulation

The main notation used in the paper is defined in Table I. Let the complex range compressed raw radar data be denoted $z_t(R_t)$, also called Real Aperture Radar (RAR) image, where $t$ is the time index, and where $z_t(R_t)$ denotes the returned radio energy at time $t$ corresponding to distance $R_t$ to the scene from the radar. The range $R_t$ is calculated as $R_t = \frac{1}{2} c_0 \tau$ where $c_0$ is the speed of light and $\tau$ is the radar pulse total propagation time to the scene. Note that $t$ and $\tau$ are different times, usually called slow and fast time respectively. Further, let $x_t$ denote the state vector of the platform, which includes position and velocity (radar pose). The backprojection method of producing the images, see Figure 3, can be expressed as integration per image pixel. For each pixel $(i, j)$ in a complex valued image $I$, the total energy from each radar pulse is integrated by summing all the values from the raw data given the range from the platform to the point in the scene corresponding to the pixel $(i, j)$. This can be expressed as

$$I_{ij} = \sum_{t=1}^{N} z_t(||p_t - s_{ij}||_2)$$  \hspace{1cm} (1)

where $p_t$ is the 3D position of the radar and $s_{ij}$ is the coordinate of the scene point that is mapped to the pixel $(i, j)$ in the image. Now, for a SAR system on the UAV platform, the pose cannot be assumed to be known. Instead, we have access to an estimated position $\hat{p}_t$, and the estimated SAR
Table I: Definition of the mathematical symbols and operators.

<table>
<thead>
<tr>
<th>Symbol/Operator</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>Time index denoting platform’s time, used for indexing motion states like position or velocity and azimuth direction of the RAR image [s]</td>
</tr>
<tr>
<td>( x_t, y_t, z_t )^T</td>
<td>Platform’s 3D position [m]</td>
</tr>
<tr>
<td>( \dot{x}_t, \dot{y}_t, \dot{z}_t )^T</td>
<td>Platform’s 3D velocity [m/s]</td>
</tr>
<tr>
<td>( \ddot{x}_t, \ddot{y}_t, \ddot{z}_t )^T</td>
<td>Platform’s 3D acceleration [m/s^2]</td>
</tr>
<tr>
<td>( x_t )</td>
<td>Platform’s states (consisting of positions, velocities and accelerations)</td>
</tr>
<tr>
<td>( P_i )</td>
<td>Platform’s states’ covariance</td>
</tr>
<tr>
<td>( [w_i^x, w_i^y, w_i^z]^T )</td>
<td>Process noise [m/s^2]</td>
</tr>
<tr>
<td>( Q_t )</td>
<td>Process noise covariance</td>
</tr>
<tr>
<td>( y_t )</td>
<td>Vector of measurements</td>
</tr>
<tr>
<td>( e_t )</td>
<td>Vector with measurement noise</td>
</tr>
<tr>
<td>( V_t )</td>
<td>Covariance of the measurement noise</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>Sampling time [s]</td>
</tr>
<tr>
<td>( f(\cdot), h(\cdot) )</td>
<td>Functions describing system dynamics and measurements respectively</td>
</tr>
<tr>
<td>( F, G )</td>
<td>Matrices describing system dynamics</td>
</tr>
<tr>
<td>( H )</td>
<td>Matrix describing measurements</td>
</tr>
<tr>
<td>( \theta )</td>
<td>Vector of estimated parameters</td>
</tr>
<tr>
<td>( z_t(R_t) )</td>
<td>Complex raw radar data (range compressed)</td>
</tr>
<tr>
<td>( I )</td>
<td>(Complex) SAR image</td>
</tr>
<tr>
<td>( I_{ij} )</td>
<td>Pixel ((i,j)) in the SAR image</td>
</tr>
<tr>
<td>( R_t )</td>
<td>Slant range from the platform at time ( t ) to some point in the imaged scene (usually midpoint) used for indexing range direction of the RAR image [m]. This range is calculated from each radar pulse as ( R_t = \frac{c_0 \tau}{2} ), where ( c_0 ) is the speed of light and ( \tau ) is the radar pulse total traveling time</td>
</tr>
<tr>
<td>( F )</td>
<td>General image focus measure</td>
</tr>
<tr>
<td>( E_t )</td>
<td>Entropy 1 image focus measure</td>
</tr>
<tr>
<td>( E_2 )</td>
<td>Entropy 2 image focus measure</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>Weighting coefficient</td>
</tr>
<tr>
<td>( \widehat{x} )</td>
<td>Estimate of ( x )</td>
</tr>
<tr>
<td>( I_i )</td>
<td>Identity ( i \times i ) matrix</td>
</tr>
<tr>
<td>( 0_{i \times i} )</td>
<td>( i \times j ) matrix with all zero elements</td>
</tr>
<tr>
<td>( O(N) )</td>
<td>Big Ordo notation, in the order of ( N )</td>
</tr>
<tr>
<td>( N(\mu, P) )</td>
<td>Gaussian (normal) distribution with mean ( \mu ) and covariance ( P )</td>
</tr>
<tr>
<td>( \arg \min_x A(x) )</td>
<td>Minimising argument of ( A(x) ) wrt ( x )</td>
</tr>
<tr>
<td>( \partial f(x)/\partial x )</td>
<td>Partial derivative of ( f(x) ) wrt ( x )</td>
</tr>
<tr>
<td>( \nabla f(x) )</td>
<td>Gradient of scalar function ( f ) wrt vector ( x )</td>
</tr>
<tr>
<td>(</td>
<td></td>
</tr>
<tr>
<td>(</td>
<td></td>
</tr>
<tr>
<td>(</td>
<td></td>
</tr>
</tbody>
</table>

In equation (2), we estimate the SAR image as follows:

\[
\hat{I}_{ij} = \sum_{t=1}^{N} z_t(\|\hat{p}_t - s_{ij}\|_2). \tag{2}
\]

This estimated SAR image will be out of focus, since all the contributions from raw data will now be scattered due to the error in position estimate, see Figure 1 for an example of this.

The key idea in this contribution is to perform a parametric focusing. To enable this, we will make use of a focus measure \( \mathcal{F}(\hat{I}) \), with the property that

\[
\mathcal{F}(\hat{I}) > \mathcal{F}(I^0), \quad \hat{I} \neq I^0, \tag{3a}
\]

\[
x_{1:N}^0 = \arg \min_{x_{1:N}} \mathcal{F}(\hat{I}) \tag{3b}
\]

where \( I^0 \) denotes the true SAR image and \( x_{1:N}^0 \) the true state sequence.

We will, however, not optimize the focus w.r.t. the pose blindly. We will optimize focus jointly with the filtering problem, in that the states obey the state dynamics and observations as well as possible.

As already noted, building up the image of size \( KM \) pixels with an aperture of \( N \) time points requires a huge computational effort (\( O(NKM) \)). It may seem that an outer loop that performs focusing will increase the computational burden at least an order of magnitude more. However, we will show that the gradient of the focus measure can be computed efficiently.

First, let \( \theta \) denote the parameters used to describe the state trajectory. It may be the whole position trajectory, \( \theta = p_{1:N}, \) or the initial states only, \( \theta = x_0, \) as two special cases. Then, define \( I_{t,ij} = z_t(R_{t}^{ij}) \) and, as before, \( R_{t}^{ij} = \|p_t - s_{ij}\|_2, \) so that (1) can be written \( I_{ij} = \sum_{t=1}^{N} I_{t,ij}. \) Then, using the chain rule, we get

\[
\frac{\partial \mathcal{F}(\theta)}{\partial \theta} = \frac{\partial \mathcal{F}}{\partial I_{t,ij}} \frac{\partial I_{t,ij}}{\partial R_{t}^{ij}} \frac{\partial R_{t}^{ij}}{\partial \theta} \approx \frac{\partial \mathcal{F}}{\partial I_{t,ij}} \frac{\partial I_{t,ij}}{\partial R_{t}} \frac{\partial R_{t}}{\partial \theta} \tag{4}
\]

The approximation that allows the dependence of \( ij \) on the range to be removed is based on small scene assumption: the range \( R_t \) to the middle of the scene does not vary much over the whole scene indexed by \( ij \). There are several advantages of using this approximation:

- The first and third partial derivatives can be derived analytically.
- The second one can be computed almost for free from any SAR algorithm, and requires only \( KM \) extra memory cells and one or two more SAR image computations.
depending on how the numerical gradient \( \partial I / \partial R \) is approximated.

More concrete formulations of (4) come later when the focus measures are introduced.

In order to evaluate the performance of the estimation methods, some performance measures are needed. A popular measure for the parameter estimate is Root Mean Square Error (RMSE) defined as

\[
\text{RMSE}(\hat{\theta}) = \sqrt{\frac{\sum_{k=1}^{N} (\hat{\theta}_k - \theta^0)^2}{N}} \tag{5}
\]

where \( \hat{\theta}_1, \ldots, \hat{\theta}_N \) are the unbiased estimates of the true scalar parameter \( \theta^0 \). To assess the quality of the obtained SAR images, the power of the error image can be used. This can be defined as

\[
\hat{P} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{M} |\hat{I}_{ij} - I^0_{ij}|^2}{KM} \tag{6}
\]

where \( \hat{I} \) is the \( K \times M \) complex SAR image obtained with the estimation procedure and \( I^0 \) is the perfect focused SAR image, i.e. created with the true trajectory.

### III. NAVIGATION FRAMEWORK

An Inertial Navigation System (INS) in an aircraft integrates accelerometer and gyroscope data and corrects the state with aiding sensors such as barometer and GPS using a general dynamics and measurement equations

\[
x_{t+1} = f(x_t, w_t) \tag{7a}
\]
\[
y_t = h(x_t) + e_t \tag{7b}
\]

where \( x_t \) are states of the system, \( w_t \) denotes the process noise with variance \( Q_t \), \( e_t \) is measurement noise with variance \( V_t \) and \( y_t \) are the measurements. Usually an Extended Kalman filter is applied to estimate the states, see e.g. [38]. In this work, a simplified, yet useful, model of the dynamics will be assumed which will give simpler expressions in the algorithms.

#### A. Aircraft Model

In this setup, the following 2-DOF linear INS time discrete dynamics is used, [38], and it is assumed that all states are measured by GPS and inertial measurements up to the time synthetic aperture is started

\[
x_{t+1} = FX_t + GW_t \tag{8a}
\]
\[
F = \begin{bmatrix} I_2 & T_sI_2 & \frac{T_s^2}{2}I_2 \\ 0_{2 \times 2} & I_2 & T_sI_2 \\ 0_{2 \times 2} & 0_{2 \times 2} & I_2 \end{bmatrix} \tag{8b}
\]
\[
G = \begin{bmatrix} T_s^3I_2 \\ T_s^2I_2 \\ T_sI_2 \end{bmatrix} \tag{8c}
\]

\[
x_t = [X_t, Y_t, v_t^X, v_t^Y, a_t^X, a_t^Y]^T \tag{8d}
\]
\[
w_t = [v_t^X, v_t^Y]^T \tag{8e}
\]
\[
y_t = [X_t, Y_t]^T \tag{8f}
\]

### Table II: Measurement error and stationary measurement error for the navigation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meas. Error (1-( \sigma ))</th>
<th>Stat. Error (1-( \sigma ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>3 m</td>
<td>0.093 m</td>
</tr>
<tr>
<td>Velocity</td>
<td>0.4 m/s</td>
<td>0.012 m/s</td>
</tr>
<tr>
<td>Acceleration</td>
<td>0.06 m/s²</td>
<td>0.015 m/s²</td>
</tr>
</tbody>
</table>

where \( T_s \) is the sampling time in seconds, \( X \) is the position in azimuth direction and \( Y \) is the position in range direction in meters, \( v^X \) and \( v^Y \) are the velocities in the \( X \)- and \( Y \)-directions in meters per second respectively and \( a^X \) and \( a^Y \) are the accelerations in \( X \)- and \( Y \)-directions in meters per second squared respectively. This model is used for the whole trajectory. Since this model is linear and time invariant the stationary Kalman filter can be used to estimate \( x_t \) giving \( \tilde{x}_t \) and its corresponding covariance \( P_t \).

#### B. Navigation Performance

Due to the fact that the system is time invariant and linear, the covariance of the estimate will converge to the stationary covariance \( \hat{P} \). This covariance can be calculated as

\[
\hat{P} = FP\hat{P}T - FP\hat{P}TH\hat{P}F^T + GQG^T
\]

where \( F \) and \( G \) are defined above, and \( H \) is as mentioned before chosen as \( H = I_6 \), since we assume that all states are measured by the navigation system. For a typical navigation system used in an UAV, the accuracy for the position, velocity and acceleration measured by the GPS and INS can be summarised according to Table II. System noise covariance, \( Q \), which represents disturbance on states, like wind turbulence, can be taken as \( \text{diag} \{0.25, 0.25\} \text{ [m}^2/\text{s}^4\]}. With these values, the stationary standard deviation is given in the third column in Table II.

### IV. FOCUS MEASURES

We here review and compare two common focus measures.

#### A. Two Entropy Measures

One common focus measure in SAR or image processing literature is image entropy calculated as

\[
E_1(I) = -\sum_{k=1}^{256} q_k \log_2(q_k) \tag{10}
\]

where \( q_k \) is an approximated grey level distribution of the \( K \times M \) grey-scale image \( |I| \), where \( I \) is the complex-valued SAR image. It can be obtained from the image histogram calculated as

\[
q_k = \frac{\# \text{ of pixel values } |I_{ij}| \in [k-1,k]}{KM} \tag{11a}
\]

\[
k \in [1, 256]. \tag{11b}
\]

The more focused the image is, the lower the entropy is, see for example [21]. Histograms for the images in Figure 1 are given in Figure 4. Note the log-scale on the \( y \)-axis. An alternative
definition of entropy, and more commonly used in the SAR context, is, [21], [34], [39],
\[
E_2(I) = - \sum_{i=1}^{K} \sum_{j=1}^{M} q_{ij} \ln(q_{ij}) \quad (12a)
\]
\[
q_{ij} = \frac{I_{ij}^2}{\sum_{i=1}^{K} \sum_{j=1}^{M} |I_{ij}|^2}. \quad (12b)
\]

B. Focus Measure Performance

Entropy 1 and 2 focus measures are tested and compared on SAR images according to Figure 5 and the results are depicted in Figure 6 where standard deviations 1 − σ, 2 − σ and 3 − σ are also drawn. These images are chosen since they represent both structured and unstructured scenes and they are also small enough to satisfy the small scene assumption needed for the gradient calculation in (4). On top of that, a small image is faster to calculate, giving the faster optimisation calculation in turn. As an additional motivation for using these small artificial scenes, another, larger and more realistic scene is examined, see Figure 7a. The scene is created by using a real SAR image over part of Washington DC and creating raw data from it. After that the same simulation is performed as for the small scenes and Entropy 1 and 2 measures are given in Figure 7b and 7c. Since the principal form of the entropy measures is essentially the same, more thorough examination will be done on the two small scenes due to above mentioned reasons. In all these simulations the state noise in model (8) is set to zero, i.e. the trajectory is completely deterministic. This is done in order to illustrate the focus measure functions $F_i$ in a two dimensional plot, since the trajectory, and consequently the focus measure, is then only dependent on the initial values.

In all these figures it can be seen that entropy 2 has a convex and pretty nice behaviour around the true value of the initial state. However it looks very flat along the velocity direction which indicates that it is very difficult to estimate that particular state. The entropy 1 measure has, on the other hand, a sharp minimum for the correct value of the initial state but many local minima. This means that the two entropy measures complement each other perfectly, and can be used in combination to obtain the global minimum of the focus measure.

V. SEARCH METHODS

As demonstrated in Section IV-B, the entropy 2 measure can be used as a coarse first step in the optimization to come close to the global minimum, and then entropy 1 can be used to obtain the global minimum. Note that a special structure of the problem (8) allows for unconstrained solution of the problem. This is due to the fact that the constraints representing the trajectory can be taken into account while calculating the gradient of the cost function as will be demonstrated in Section V-C.
A. Joint Optimization of Trajectory and Focus

Since the focus of the image depends on the unknown trajectory, one solution is to solve the following minimisation problem

\[
\hat{\theta} = \arg \min_{\theta} g(\theta)
\]

\[
g(\theta) = \gamma_F E_i(x_{0:N}) + \gamma_s \sum_{i=1}^{N} \| y_t - h(x_t) \|_{V_t^{-1}} + \| w_t \|_{Q_t^{-1}}
\]

subject to \( \theta = [x_0^T, w_{1:N}^T]^T \)

where \( \gamma_F \) and \( \gamma_s \) are the weights (and \( \gamma_F + \gamma_s = 1 \)) and measurement equation \( h(x) \) and system dynamics \( f(x, w) \) are defined as in Section III. In Equation (13b), \( E_i(x_{0:N}), i \in \{1, 2\} \), is a function of the SAR image \( I \) created from the radar measurements and is of the type “how focused is the image?” according to Section IV.

B. Gradient Search

Gradient search methods will be exemplified here with a couple of examples with different trajectories and errors in them. Only two states and their initial values are considered (\( v_0^\chi \) and \( \alpha_0^\chi \)), for illustrative purposes. In general, the minimisation should be applied for these states for all or at least some of the time instants along the trajectory. Such an example will be studied later.

A gradient search can, for the general problem (13a), be formulated as

\[
\theta^{k+1} = \theta^k + \mu_k H(\theta^k)^{-1} \nabla g(\theta^k)
\]  

(14a)

\[
\nabla g(\theta) = \frac{\partial}{\partial \theta} g(\theta)
\]  

(14b)

where \( \mu_k \) is step size with \( \mu_0 = 1 \) and \( H(\theta) \) is some (positive definite) matrix. The step size is used to ensure that each iteration actually gives a decreasing value of the function. The initial estimate, \( \theta^0 \), can be taken as the usual estimate from the navigation system. In the simplest case \( H \) can be chosen as the identity matrix and the procedure becomes a pure gradient search. The disadvantage of such procedure is the slow convergence, especially if the function to be minimised is ridge-like like entropy 1 focus measure. If \( H \) is chosen as the Hessian of \( g \), the procedure becomes a Newton search. The Newton search has a fast convergence, and is to prefer if the Hessian is available. In many cases the Hessian is either not available or very difficult to obtain, as in the case considered here, and some approximate methods must be applied. One option is a quasi-Newton method, and BFGS in particular, where the Hessian is approximated by utilising gradients of the function during the search, see [40]. The general gradient search procedure is summarised in Algorithm 1.

In all these procedures it is essential to obtain the gradient of the loss function. Because of the special structure of the focus measure function and the SAR processing algorithm, the complete analytical gradient is hard to obtain. For example, for the entropy 1 measure it is hard to differentiate a histogram of the image. In this case numerical methods must be used. However, for the entropy 2 measure it is possible to obtain analytical expressions for most part of the gradient. Numerical gradient calculation requires that one new SAR image is created for each parameter and step in the inner loop in Algorithm 1 in each iteration. The analytical gradient requires only one new image in the inner loop execution per iteration. The SAR image creation process is in this case main bottleneck when it comes to computation time, since gradient calculation is very fast compared to it. How gradient for entropy 2 is calculated will be described in the next subsection.

C. Calculating the Gradient

The calculations to obtain an analytical gradient of the entropy 2 function will be presented. The key to doing this is the chain rule for gradient calculation (see also Section II and Equation (4)),

\[
\frac{\partial E_2}{\partial \theta} = \frac{\partial E_2}{\partial q} \frac{\partial q}{\partial I} \frac{\partial I}{\partial R} \frac{\partial R}{\partial \theta}
\]

(15)

In order to apply the chain rule, first the decomposition chain of the entropy 2 focus measure will be considered and then, all partial derivatives will be presented.
From the figure it can be seen that the range function of the states, the geometry setup as in Figure 8 can preliminaries are needed. In order to express range as a function of the states, some SAR geometry
of the range as a function of the states, the geometry setup as in Figure 8 can
preliminaries are needed. In order to express range as a function of the states, some SAR geometry
preliminaries are needed. In order to express range as a function of the states, some SAR geometry

\[
\frac{\partial g}{\partial R} = \frac{1}{2} \sqrt{\sum_{i=1}^{K} \sum_{j=1}^{M} q_{ij} \ln q_{ij}} - \sum_{i=1}^{K} q_{i} \ln q_{i}
\]

(16)

where last equality is simply reformulation of the double sum by vectorising the image.

In the second factor, each \( q_{i} \) is obtained by

\[
q_{i} = \frac{|I_{i}|^2}{\sum_{j=1}^{M} |I_{j}|^2}
\]

(17)

meaning that \( q_{i} \) is a function of the absolute value of the complex-valued SAR image.

For the third factor, we need to obtain the derivative \( \partial |I|/\partial R \). In the creation of the image, a back-projection sum is evaluated and all partial images are summed up. Each partial image is a function of one column in the RAR image and the range from the platform to each pixel in the SAR image, see Section II. Unfortunately it is not easy, if not impossible, to obtain analytical expression for the derivative \( \partial |I|/\partial R \). However, this value can simply be obtained during image creation by means of numerical derivation. The cost for that procedure is memory demand and execution time which both are doubled. But this increase in cost is independent of the state, and thus constant no matter how many parameters optimisation is performed over. A straightforward numerical gradients would give a cost that is increasing linearly with the number of parameters.

The last factor that needs to be calculated is the gradient of the range as a function of the states, \( R(x_{t}) \). To calculate an analytical expression of this function some SAR geometry preliminaries are needed. In order to express range as a function of the states, the geometry setup as in Figure 8 can be considered. From the figure it can be seen that the range

\[
R_{t} = \sqrt{(X_{m} - X_{t})^2 + (R_{y} - Y_{t})^2 + (Z_{0} - Z_{t})^2}
\]

(18a)

\[
R_{y} = \sqrt{R_{m}^2 - Z_{0}^2}
\]

(18b)

i.e. as a function of the trajectory. Note that \( Z_{t} \), the altitude of the platform, is assumed to be known here. This can be achieved by measuring it with, for example, barometric sensors which is always done in the aircraft applications. Here the exact expression for the range along the trajectory is used, unlike in most of the SAR literature, where approximate and linearised expressions are used, see for example [23]. This is due to the fact that in low frequency SAR application, as the one considered here, the ratio between range and trajectory length is not negligible due to the lobe width. If approximate expressions are used, too large errors would be introduced in the beginning and the end of the trajectory.

Next, the dynamical model (8) can be used to express the position states used in the range expression above as a function of any other state by using that

\[
x_{t} = F^{t-k}x_{k}, \quad t > k
\]

(19)

Note that the state noise term, \( w_{t} \), is neglected in the following since it is equivalent to optimise over noise and over acceleration states, \( a_{t} \), so the latter one is used here to simplify the expressions and reduce the amount of variables in the problem. From (19) any position can be expressed explicitly as

\[
X_{t} = X_{0} + T_{s}(t-k)\nu_{k}^{x} + \frac{T_{s}^2(t-k)^2}{2}a_{k}^{x}
\]

(20a)

\[
Y_{t} = Y_{0} + T_{s}(t-k)\nu_{k}^{y} + \frac{T_{s}^2(t-k)^2}{2}a_{k}^{y}
\]

(20b)

If these expressions are used in (18a), we can easily obtain partial derivatives of the range with respect to the velocities and accelerations in arbitrary time points.

Now, we have everything needed to calculate the gradient of the focus measure with respect to the trajectory states. The
Table III: SAR parameters used for the simulated data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radar center frequency</td>
<td>53.125 MHz</td>
</tr>
<tr>
<td>Aperture length ($2X_m$)</td>
<td>2770 m</td>
</tr>
<tr>
<td>Ground range to target ($R_g$)</td>
<td>2182 m</td>
</tr>
<tr>
<td>Nominal altitude ($2h_0$)</td>
<td>1000 m</td>
</tr>
<tr>
<td>Nominal speed ($v_0^x$)</td>
<td>100 m/s</td>
</tr>
</tbody>
</table>

Partial derivatives are, in turn (for $t > k$),
\[
\frac{\partial E_2}{\partial q_i} = -\ln q_i - 1 \quad (21a)
\]
\[
\frac{\partial q_i}{\partial |I_j|} = \begin{cases} \frac{2|I_i| \sum |I|^2 - 2|I_i|^2 |I_j|}{\left(\sum |I|^2\right)^2}, & i = j \\ \frac{-2|I_i|^2 |I_j|}{\left(\sum |I|^2\right)^2}, & i \neq j \end{cases} \quad (21b)
\]
\[
\frac{\partial R_t}{\partial v_k^X} = \frac{(X_m - X_t)T_s(t - k)}{R_t} \quad (21c)
\]
\[
\frac{\partial v_k^X}{\partial R_t} = \frac{(R_g - Y_t)T_s(t - k)}{R_t} \quad (21d)
\]
\[
\frac{\partial v_k^Y}{\partial v^X} = \frac{-(X_m - X_t)T^2_s(t - k)^2}{2R_t} \quad (21e)
\]
\[
\frac{\partial a_k^X}{\partial a_k^Y} = \frac{-(R_g - Y_t)T^2_s(t - k)^2}{2R_t} \quad (21f)
\]

and $\partial |I|/\partial R$ is numerically calculated during image formation. Now, at least for entropy 2 focus measure, we can calculate the gradient (semi-) analytically and use it in the minimisation procedure. The second term in (13b) is easy to differentiate, since it is a quadratic form and $h(x)$ is a linear function in this case.

VI. NUMERICAL EXAMPLES FOR SIMULATED IMAGES

In order to demonstrate the behaviour of the gradient search for this setup, the SAR image from Figure 5a is used in two different experiments. In these simulations SAR parameters according to Table III are used.

A. Two-Dimensional Optimization

To be able to illustrate the convergence of the solution, only two optimisation variables are considered here, $\theta = [v_0^X, a_0^Y]^T$ and the algorithm is initiated with random starting points $\theta^0$ based on the stationary covariance of the states in the system. Those initial values are $[100.005, 0.005]^T$, $[99.99, 0.01]^T$, $[99.995, 0.02]^T$, $[100.02, -0.01]^T$ and $[100.005, -0.035]^T$.

The trajectories generated with these initial values are illustrated in Figure 9. In Figure 10a, the gradient search based entropy 2 measure is illustrated and we can see that the solutions converge to the flat ridge-like area close to the correct acceleration, but not necessarily to the correct velocity. In Figure 10b, the gradient search where the entropy 1 measure is used is depicted. In this case the algorithm is initiated with the solution from the entropy 2 search. It can be seen that this minimisation strategy works pretty well, although one solution is stuck in a local minimum. In that case the velocity error is the largest one of all errors. Note also that only the focus measure is used to find estimate of the states i.e. $\gamma_s$ is set to zero while $\gamma_f$ is set to one in Equation (13b).

Figure 9: Trajectory examples for different initial values of the velocity, $v_0^X$ and acceleration, $a_0^Y$.

Figure 10: Search trajectory for five different values of $x_0$ using two different entropy measures.
Figure 11: Resulting images from the minimisation procedure with starting point [100.02, −0.01]T.

It is interesting to see how the image created with the solution that is stuck in the local minimum of the entropy 1 measure looks like compared to the unfocused image that is initialised with. As illustrated in Figure 11, it can be seen that the image created with values from the minimisation procedure is very close to the focused image and much better than the unfocused images that are initialised with. The probable explanation for this comes from the fact that small azimuth direction velocity errors do not influence the final image much due to the quantisation effects. However the estimate of the navigation states is not correct.

B. High-Dimensional Optimization

In the second example a more realistic setup is done. The optimisation problem to be solved is

$$\hat{t} = \arg \min_{\theta} \gamma_F E_{1,2}(X_{0:N}) + \gamma_s \sum_{t=1}^{N} \|a_t^Y - a_t^Y\|_{V_t}^2 \quad (22a)$$

subject to

$$\theta = [v_0^X, a_0^Y, a_{[N/4]}^Y, a_{[N/2]}^Y, a_{[3N/4]}^Y]^T \quad (22b)$$
$$\gamma_F = 0.99, \ \gamma_s = 0.01 \quad (22c)$$
$$x_{t+1} = F x_t, [X_0, Y_0, v_0^Y, a_0^X]^T = 0_{4 \times 1}, \ P_0 = \infty \cdot I_2 \quad (22d)$$
$$a_t^X = 0, \ t \in \{0 : N\} \quad (22e)$$
$$a_t^Y = \begin{cases} a_t^Y, & t \in \{0 : [N/4] - 1\} \\ a_{[N/4]}^Y, & t \in \{[N/4] : [N/2] - 1\} \\ a_{[N/2]}^Y, & t \in \{[N/2] : [3N/4] - 1\} \\ a_{[3N/4]}^Y, & t \in \{[3N/4] : N\} \end{cases} \quad (22f)$$

where $a_t^Y$ is the measured acceleration in $Y$-direction with additive white Gaussian noise with $V_t = 0.0022 \text{ m}^2/\text{s}^4$. $E_{1.2}(X_{0:N})$ is either entropy 2 or entropy 1, exactly as in the previous example. Here it is assumed that a change in $Y$-direction acceleration will behave in a step like manner only a few times during the SAR image generation and that the amplitude of the step is arbitrary. It is also assumed that the acceleration in $X$-direction will vary slowly due to the platforms inherited inertia in this direction, so it can be assumed to be zero. The meaning of $P_0$ in (22d) is that there is no prior information about the initial values of the trajectory. Another spline-like interpretation of the setup in (22) is to find a best trajectory by fitting the second order polynomials between four points evenly spaced along the trajectory. Both scenes from Figure 5 are used and 30 Monte Carlo simulations are performed in order to evaluate the performance of the estimation procedure.

The resulting RMSE of the parameters and the mean value of the error image power are presented in Table IV and Table V for both structured and unstructured scene. Here, the actual acceleration is presented instead of the process noise value, since it is more physically interesting. It can be noticed that

<table>
<thead>
<tr>
<th>Parameter</th>
<th>RMSE (opt. with $E_2$)</th>
<th>RMSE (opt. with $E_1$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_0^X$</td>
<td>$11.2 \cdot 10^{-3}$ m/s</td>
<td>$11.2 \cdot 10^{-3}$ m/s</td>
</tr>
<tr>
<td>$v_0^Y$</td>
<td>$11.61 \cdot 10^{-4}$ m/s</td>
<td>$10.98 \cdot 10^{-4}$ m/s</td>
</tr>
<tr>
<td>$a_{[N/4]}^Y$</td>
<td>$6.63 \cdot 10^{-5}$ m/s</td>
<td>$6.52 \cdot 10^{-5}$ m/s</td>
</tr>
<tr>
<td>$a_{[N/2]}^Y$</td>
<td>$9.31 \cdot 10^{-4}$ m/s</td>
<td>$8.86 \cdot 10^{-4}$ m/s</td>
</tr>
<tr>
<td>$a_{[3N/4]}^Y$</td>
<td>$7.77 \cdot 10^{-4}$ m/s</td>
<td>$7.58 \cdot 10^{-4}$ m/s</td>
</tr>
<tr>
<td>Mean value of the error image power</td>
<td>1348</td>
<td>1242</td>
</tr>
</tbody>
</table>

Table V: RMSE for the estimated parameters and the mean value for the error image power for the unstructured scene.
erations were about 10 for the entropy 2 case and 7 for the entropy 1 case. Additional step size calculations were about 2 per iteration giving 20 and 14 gradient calculations in total. As previously mentioned, the main computational speed bottleneck is the SAR image creation in each iteration and for the used images it is about 1 minute. That gives an average time of about 20 minutes for entropy 2 measure and 70 minutes for the entropy 1 measure (since there are 5 parameters in total). This time can be decreased if faster image creation procedure is used.

Figure 13: Resulting images from the gradient search minimisation.

(a) Image of the structured scene after minimisation with entropy 2 as focus measure.
(b) Image of the structured scene after minimisation with entropy 1 as focus measure.
(c) Image of the unstructured scene after minimisation with entropy 2 as focus measure.
(d) Image of the unstructured scene after minimisation with entropy 1 as focus measure.

Table VI: Error in the estimated parameters for the two scenes after entropy 1 minimisation procedure. These are the parameter values used to create images in Figure 13b and Figure 13d

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Structured scene</th>
<th>Unstructured scene</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_0^X$</td>
<td>$9.236 \times 10^{-3}$ m/s</td>
<td>$10.13 \times 10^{-3}$ m/s</td>
</tr>
<tr>
<td>$v_0^Y$</td>
<td>$-3.057 \times 10^{-4}$ m/s$^2$</td>
<td>$-1.707 \times 10^{-4}$ m/s$^2$</td>
</tr>
<tr>
<td>$a_0^Y$</td>
<td>$-0.553 \times 10^{-4}$ m/s$^2$</td>
<td>$-1.733 \times 10^{-4}$ m/s$^2$</td>
</tr>
<tr>
<td>$a_1^X$</td>
<td>$11.15 \times 10^{-4}$ m/s$^2$</td>
<td>$10.25 \times 10^{-4}$ m/s$^2$</td>
</tr>
<tr>
<td>$a_1^Y$</td>
<td>$1.384 \times 10^{-4}$ m/s$^2$</td>
<td>$1.234 \times 10^{-4}$ m/s$^2$</td>
</tr>
<tr>
<td>Error power</td>
<td>53.37</td>
<td>53.12</td>
</tr>
</tbody>
</table>

Table VII: SAR parameters used for the real data.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radar center frequency</td>
<td>52.1875 MHz</td>
</tr>
<tr>
<td>Aperture length ($2X_m$)</td>
<td>12187 m</td>
</tr>
<tr>
<td>Ground range to target ($R_g$)</td>
<td>8850 m</td>
</tr>
<tr>
<td>Nominal altitude ($Z_0$)</td>
<td>3015 m</td>
</tr>
<tr>
<td>Nominal speed ($v_0^c$)</td>
<td>128 m/s</td>
</tr>
</tbody>
</table>

VII. EXAMPLE WITH REAL SAR IMAGE

Here, we illustrate the estimation results using data from the CARABAS II system [41] collected in western Sweden. The trajectory and the SAR image obtained by the proposed estimation method are compared to the image created with the GPS based trajectory, which is assumed to be the ground truth. The SAR image used for the estimation is illustrated in Figure 14a where the GPS based trajectory is used to generate the image. SAR parameters used for the real data are given in Table VII.

For the real data case, the optimisation problem to be solved is formulated according to

$$\hat{\theta} = \arg \min_{\theta} \gamma_F E_2(x_{0:N}) + \gamma_s \sum_{i=1}^{N} ||a_i^m - a_i||^2_{V_t^{-1}}$$  \hspace{1cm} (23a)

subject to

$$\theta = [v_0^X, v_0^Y, a_0^X, a_0^Y, a_1^X, a_1^Y]^T$$  \hspace{1cm} (23b)

$$i \in \{0 : [N/100]\}$$  \hspace{1cm} (23c)

$$\gamma_F = 0.36, \gamma_s = 0.64$$  \hspace{1cm} (23d)

$$x_{t+1} = Fx_t, [X_0, Y_0]^T = 0_{2 \times 1}, P_0 = \infty \cdot I_4$$  \hspace{1cm} (23e)

$$a_i = \begin{cases} a_0, & i \in \{0 : [N/200]\} - 1 \\ a_{[N/200]}, & i \in \{[N/200] : [N/100]\} - 1 \\ \vdots \end{cases}$$  \hspace{1cm} (23f)

where the variables are defined as

$$a_t = [a_t^X, a_t^Y]^T$$  \hspace{1cm} (24a)

$$a_t^m = [a_t^mX, a_t^mY]^T$$  \hspace{1cm} (24b)

$$V_t = \text{diag}\{V_t^X, V_t^Y\} = \text{diag}\{0.1, 0.1\} [m^2/s^2]$$  \hspace{1cm} (24c)

$a_t^mX$ and $a_t^mY$ are the accelerations of the platform in $X$- and $Y$-directions measured by the navigation system. Note that in this case only the entropy 2 focus measure is used due to the computational load to calculate the numerical gradient for the entropy 1 measure. However, according to the results in Section VI, the improvement of the estimates by using
additional optimisation with entropy 1 is small and therefore it is omitted here. Also, the weights and the covariance of the acceleration measurements are seen as tuning parameters.

Results from the optimisation procedure, which takes five steps to converge, is illustrated in Figure 14b, where the estimated trajectory is used to generate the image and Figure 15a where error between GPS and estimated trajectory is shown. That error can be compared to the error in the trajectory with the initial values of the parameters, $\theta_0$, shown in Figure 15b. It can be seen from these two plots that improvement in $Y$-direction is much less than improvement in $X$-direction. The total loss function (23a) and entropy 2 measure as a function of the iteration number defined through Algorithm 1 are depicted in Figures 16a and 16b, respectively. The image resulting from the estimated trajectory is hard do distinguish from the GPS based trajectory image, except that it is shifted in range direction. This ambiguity is, unfortunately, unobservable in the auto-focusing process, i.e. the method is invariant to the translation of the image.

Computational time in the real data case is five iterations and about 10 gradient calculations in total (including inner loop in the Algorithm 1). Time for processing the SAR image was 5 minutes which gives total time of about 50 minutes.

VIII. CONCLUSIONS AND FUTURE WORK

An iterative optimisation method is presented based on a decentralised sensor fusion framework, which is intended to provide better focused SAR images on future cheap and small SAR platforms. The approach is based on jointly optimizing a focus measure and the error in the navigation states. As was concluded from simulation examples of the simple scene and real SAR data, the method works fairly well, although not all the states are observable and some errors in these are still present. Nevertheless, even if the whole navigation state vector can not be corrected, the resulting SAR image is much more focused after optimization than the original one. An important theoretical contribution, to reach the requirements on computation complexity, is an analytical expression for the gradient of a focus measure. The result enables an implementation of gradient search algorithms that adds only marginal complexity to the SAR image creating process.

In the derivation of the focus measure gradient needed for the optimisation process, a small scene approximation is used. Since SAR images can be quite large, an extension of using several small images from one large, each with different range, can be applied. This might give more observability of the parameters in the loss function based on entropy.

ACKNOWLEDGMENTS

The authors would like to thank Lars Ulander and Anders Gustavsson from Swedish Defence Research Agency (FOI) in Linköping, Sweden, for providing real CARABAS II data and help with these. Our thanks go also to Hans Hellsten from Saab Electronic Defence Systems in Göteborg, Sweden, for all help with the SAR system theory.

This work was supported by the Industry Excellence Center LINK-SIC founded by The Swedish Governmental Agency for Innovation Systems, VINNOVA and Saab AB.

REFERENCES


