

Navigation and SAR focusing with Map Aiding

Zoran Sjanic and Fredrik Gustafsson *Fellow, IEEE*

Abstract—A method for fusing Synthetic Aperture Radar (SAR) images with optical aerial images is presented. This is done in a navigation framework, where the absolute position and orientation of the flying platform, as computed from the inertial navigation system, is corrected based on the aerial image coordinates taken as ground truth. The method is suitable for new low-price SAR systems for small unmanned vehicles. The primary application is surveillance and to some extent remote sensing, where the SAR image provides complementary information by revealing reflectivity to microwave frequencies.

The method is based on first applying an edge detection algorithm to the images and then optimising the most important navigation states by matching the two binary images. To get a measure of the estimation uncertainty, we embed the optimisation in a least squares framework, where an explicit method to estimate the (relative) size of the errors is presented. The performance is demonstrated on real SAR and aerial images, leading to an error of only a few pixels (around 4 meters in our case), which is a quite satisfactory performance for applications like surveillance and navigation.

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

I. INTRODUCTION

A radar mounted on a flying platform, like an aircraft or a satellite, can be used to get an image of the surroundings by taking intensity (or radar cross section) of the reflections and map it to pixels. This kind of image would be of pretty bad quality, since the resolution will be decided by the radar lobe width which in turn is decided by the antenna length and the frequency of the radar. For realistic antenna lengths found on the flying platforms, this resolution is in range of several tenths of meters or more. By taking many radar echoes from the same area by moving the radar antenna and in this way creating a large synthetic antenna, images with much higher resolution can be created. This is the basics of the Synthetic Aperture Radar (SAR) imaging, [1]. For more detailed description of SAR and SAR images see e.g., [2]. With modern SAR systems the resolutions in images can be as good as a couple of decimetres, giving very detailed images of the scene. The knowledge of the flown trajectory is very important in the image creation principle and errors in the trajectory will lead to defocused SAR images. A process to correct for these image defects is called autofocusing. There exist many autofocusing methods, of which some are based on the raw radar data and others on the already processed SAR image, [2]–[7]. All these methods use only SAR images without any prior information of the scene to perform the focusing.

The goal of this work is to match SAR images with optical images or map information, e.g., Google Maps. The

fusion of information from these sources is then utilised for autofocusing and correction of the navigation trajectory. The assumptions are that the most focused images will also match the optical map in the best way and in turn correspond to the best possible trajectory (giving the best focus) and best possible absolute position on the map (giving the global navigation ability). Traditionally, SAR images are usually used for surveillance and remote sensing purposes, but some cases where they are used for navigation purposes have also been studied, see e.g., [8]. The method can be useful as an alternative to high precision navigation aids, such as Global Navigation Satellite System (GNSS), of which GPS NavStar is the most famous one, to stabilise inertial based navigation systems which are known to be prone for long term drift. The method has many similarities to the e.g., terrain aided navigation [9], where an altitude database of the terrain is used to support navigation. Other similar methods are visual odometry, [10], and a method of aided navigation where optical cameras and maps are used to navigate by matching the camera images and the map, see e.g., [11], [12]. However, the fusion of the SAR and optical map images is not as trivial task as to match optical camera images to the map images, since the SAR images have quite different properties than the optical images. The SAR images show the reflectivity of the scene for microwave frequencies instead of visible light frequencies. This implies that completely different information can be contained in the SAR images compared to the optical images, although some of the features in the images are clearly very similar. This makes the fusion of SAR and optical images a promising method for surveillance and, in some cases, remote sensing applications. As a navigation tool, SAR is not sensitive to occlusions from clouds like optical sensors are, giving a less weather sensitive position sensor.

As means for extracting useful information from the images, an edge detector (Canny edge detector) and a modified image matching method (Chamfer matching) will be used in order to match SAR images to the optical map images. The results of the matching and focusing method will be illustrated on real SAR and optical images that are depicted in Figures 1 and 2. The SAR image is created by a radar operating in Ku-band (12 – 18 GHz) with a nominal resolution of 1 m in both directions. This work is an extension of [13], where only image matching was considered with the assumption that autofocusing has already been performed.

The paper is organised as follows, Section I introduces the work, Section II explains the SAR imaging principle. Section III introduces the navigation models and defines the basic SAR geometry that relates the image and the flying platform while Section IV introduces the image matching approaches. In Section V, estimation of the kinematic trajectory parameters is explained and in Section VI the results from the



Figure 1: SAR image of Washington D.C. Image: Sandia National Laboratories.



Figure 2: Optical image of Capitol Hill, Washington D.C. with surroundings. Image: Google Maps.

matching and kinematic estimation are shown. In Section VII, conclusions are given and some future work is discussed.

II. SAR IMAGING PRINCIPLE

SAR imaging is based on a moving platform that passes the scene that shall be imaged. During the movement, the platform transmits radar pulses which hit the scene and return to the platform with a certain time delay which is proportional to the range to the scene. This returned signal is filtered with a matched filter and then sampled. Each reflector in the scene will contribute with its reflected power which will then be placed in the appropriate range bin. The range is determined as a product between signal propagation speed (usually speed of light) and delay time. In this way a single scene transfer function is obtained, denoted $g(R)$. Now this process can be repeated during platform movement, and all the stored transfer functions are stored in a two-dimensional array $g_t(R)$. Basically, this raw data, $g_t(R)$, is an example of a real aperture radar (RAR). The resolution in such radar system is proportional to the antenna lobe width and is usually quite poor. One important thing to notice is that the lobe width is inversely proportional to the antenna size, i.e., the larger antenna the smaller lobe we can obtain. The idea behind SAR is to artificially synthesize a big antenna by moving the platform. Traditionally, this operation is performed in the

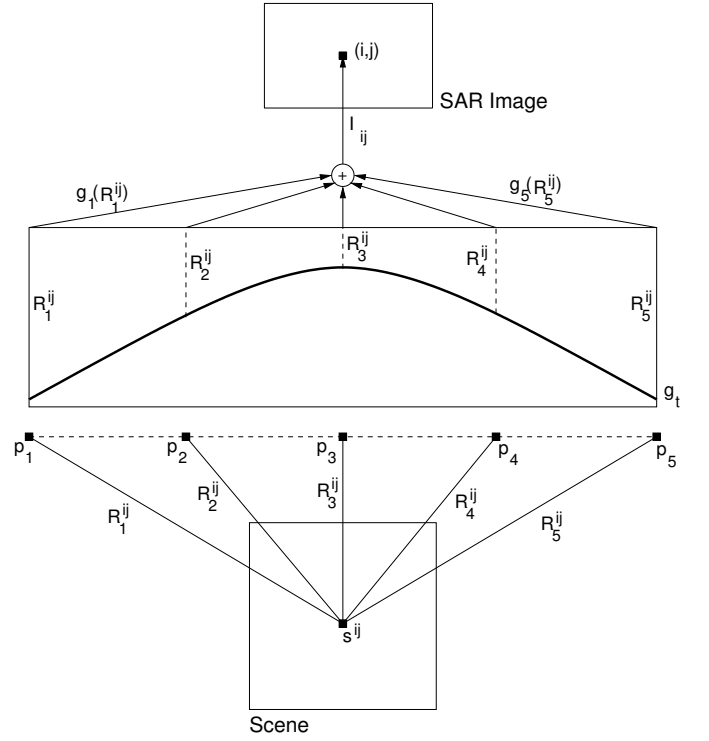


Figure 3: The global backprojection method for creating SAR images. The scene consists of only one point target in this illustration. The figure is not to scale.

frequency domain using FFT like methods, e.g., the Fourier-Hankel method [14]–[16] or the ω -K migration methods [17]–[19]. The common denominator of these methods is that they assume that the aircraft's (or antenna's) flown path is linear 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 non-linear trajectory but then the methods become computationally inefficient. Another method that can be used is so called global back-projection method that will be outlined below.

Given the raw (complex) data $g_t(R)$ we can back-project each radar echo on the image giving the subimage I_t and each reflector will create a circle in each subimage. The total image can then be created by summing up all the subimages along the synthetic aperture, [20], (i.e., solving the back-projection integral in discrete time)

$$I = \sum_{t=1}^N I_t \quad (1)$$

Another way of creating the image is to integrate the raw data for each pixel in the image I_{ij} as

$$I_{ij} = \sum_{t=1}^N g_t(R_t^{ij}) \quad (2a)$$

$$R_t^{ij} = \|p_t - s^{ij}\|_2 \quad (2b)$$

where p_t is the position of the platform and s^{ij} is the position in the scene which corresponds to the pixel (i, j) . This method is schematically illustrated in Figure 3 for a simple scene with

only one point target and where only a few platform positions are considered, here $N = 5$.

III. MOTION MODELS

Precise knowledge of the antenna position p_t in (2b) is apparently crucial. The on-board inertial navigation system (INS) provides a nominal trajectory $\bar{p}_{1:N}$ that can be used to construct a first (un-focused) SAR image. Our approach to focus the SAR image is based on computing a refined trajectory $p_{1:N}$. In the sequel, we will implicitly only model the deviation from the nominal trajectory, so p_t denotes the difference of the true position and the INS position. In this work we assume that INS is of approximately tactical or less performance usually used in unmanned aerial vehicles. The dominating source of error for these sensors is acceleration bias which has a typical value of $5 \cdot 10^{-3} - 5 \cdot 10^{-2} [\text{m/s}^2]$. With this size of acceleration bias the accumulated velocity and position errors after 120 s are between 0.6 and 6 [m/s] and 36 and 360 [m] respectively. The time of 120 seconds is taken as a typical time for SAR image acquisition, and hence it is assumed that the corrections to the navigation system are performed with this time interval.

This error trajectory of the platform is assumed to follow a simple second order dynamics expressed in discrete time as

$$p_{t+1} = p_t + T_s v_t + \frac{T_s^2}{2} a_t \quad (3a)$$

$$v_{t+1} = v_t + T_s a_t \quad (3b)$$

$$a_{t+1} = a_t \quad (3c)$$

where $p_t = [X_t \ Y_t \ Z_t]^T$ is the platform's position relative to the beginning of the synthetic aperture and $v_t = [v_t^X \ v_t^Y \ v_t^Z]^T$ is its velocity. This model of the kinematics allows us to calculate the whole trajectory if all the initial states, p_0 and v_0 , and acceleration sequence, $a_{0:N}$, are known. The initial position is basically arbitrary since the SAR image can be translated freely, but the initial velocity is not, since it is influencing the trajectory's shape. It is therefore no loss of generality to take $p_0 = 0_{3 \times 1}$, to define the navigation frame. In that case, the trajectory is related to this zero-frame. However, this frame can be translated and rotated and the trajectory will follow as a rigid body. Usually, the accelerations are measured by the onboard inertial measurement unit (IMU), but these are not perfect, and that will also cause an error in the trajectory. These errors will in turn cause the SAR image to become out of focus as mentioned before. So the best focus, i.e., the sharpest image should be produced if correct kinematic states are used. The main approach in this work is basically to use optical images or maps to match the SAR image to these in order to determine the initial states, i.e., both trajectory parameters and global position. This can be done by minimising some criterion that depends on parameters for SAR image's global position and orientation and kinematic states from the model in Equation (3). These parameters are collected into a vector θ , which can for example be $\theta = [r \ a \ \chi \ v_0^X \ v_0^Y \ a_0^X \ a_0^Y]^T$, where r , a and χ are the pixel positions and orientation of the SAR image relative to the optical map image and the other ones are the interesting kinematic states. Since these will be

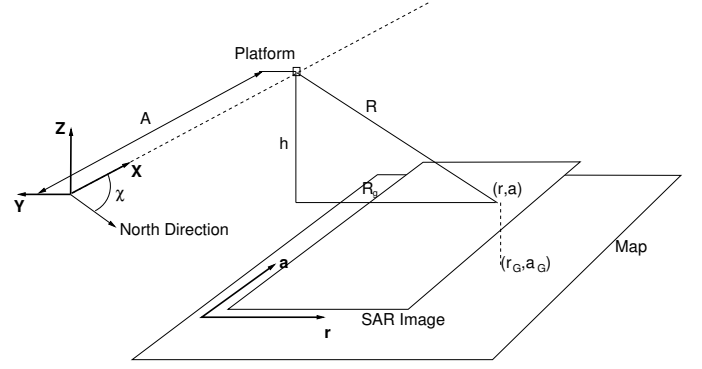


Figure 4: Basic SAR geometry. Also a map is depicted in order to illustrate the used notation in the matching procedure. Figure is not to scale.

treated in a slightly different manner, we divide θ into an image part, θ_I , and a kinematic part, θ_K , as $\theta = [\theta_I^T \ \theta_K^T]^T$. One reason is the fact that the image matching procedure to estimate θ_I is the difficult part of the procedure. The variation of the kinematic part of the parameters, θ_K , will just create different SAR images and each such image must be matched to the optical image. In Section IV the image matching approach (the one that estimates θ_I) will be described in detail. The main assumption is that the best focused (sharpest) image will give the best matching criterion and in that way global position of the platform can be resolved and the most focused image created.

A. SAR Geometry

Aerial images and maps are orthorectified, and for matching purposes the SAR image also needs to be orthorectified. This section describes the geometrical transformations to get a SAR image in the horizontal plane in order to be related with the map.

The basic SAR geometry is illustrated in Figure 4. SAR images use a coordinate system consisting of azimuth direction which is parallel to antenna's moving direction (in most cases the same as the platform's moving direction), here called X , and range direction which is perpendicular to antenna's moving direction, here called Y . Range direction can either be slant range or ground range coordinates. SAR images are naturally slant range images and the advantage of using those is that there is no need for transforming these to ground range images. But if we want to match these to the map images, the map must be transformed to the slant range image. Under the flat terrain assumption this is straight forward. However, if the terrain variation is present, we need some terrain height information in order to accurately reproduce geometrical distortions present in SAR images, such as layover or shadowing, if the terrain variations in the imaged area are considerable.

In order to relate the navigation frame and the SAR image pixels some basic geometry must be defined. Referring to Figure 4 the following notation is defined:

- r is the pixel's coordinate in the SAR image range direction expressed in the image frame.

- a is the pixel's coordinate in the SAR image azimuth direction expressed in the image frame.
- A is the position of the pixel in the azimuth direction expressed in the navigation frame (X).
- R_g is the position of the pixel in the range direction expressed in the navigation frame (Y).
- h is the pixel's altitude below the navigation frame (Z).
- R is the slant range from the X -axis of the navigation frame to the pixel point on the ground (parallel with the Y -axis).

Given now any SAR image pixel's coordinate, (r, a) , its position in the navigation frame (with coordinates $[A \ R_g \ h]^T$ which is the same as s^{ra} in Equation (2b)) can be calculated as follows; First, the pixel's X -direction is simply obtained as

$$A = a\Delta_A \quad (4)$$

where Δ_A is the SAR image resolution in the azimuth direction. Next, the Y -direction coordinate of the pixel is obtained by using the right triangle with sides R , h and R_g and the Pythagorean theorem

$$R_g = \sqrt{R^2 - h^2} = \sqrt{(r\Delta_R + R_0)^2 - h^2} \quad (5)$$

where Δ_R is the SAR image resolution in the range direction and R_0 is minimum slant range in the image (first row in the SAR image is on the range R_0). Last, Z -direction is taken directly as the platform's altitude, h .

Assuming we have a successful matching of the SAR image to the optical image, then a correspondence between the SAR image pixels and the optical image pixels is obtained giving the true geographical positions of the SAR image pixels, (r_G, a_G) , since the optical image pixels have known geographical positions, see Figure 4. The platform's average direction of the flight, χ , i.e., the angle between the North-axis and the X -axis of the navigation frame, is part of the matching results and is obtained directly, since this is the rotation of the SAR image relative to the map image. With these given, the navigation frame's position in global coordinates can readily be obtained by means of rotation and translation of the pixel's coordinates in the following way

$$\hat{p}_0 = \begin{bmatrix} \hat{X}_0 \\ \hat{Y}_0 \\ \hat{Z}_0 \end{bmatrix} = \underbrace{\begin{bmatrix} a_G \\ r_G \\ 0 \end{bmatrix}}_{\mathcal{T}} + \underbrace{\begin{bmatrix} \cos(\chi) & \sin(\chi) & 0 \\ \sin(\chi) & -\cos(\chi) & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathcal{R}(\chi)} \begin{bmatrix} A \\ R_g \\ h \end{bmatrix} \quad (6)$$

where $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T$ is the navigation frame's position in the global coordinates, \mathcal{T} is the translation vector (consisting of map coordinates obtained by matching procedure) and $\mathcal{R}(\chi)$ is the matrix representation of the rotation around Z -axis with angle χ . Note that Z -coordinate is the same as altitude h , which we in general can obtain directly from the barometric measurements. If these measurements are not available or have bad performance, the approach from [13] can be used to obtain an altitude estimate.

Since the kinematic parameters are also part of the matching results, any position in the trajectory can now be calculated by using Equation (3). Note that these calculations are valid under the flat earth approximation, which is valid if the SAR image

is fairly close to the platform. This is true for most aircraft (but not satellites). The procedure described above is summarised in Algorithm 1. Note that if several patches are needed or wanted to be matched the procedure is simply repeated for all patches. After that a single estimate can be obtained by e.g., calculating a weighted mean from all estimates using their covariance as weights.

Algorithm 1 SAR-OptMAP Global Position Calculation

Require: SAR image coordinates (r, a) , SAR image parameters R_0, Δ_R, Δ_A , information from navigation system $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T, \hat{\chi}$ and optical image or map with known geographical coordinates

Ensure: $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T$ and its covariance

- 1: Calculate A with a, a_{last} and Δ_A from (4)
 - 2: Calculate R with r, R_0 and Δ_R from (5)
 - 3: Match the SAR image to the map image to obtain r_G, a_G, χ and kinematic states using $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T$ and $\hat{\chi}$ to initialise the search
 - 4: Calculate $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T$ with h, A, R_g, χ, r_G and a_G from (6)
 - 5: (Optional) Calculate any trajectory position in the global coordinates, \hat{p}_t , by using model (3) and estimated kinematic parameters
-

IV. IMAGE MATCHING APPROACH

In step 3 of Algorithm 1, we simply defined a matching step which directly delivers the interesting parameters. In this section, one such matching algorithm will be described in more detail since it is a prerequisite in order to obtain the parameters used to calculate a platform's position and orientation. The matching between those images can be obtained in many ways, for example by simple correlation or by using image point features extracted by some point feature detectors, like Harris corner detector [21] or SIFT detector [22]. However, although the SAR and optical map images can share many similarities, in particular over man-made structured environments, they can be very different in their structure and appearance. For example, structures like rooftops can have completely different intensities, very bright in the SAR images and very dark in the optical images, and the above-mentioned methods might not work satisfactory. In this case it might be better to increase the feature complexity one level and use the lines (edges) in the images. See [23] or [24] for examples where edges are used as image features. Exactly as for point features, there are several well known edge detectors, where Sobel, Prewitt and Canny [25] are maybe the most known ones. Since the Canny edge detector is quite robust to noise, it is suggested as the detector in the approach described here. By applying this detector to SAR and optical images, two binary edge images are obtained. The next problem to be solved is to match these SAR and optical map binary edge images to each other. One well known method for parametric matching of templates to the image is so called Chamfer matching method, see [26], [27] or [28]. Since this is quite a robust matching method, it will be the basis of the approach proposed here. Next, a short

description of the Canny edge detector will be given as well as an introduction to Chamfer matching and the modifications we propose for this particular application.

A. Edge Detector

The Canny edge detector uses image gradient and thresholding to detect edges in the images like many other detectors. Its main advantage is better robustness to the noise in the images. This is obtained by using hysteresis with two thresholds, one high and one low. This avoids the problem of broken edges, or streaking, which is almost always present in detectors with only one threshold. The higher threshold is used to detect edges, just as in any detector, while the lower one is used to implement hysteresis and keep an edge even if the gradient response would fall under the higher threshold. The general problem of threshold tuning still remains. Individual thresholds for different images must be found on a case by case basis. In this work we are using an existing Canny edge detector implemented in the Image Processing Toolbox in Matlab.

B. Chamfer Image Matching

The basics of the Chamfer image matching is the distance transform of the edge image to which the template image is to be matched. In this context the template image is not considered to be the actual SAR image but rather the binary image consisting of extracted edges. The distance transform is calculated by assigning the pixels in the binary image a value of the distance to the closest nonzero pixel. The distance metric is usually Euclidean, but also Manhattan distance (1-norm) or even maximum norm can be used. As an illustrative example, consider a simple binary 7×7 image represented as a matrix

$$I = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The distance transform of this image using Euclidean distance is

$$D = \begin{bmatrix} \sqrt{5} & 2 & \sqrt{2} & 1 & \sqrt{2} & 2 & \sqrt{5} \\ \sqrt{2} & 1 & 1 & 0 & 1 & 1 & \sqrt{2} \\ 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 & 1 & 0 & 1 \\ \sqrt{2} & 1 & \sqrt{2} & 1 & 1 & 0 & 1 \\ \sqrt{5} & 2 & 1 & 0 & 0 & 0 & 1 \\ \sqrt{10} & \sqrt{5} & \sqrt{2} & 1 & 1 & 1 & \sqrt{2} \end{bmatrix}$$

Now the idea in Chamfer matching is to overlay the edge pixels of the binary template image, T , on the distance image for different translation, rotation and scaling values and calculate some loss function as some metric based on the values in the distance transform image that are hit by the template edge pixels, for example, the total sum of the values. From the implementation point of view this is equivalent to

taking a whole binary template image as a matrix and element-wise multiplying it with the distance transform image. The reason is simply the fact that edge pixels have value 1 and non-edge ones have value 0. This can be written as

$$\eta(\theta_I) = D \odot \tilde{T}(\theta_I) \quad (7a)$$

$$C(\theta_I) = f(\eta(\theta_I)) \quad (7b)$$

where $\eta(\theta_I)$ is the matrix resulting from the element-wise product (\odot) of the extended template image, \tilde{T} , and the distance transform image, D . In the general case $\theta_I = [r \ a \ \chi \ s_r \ s_c]^T$, where we introduced image scaling parameters s_r and s_c . Sometimes it is possible to take the subset of the θ_I if, for example, some of the parameters are known or not estimated. The extended template image, \tilde{T} , has been created by first rotating the original template image with χ degrees and scaling it s_r and s_c times in row and column directions respectively. The binary image created from this template is then extended with zeros to the size of D in such a way that the upper left corner of the template image is on the coordinate (r, a) . Here, $f : \mathbf{R}^{\text{size}(D)} \rightarrow \mathbf{R}_+$ is some positive and monotonously increasing function. This means that for correct matching parameters, the loss function $C(\theta_I)$ would obtain its minimum value and the parameter estimates are obtained as

$$\hat{\theta}_I = \arg \min_{\theta} C(\theta_I) \quad (8)$$

If the template which is to be matched to the image above is

$$T = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

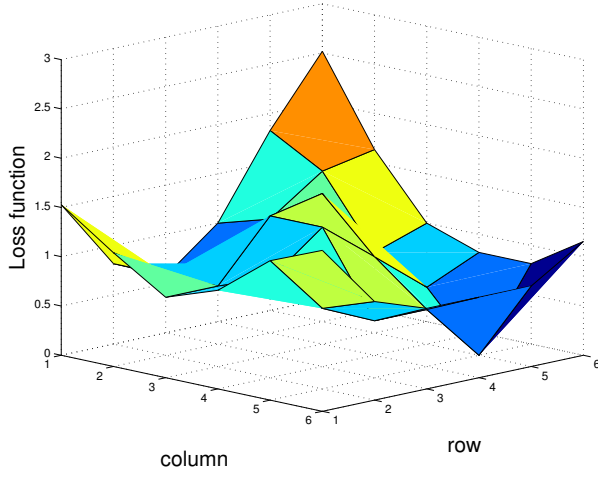
and only translation is considered, i.e., $\theta_I = [r \ a]^T$, the surf plot of the resulting loss function, $C(\theta_I)$, is according to Figure 5a. It can be seen that the minimum value is obtained for the translation parameters $r = 5$ (row) and $a = 5$ (column) which is the best possible match. The function f used here is the RMSE value of the element-wise product of the distance transform image and the extended template image

$$C(r, a) = \sqrt{\frac{1}{N_{\text{nz}}} \sum_{k=1}^7 \sum_{l=1}^7 \eta_{k,l}(r, a)^2}, \quad (9)$$

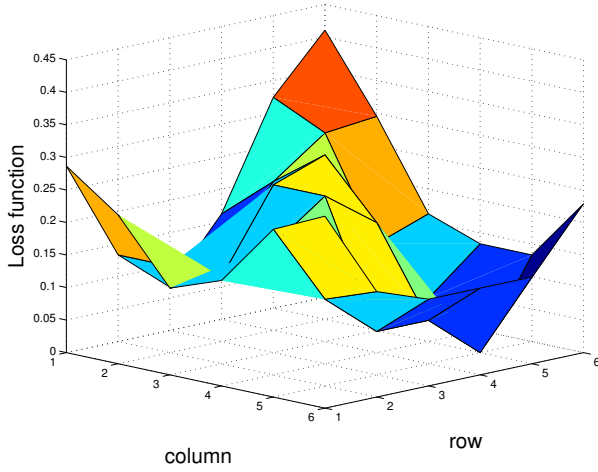
where $\eta_{k,l}(r, a)$ is the matrix of values according to (7a), except that rotation and scaling are not considered. Here, N_{nz} is the amount of nonzero elements in the extended template image, in this example $N_{\text{nz}} = 3$. If rotation, χ , or scalings, s_r and s_c , are also considered as unknown parameters, the example search illustrated above must be repeated for each considered value of the rotation and scalings. In this case the solution is given by the minimum value of the total cost function.

C. Modified Matching Approach

In this work, a slightly modified loss function is proposed, which bears more similarity to the well known least squares approach. The reason is that we need an uncertainty measure to the position estimate, otherwise higher level fusion with



(a) Chamfer matching loss function, $C(r, a)$, of the matching example.



(b) Nonlinear least squares loss function, $V(r, a)$, of the matching example.

Figure 5: Different loss functions for the matching example.

the on-board navigation system would be problematic. To get a statistically correct measure of covariance is a complicated problem, but at least we get a matrix that has the most essential properties of a covariance matrix: it is a positive definite symmetric matrix, it reveals parameter identifiability problems by having a high condition number, and it shows relative size of estimation errors by having different size of the diagonal elements.

First, the distance transformed binary map image, D , is transformed as

$$\tilde{D} = \exp(-D) \quad (10)$$

where $\exp(\cdot)$ function acts elementwise. This will basically “invert” the distance transform making zero valued pixels become ones and high valued pixels become low valued. For

the example above \tilde{D} looks like

$$\tilde{D} = \begin{bmatrix} 0.11 & 0.14 & 0.24 & 0.37 & 0.24 & 0.14 & 0.11 \\ 0.24 & 0.37 & 0.37 & 1 & 0.37 & 0.37 & 0.24 \\ 0.37 & 1 & 0.37 & 0.37 & 0.37 & 1 & 0.37 \\ 0.37 & 1 & 0.37 & 0.14 & 0.37 & 1 & 0.37 \\ 0.24 & 0.37 & 0.24 & 0.37 & 0.37 & 1 & 0.37 \\ 0.11 & 0.14 & 0.37 & 1 & 1 & 1 & 0.37 \\ 0.04 & 0.11 & 0.24 & 0.37 & 0.37 & 0.37 & 0.24 \end{bmatrix}$$

Let $\xi(\theta_I)$ be the $N_{nz} \times 1$ vector of values from \tilde{D} hit by the translated, rotated and scaled edge pixels of the binary template, $T(\theta_I)$. Then we have the following relation

$$\mathbf{1}_{N_{nz}} = \xi(\theta_I) + e \quad (11)$$

where $\mathbf{1}_{N_{nz}}$ is the $N_{nz} \times 1$ vector of ones and e is some noise. This relation can be interpreted as a measurement equation which is a function of parameter vector θ_I , and then the minimisation criterion can be written as

$$\hat{\theta}_I = \arg \min_{\theta_I} V(\theta_I), \quad (12a)$$

$$V(\theta_I) = \frac{\|\mathbf{1}_{N_{nz}} - \xi(\theta_I)\|_2^2}{2N_{nz}} = \frac{1}{2N_{nz}} \sum_{k=1}^{N_{nz}} (1 - \xi_k(\theta_I))^2 \quad (12b)$$

which is a nonlinear least squares formulation. The modified loss function, $V(r, a)$, for the example is depicted in Figure 5b. This loss function has a very similar shape as the original one, $C(r, a)$, but it is a little bit steeper close to the minimum. Since both of these loss functions are defined on a grid of discrete values, the minimisation procedure can be performed as a global grid search.

Besides the parameter values, $\hat{\theta}_I$, it is also desirable to estimate the covariance which in turn can be used to estimate the covariance of the estimated navigation parameters, position of the navigation frame, $[\hat{X}_0 \ \hat{Y}_0 \ \hat{Z}_0]^T$, and track angle, $\hat{\chi}$. These covariances can then be used for weighting purpose in Algorithm 1. The covariance can be estimated by assuming a locally quadratic function around the minimum value of the loss function, $V(\hat{\theta}_I)$, and estimating the Hessian matrix, H . This can be done by solving the overdetermined linear system of equations originating from the following relation

$$V(\hat{\theta}_I + \Delta) \approx V(\hat{\theta}_I) + \Delta^T H \Delta \quad (13)$$

where a Taylor expansion around $\hat{\theta}_I$ is performed for some Δ assuming that the gradient is zero (since $V(\hat{\theta}_I)$ is a stationary point, it is the minimum value). Then the covariance of the parameter estimates can be estimated as

$$\text{Cov}(\hat{\theta}_I) = \hat{\lambda} H^{-1} \quad (14)$$

where $\hat{\lambda} = V(\hat{\theta}_I)$, see [29]. Note that in the example above we obtain the covariance which equals zero for both parameters and it is natural since the template fits perfectly, and there is no uncertainty. In the general case, however, the template will not fit perfectly and there will always be some uncertainty in the estimates.

V. KINEMATIC PARAMETER ESTIMATION

The procedure described in Section IV considers how to estimate the image parameters θ_I given a SAR image produced with some trajectory $\hat{p}_{0:N}$. By varying the values of the kinematic part of the parameter vector, θ_K , and using the model (3) and possibly accelerations measured by the onboard IMU, different SAR images $I(\theta_K)$ can be obtained. Each of these images can now be matched according to the solution of Equation (12) which produces another loss function,

$$J(\theta_K, \hat{\theta}_I) = V_{\theta_K}(\hat{\theta}_I), \quad (15)$$

where $V_{\theta_K}(\hat{\theta}_I)$ is the value of the image matching loss function obtained for a SAR image created with θ_K as kinematic parameters. The loss function J is, exactly as V , a non-convex function with many local minima, implying that a grid search is the best option to find a solution according to

$$\hat{\theta}_K = \arg \min_{\theta_K} J(\theta_K, \hat{\theta}_I). \quad (16)$$

This will give the total solution $\hat{\theta} = [\hat{\theta}_I^T \hat{\theta}_K^T]^T$, with the best focus in this metric, and an accurate global position. The covariance estimation approach from Equations (13) and (14) can be used here as well in order to obtain covariance of the kinematic parameters.

VI. RESULTS

In this section we will present both the results for the image matching approach and the kinematic parameter estimation. The image matching approach will be presented in more detail on two example patches from the SAR image assuming the focused image. For the kinematic parameter estimation, a low-frequency SAR simulation environment, CARABAS II in particular, [30], is used where it is possible to vary trajectories and create different SAR images. This environment makes it possible to create realistic SAR images of the scene and to evaluate the focusing results in a controlled manner.

A. Results of the Image Matching Approach

In order to show the results of the matching procedure described in Section IV, two patches from the SAR image in Figure 1 are matched to the optical image in Figure 2. These two patches are depicted in Figure 6. The optimised parameters are translations and rotation, and the scaling is fixed beforehand to minimise the parameter space and speed up the search. It is the optical image that has been fixed north up and the SAR image that has been rotated. In that case the flight direction angle, χ , is directly obtained. It should also be pointed out that in the search for the matching parameters, information from the navigation system is used to narrow down the search space and in that way prune possible false solutions due to a possibly too similar environment. The results are presented both graphically, where SAR image patches are overlaid on the optical image, and in a table with an error and a standard deviation of the estimates.

The SAR patches rotated with the angle $\hat{\chi}$ obtained in the optimisation are depicted in Figures 7a and 8a. The result of the matching is depicted in Figures 7c and 8c by overlaying

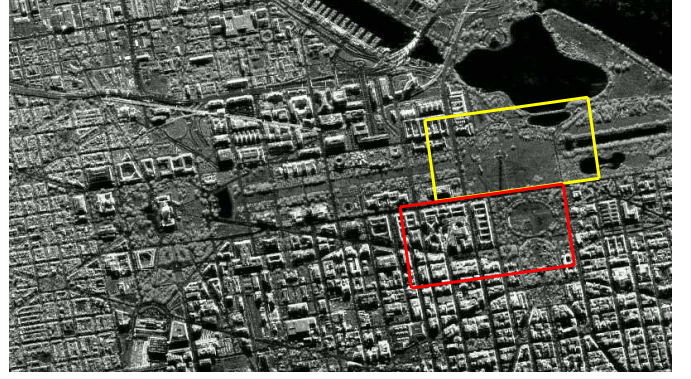


Figure 6: Zoomed part of the SAR image with the two patches used in the Chamfer matching procedure.

Parameter	Error	Standard Deviation
r [pixels]	0, 2	5.73, 7.73
a [pixels]	0, -3	6.82, 7.40
χ [degrees]	1, 1	5.26, 6.49

Table I: Errors and standard deviations of the parameter estimates for the two different example SAR patches.

the SAR image patch on the optical image on the solution pixels, (\hat{r}, \hat{a}) . The errors and the standard deviations of the estimates for these two cases are presented in Table I.

B. Results for the Kinematic Parameters Estimation

For the kinematic parameter estimation results, one patch from the SAR image is chosen, see Figure 9a. Also, the same patch created with a non-correct trajectory, causing defocusing, is depicted in Figure 9b. These patches are produced with the CARABAS II simulation environment, and it can be seen that they are quite realistic. First, a simple case where initial velocity in X -direction is unknown while all the other parameters are known is examined. In this case, $\theta_K = v_0^X$, and the error in speed was varied, first between -4% and 4% . The resulting loss function $J(\theta_K, \hat{\theta}_I)$ is depicted in Figure 10. Here it can be seen that the minimum value is obtained for the correct initial velocity, and furthermore there is no matching error. Note that grid has higher resolution in the middle of the plot. Another simple case that is examined concerns variation of initial acceleration in the Y -direction, while all the other parameters are known. In this case $\theta_K = a_0^Y$ and the resulting loss function is presented in Figure 11. Even here the minimum value is obtained for the correct value of the acceleration, and there was no matching error in this case either. In both cases it can be seen that loss function has much more irregular behaviour close to the optimum and that some of the loss function values are close to the value for the correct acceleration. If both initial velocity in X -direction and acceleration in Y -direction are set as parameters, i.e., $\theta_K = [v_0^X \ a_0^Y]^T$, the loss function is according to Figure 12 for both large and small error in states. In this case the global minimum of the loss function is obtained for a correct value if the error was large, i.e., the grid resolution was coarse, while a non-correct value of the parameters is obtained for the small

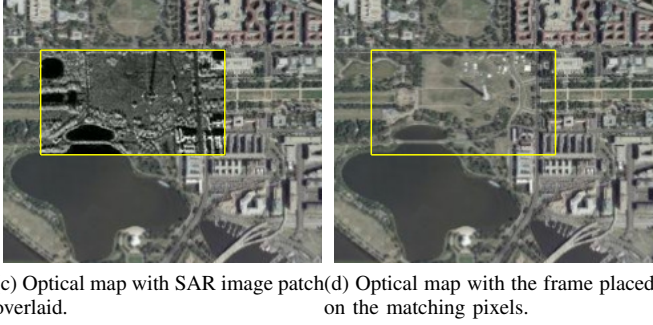
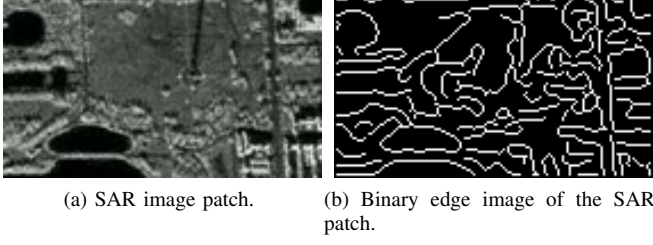


Figure 7: Example 1: Illustration of the proposed matching procedure starting with the SAR image patch, going through extracted edge image and showing the results of the matching by overlaying SAR patch on the optical image map. Also, only a frame is placed on the optical map for comparison.

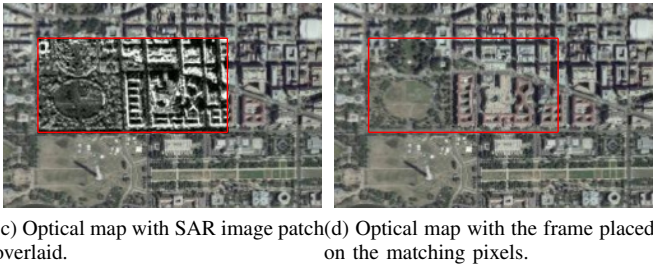
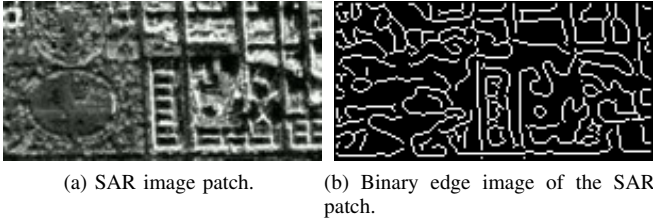


Figure 8: Example 2: Illustration of the proposed matching procedure starting with the SAR image patch, going through extracted edge image and showing the results of the matching by overlaying SAR patch on the optical image map. Also, only a frame is placed on the optical map for comparison.

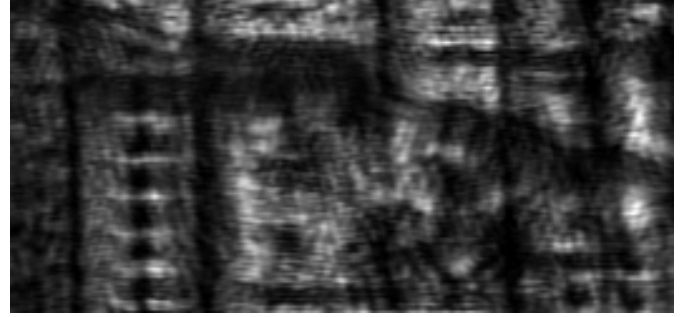


Figure 9: SAR image patch used for the evaluation of the kinematic parameters estimation. Both focused and unfocused images are shown.

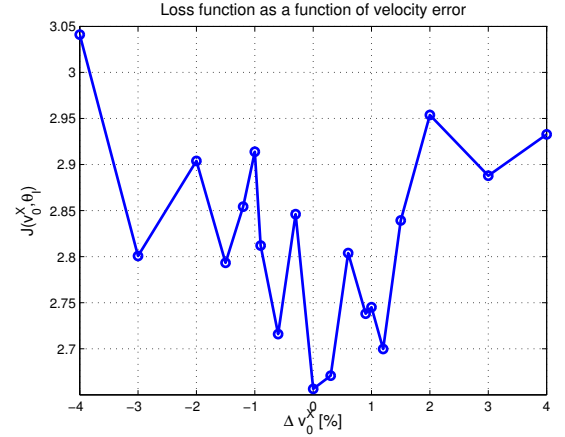


Figure 10: The value of the loss function as a function of a percentual error in initial velocity in X-direction.

error case, $\hat{\theta}_K = [-0.30 \ -0.02]^T$ ($[\% \text{ m/s}^2]$) although the correct value was the second smallest. This error is, however, quite small giving a trajectory RMSE error of about 6 m and the SAR image patch resulting from this trajectory is shown in Figure 13. The matching errors in this case are 1 and 2 pixels in range and azimuth directions, respectively, and the error in rotation of the patch is 0.5° . We see that the actual difference in the image quality is hard to distinguish with the naked eye, and that the navigation parameter estimates are also quite good. Note also that in all cases it is obvious that the loss function is highly non-convex and that a grid based

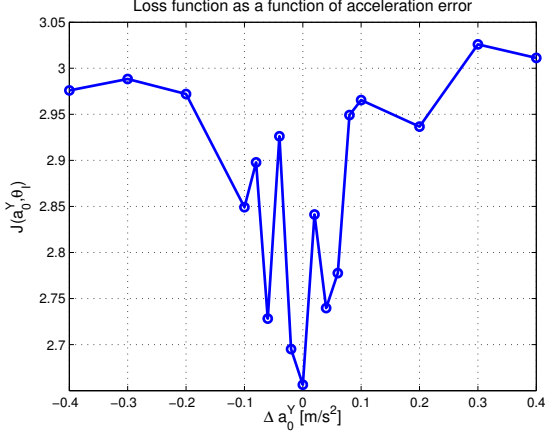
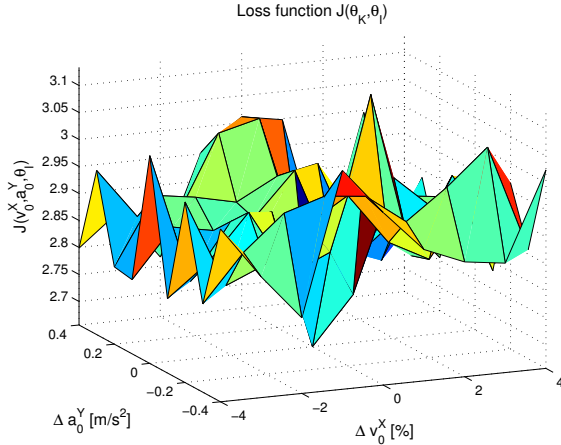
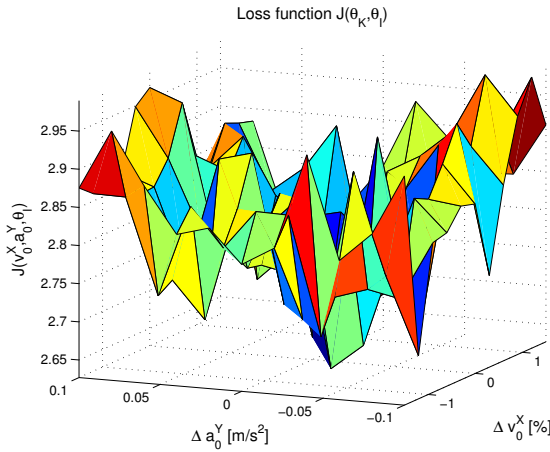


Figure 11: The value of the loss function as a function of an error in initial acceleration in Y -direction.



(a) Loss function for the large initial velocity and acceleration error.



(b) Loss function for the small initial velocity and acceleration error.

Figure 12: The value of the loss function as a function of an error in initial velocity in X -direction and acceleration in Y -direction.

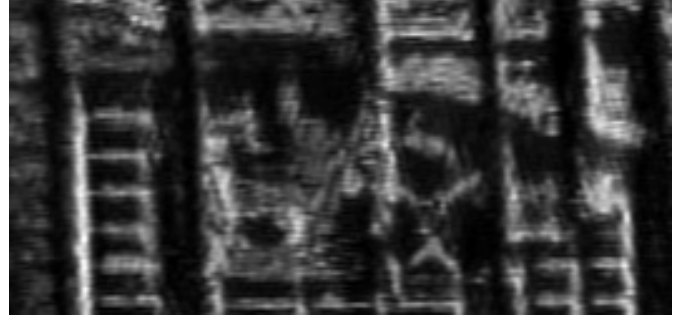


Figure 13: SAR image patch obtained with the estimated trajectory given the values from the small error case.

search is necessary. This also implies that the grid resolution will set the accuracy limit, and the number of operations grows exponentially with the number of grid points. However, it is possible to evaluate each grid point individually, which suits parallel computation architectures.

C. Discussion

In this section a discussion on the method's results and performance and some possible improvements is provided. The performance of the method for both patch matching and kinematic parameter estimation is highly dependent on different quantization effects. Both optical and SAR images have finite pixel resolution. This sets the limit on the performance of the edge detector and on the matching performance. Furthermore, the standard implementation of the Canny detector uses greyscale images which implies that additional quantization is present. All these effects will put the limit on matching and rotation estimation performance. The obtained performance for the examples studied here is fairly good and in the magnitude of what can be expected.

For the kinematic parameters, the grid resolution will naturally set the limit on the performance. The finer grid, the better possibility to get good performance. But if the resolution is too small, the difference in the trajectories created with neighboring grid parameter values will not be enough to make SAR images different enough from the focusing point of view. Then, in practice, only the numerical accuracy and their effects will dominate. The grid size and resolution set also the limit on the execution speed. Therefore the grid resolution and size are seen as tuning parameters.

In [31], an auto-focusing approach based on the SAR image only is exploited using image entropy as a focus measure. The entropy is here defined as

$$E = - \sum_{i=1}^M \sum_{j=1}^K q_{ij} \log q_{ij} \quad (17a)$$

$$q_{ij} = \frac{|I_{ij}|^2}{\sum_k \sum_l |I_{kl}|^2} \quad (17b)$$

where I_{ij} is the complex-valued pixel (i, j) in the SAR image. Note that E is a function of θ_K . Then, we could combine the entropy and the loss function $J(\theta)$ around the global minimum to improve the estimation results. In the case used above, the

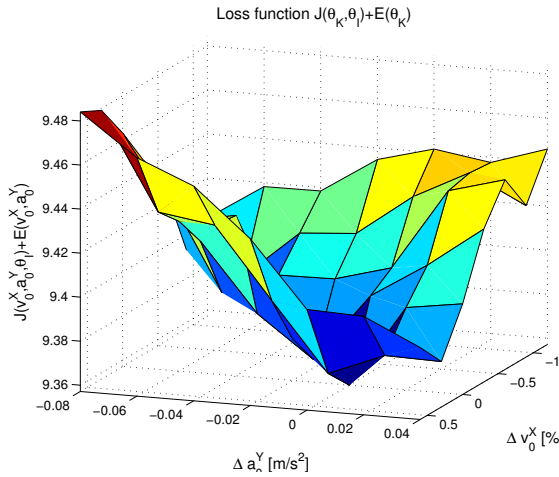


Figure 14: Combined entropy, E , and loss function, J , around the global minimum of the loss function.

combined loss function $J + E$ is shown in Figure 14. We see that for this function a correct value of the parameters is obtained.

VII. CONCLUSIONS AND FUTURE WORK

A method of matching SAR images and optical images is presented, for the primary purpose of autofocus and adding the radar reflectance image to ordinary images as complementary information, which can be useful in, primarily, surveillance, and, in some cases, remote sensing applications. The method is based on the pattern matching algorithm called Chamfer matching, which is modified to resemble a least squares formulation and a grid based optimisation of the kinematic parameters. For both cases a statistical performance measure, covariance, of the estimates can also be obtained. The evaluation of the results is performed on the SAR image and optical map image, and both matching performance and autofocus performance is evaluated based on a couple of SAR image patches. The obtained results on the real SAR images and very simple optical map images from Google Maps show that the performance of the matching and autofocus methods is fairly good, with small errors and variance, even with these simple means. This performance is good enough for surveillance, but probably not for many remote sensing applications. For that kind of applications a smaller matching error should be obtained. It must also be pointed out that this method assumes a variation in the scene in order to work. The environment where edge features are hard to extract or missing will of course give much poorer results.

As a future extension of this work, as preliminary results show in the discussion above, entropy measure can be incorporated in the total cost function to eventually improve the results of the kinematic parameters estimation and in turn the autofocus performance. As a future application, the methodology can also be used as an all-weather GNSS-like support and backup for the inertial navigation system.

ACKNOWLEDGEMENTS

This work has been supported by the Industry Excellence Center LINK-SIC founded by The Swedish Governmental Agency for Innovation Systems (VINNOVA) and Saab AB. The thanks also go to Dr. Lars Ulander at Swedish Defense Research Agency (FOI) for the CARABAS II simulation environment.

REFERENCES

- [1] L. J. Cutrona, W. E. Vivian, E. N. Leith, and G. O. Hall, "A high-resolution radar combat-surveillance system," *IRE Transactions on Military Electronics*, vol. MIL-5, no. 2, pp. 127–131, April 1961.
- [2] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, ser. The SciTech Radar and Defense Series. SciTech, 2004.
- [3] D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and C. V. J. Jakowatz, "Phase gradient autofocus - a robust tool for high resolution SAR phase correction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, no. 3, pp. 827–835, July 1994.
- [4] M. Xing, R. Jiang, X. and Wu, F. Zhou, and Z. Bao, "Motion Compensation for UAV SAR Based on Raw Radar Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 8, pp. 2870–2883, August 2009.
- [5] J. R. Fienup, "Phase Error Correction by Shear Averaging," in *Signal Recovery and Synthesis*. Optical Society of America, June 1989, pp. 134–137.
- [6] R. L. J. Morrison and D. C. J. Munson, "An experimental study of a new entropy-based SAR autofocus technique," in *Proceedings of International Conference on Image Processing, ICIP 2002*, vol. 2, September 2002, pp. II–441–4.
- [7] L. Xi, L. Guosui, and J. Ni, "Autofocusing of ISAR images based on entropy minimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 4, pp. 1240–1252, October 1999.
- [8] M. Greco, G. Pinelli, K. Kulpa, P. Samczynski, B. Querry, and S. Querry, "The study on SAR images exploitation for air platform navigation purposes," in *Proceedings of the International Radar Symposium (IRS)*, 2011, sept. 2011, pp. 347–352.
- [9] L. Hostettler and R. Andreas, "Nonlinear Kalman filtering techniques for terrain-aided navigation," *Automatic Control, IEEE Transactions on*, vol. 28, no. 3, pp. 315–323, 1983.
- [10] D. Scaramuzza and F. Fraundorfer, "Visual Odometry [Tutorial]," *Robotics Automation Magazine, IEEE*, vol. 18, no. 4, pp. 80–92, dec. 2011.
- [11] F. Lindsten, J. Callmer, H. Ohlsson, D. Törnqvist, T. Schön, and F. Gustafsson, "Geo-referencing for UAV navigation using environmental classification," in *Robotics and Automation (ICRA), 2010 IEEE International Conference on*, may. 2010, pp. 1420–1425.
- [12] B. Grelsson, M. Felsberg, and F. Isaksson, "Efficient 7D Aerial Pose Estimation," in *IEEE Workshop on Robot Vision 2013, Clearwater Beach, Florida, USA, January 16-17, 2013*, 2013.
- [13] Z. Sjanic and F. Gustafsson, "Fusion of Information from SAR and Optical Map Images for Aided Navigation," in *Proceedings of 15th International Conference on Information Fusion*, Singapore, July 2012.
- [14] J. A. Fawcett, "Inversion of N-Dimensional Spherical Averages," *SIAM Journal on Applied Mathematics*, vol. 45, no. 2, pp. 336–341, 1985. [Online]. Available: <http://www.jstor.org/stable/2101820>
- [15] H. Hellsten and L. E. Andersson, "An inverse method for the processing of synthetic aperture radar data," *Inverse Problems*, vol. 3, no. 1, p. 111, 1987. [Online]. Available: <http://stacks.iop.org/0266-5611/3/i=1/a=013>
- [16] L. E. Andersson, "On the Determination of a Function from Spherical Averages," *SIAM Journal on Mathematical Analysis*, vol. 19, no. 1, pp. 214–232, 1988. [Online]. Available: <http://link.aip.org/link/?SJM/19/214/1>
- [17] C. Cafforio, C. Prati, and F. Rocca, "SAR data focusing using seismic migration techniques," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, no. 2, pp. 194–207, March 1991.
- [18] F. Rocca, "Synthetic Aperture Radar: a New Application for Wave Equation Techniques," *Stanford Exploration Project SEP-56*, pp. 167–189, 1987.
- [19] A. S. Milman, "SAR Imaging by Omega-K Migration," *International Journal of Remote Sensing*, vol. 14, no. 10, pp. 1965–1979, 1993.
- [20] F. Natterer, *The Mathematics of Computerised Tomography*. New York: Wiley, 1986.

- [21] C. Harris and M. Stephens, "A combined corner and edge detection," in *Proceedings of The Fourth Alvey Vision Conference*, 1988, pp. 147–151.
- [22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, pp. 91–110, November 2004.
- [23] J. D. Wagner, "Automatic Fusion of SAR and Optical Imagery," September 2007, Master's Thesis, Leibniz Universität, Hannover, Germany.
- [24] C. Taylor and D. Kriegman, "Structure and motion from line segments in multiple images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 11, pp. 1021–1032, nov. 1995.
- [25] J. Canny, "A Computational Approach to Edge Detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-8, no. 6, pp. 679–698, November 1986.
- [26] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf, "Parametric correspondence and chamfer matching: two new techniques for image matching," in *Proceedings of the 5th international joint conference on Artificial intelligence - Volume 2*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1977, pp. 659–663. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1622943.1622971>
- [27] G. Borgefors, "Hierarchical Chamfer Matching: A Parametric Edge Matching Algorithm," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 10, pp. 849–865, November 1988. [Online]. Available: <http://dl.acm.org/citation.cfm?id=56917.56924>
- [28] J. Ericsson and A. Thid, "Automatic SAR Image to Map Registration," January 2006, Master's Thesis, Chalmers, Göteborg, Sweden.
- [29] F. Gustafsson, L. Ljung, and M. Millnert, *Digital Signal Processing*. Lund: Studentlitteratur, 2010.
- [30] H. Hellsten, L. M. Ulander, A. Gustavsson, and B. Larsson, "Development of VHF CARABAS II SAR," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 2747, June 1996, pp. 48–60.
- [31] Z. Sjanic and F. Gustafsson, "Simultaneous Navigation and Synthetic Aperture Radar Focusing," Department of Electrical Engineering, Linköping University, Tech. Rep. LiTH-ISY-R-3063, 2013.



Zoran Sjanic received the M.Sc. degree in computer science and engineering 2001 and PhD in Automatic Control in 2013 both from Linköping University. His research interests are sensor fusion for navigation of manned and unmanned aircraft, Simultaneous Localisation and Mapping (SLAM) and nonlinear estimation methods.

He is also employed by Saab Aeronautics in Linköping, Sweden, since 2001 where he works at the Navigation Department as a system engineer.

He also worked as a technical manager for the navigation system in both Gripen fighter aircraft and Skeldar UAV before starting his PhD studies in 2008.



Fredrik Gustafsson is professor in Sensor Informatics at Department of Electrical Engineering, Linköping University, since 2005. He received the M.Sc. degree in electrical engineering 1988 and the Ph.D. degree in Automatic Control, 1992, both from Linköping University. During 1992-1999 he held various positions in automatic control, and 1999-2005 he had a professorship in Communication Systems. His research interests are in stochastic signal processing, adaptive filtering and change detection, with applications to communication, vehicular, air-

borne, and audio systems. He is a co-founder of the companies NIRA Dynamics (automotive safety systems), Softube (audio effects) and SenionLab (indoor positioning systems).

He was an associate editor for IEEE Transactions of Signal Processing 2000-2006 and is currently associate editor for IEEE Transactions on Aerospace and Electronic Systems and EURASIP Journal on Applied Signal Processing. He was awarded the Arnberg prize by the Royal Swedish Academy of Science (KVA) 2004, elected member of the Royal Academy of Engineering Sciences (IVA) 2007, elevated to IEEE Fellow 2011 and awarded the Harry Rowe Mimno Award 2011 for the tutorial "Particle Filter Theory and Practice with Positioning Applications", which was published in the AEES Magazine in July 2010.