

RADAR SLAM using Visual Features

Jonas Callmer, David Törnqvist, Fredrik Gustafsson, Henrik Svensson and Pelle Carlbom

Linköping University Post Print

N.B.: When citing this work, cite the original article.

The original publication is available at www.springerlink.com:

Jonas Callmer, David Törnqvist, Fredrik Gustafsson, Henrik Svensson and Pelle Carlbom, RADAR SLAM using Visual Features, 2011, EURASIP Journal on Advances in Signal Processing, (2011), 71, .

<http://dx.doi.org/10.1186/1687-6180-2011-71>

Licensee: Hindawi Publishing Corporation / Springer Verlag (Germany) / SpringerOpen

<http://www.springeropen.com/>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-72583>

RESEARCH

Open Access

Radar SLAM using visual features

Jonas Callmer^{1*}, David Törnqvist¹, Fredrik Gustafsson¹, Henrik Svensson² and Pelle Carlbom³

Abstract

A vessel navigating in a critical environment such as an archipelago requires very accurate movement estimates. Intentional or unintentional jamming makes GPS unreliable as the only source of information and an additional independent supporting navigation system should be used. In this paper, we suggest estimating the vessel movements using a sequence of radar images from the preexisting body-fixed radar. Island landmarks in the radar scans are tracked between multiple scans using visual features. This provides information not only about the position of the vessel but also of its course and velocity. We present here a navigation framework that requires no additional hardware than the already existing naval radar sensor. Experiments show that visual radar features can be used to accurately estimate the vessel trajectory over an extensive data set.

1. Introduction

In autonomous robotics, there is a need to accurately estimate the movements of a vehicle. A simple movement sensor like a wheel encoder on a ground robot or a pit log on a vessel will under ideal circumstances provide quite accurate movement measurements. Unfortunately, they are sensitive to disturbances. For example, wheel slip due to a wet surface will be interpreted incorrectly by a wheel encoder, and strong currents will not be correctly registered by the pit log why a position estimate based solely on these sensors will drift off. In applications like autonomous robotics, the movement accuracy needs to be high why other redundant movement measurement methods are required.

A common approach is to study the surroundings and see how they change over time. By relating the measurements of the environment k seconds ago to the present ones, a measurement of the vehicle translation and rotation during this time interval can be obtained. A system like this complements the movement sensor and enhances the positioning accuracy.

Most outdoor navigation systems such as surface vessels use global navigation satellite systems (GNSS) such as the Global Positioning System (GPS) to measure their position. These signals are weak making them very vulnerable to intentional or unintentional jamming [1-3]. A supporting positioning system that is redundant of the satellite signals is therefore necessary. By estimating the

vessel movements using the surroundings, a mean of measuring the reliability of the GPS system is provided. The movement estimates can also be used during a GPS outage providing accurate position and movement estimates over a limited period of time. This support system could aid the crew in critical situations during a GPS outage, avoiding costly PR disasters such as running aground.

For land-based vehicles or surface vessels, three main sensor types exist that can measure the environment: cameras, laser range sensors and radar sensors. Cameras are very rich in information and have a long reach but are sensitive to light and weather conditions. Laser range sensors provide robust and accurate range measurements but also they are very sensitive to weather conditions. The radar signal is usually the least informative signal of the three and is also quite sensitive to what the signals reflect against. On the other hand, the radar sensor works pretty much equally well in all weather conditions.

In this paper, radar scan matching to estimate relative movements is studied. The idea is to use the radar as an imagery sensor and apply computer vision algorithms to detect landmarks of opportunity. Landmarks that occur during consecuting radar scans are then used for visual odometry, that gives speed, relative position and relative course. The main motivation for using visual features to match radar scans instead of trying to align the radar scans is that visual features are easily matched despite large translational and rotational differences, which is more difficult using other scan matching techniques.

* Correspondence: callmer@isy.liu.se

¹Division of Automatic Control, Linköping University Linköping, Sweden
Full list of author information is available at the end of the article

The landmarks can optionally be saved in a map format that can be used to recognize areas that have been visited before. That is, a by-product of the robust navigation solution is a mapping and exploration system.

Our application example is based on a military patrol boat, Figure 1, that often maneuvers close to the shore in high speeds, at night, without visual aid in situations where GPS jamming or spoofing cannot be excluded. As the results will show, we are able to navigate in a complex archipelago using only the radar and get a map that is very close to ground truth.

To provide a complete backup system for GPS, global reference measurements are necessary to eliminate the longterm drift. The surface navigation system in [4,5], assumed that an accurate sea chart is available. The idea was to apply map matching between the radar image and the sea chart, and the particle filter was used for this mapping. Unfortunately, commercial sea charts still contain rather large absolute errors of the shore, see [1,2], which makes them less useful in blind navigation with critical maneuvers without visual feedback.

The radar used in these experiments measures the distances to land areas using 1,024 samples in each direction, and a full revolution is comprised of roughly 2,000 directions. Each scan has a radius of about 5 km giving a range resolution of roughly 5 m. These measurements are used to create a radar image by translating the range and bearing measurements into Cartesian coordinates. An example of the resulting image is shown in Figure 2.

The radar image gives a birds eye view of the surrounding islands and by tracking these islands, information about how the vessel is moving is obtained. We use the Scale-Invariant Feature Transform (SIFT) [6] to extract trackable features from the radar image which are subsequently matched with features from later scans. These features are shown to be distinct and stable enough to be used for island tracking. Other feature detectors like Speeded Up Robust Features (SURF) [7]



Figure 1 The high-speed patrol boat type used for the data acquisition. Note the backwash created by the jet propulsion system. Courtesy of Dockstavarvet AB.

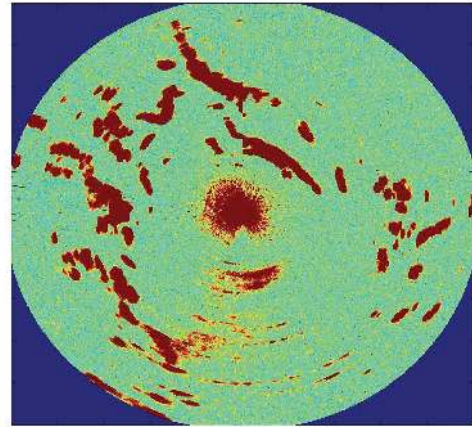


Figure 2 Typical radar image showing the islands surrounding the vessel. The radar disturbances close to the vessel are caused by the vessel and the waves. Behind the vessel (lower part of the image), the striped-shaped disturbances are the result of backscatters reflecting the radar pulses.

could equally well have been used. When these features are tracked using a filter, estimates of the vessel movements are obtained that over time give an accurate trajectory estimate.

The outline is as follows; Section II gives an overview of the related work followed by a theoretical filtering framework in Section III. In Section IV, the performance of SIFT is evaluated on radar images, and the trajectory estimation performance on experimental data is given in Section V. The paper then ends in Section VI with conclusions and suggested future work.

II. Background and relation to slam

The approach in this contribution is known as the Simultaneous Localization And Mapping (SLAM) problem. Today, SLAM is a fairly well-studied problem with solutions that are reaching some level of maturity [8,9]. SLAM has been performed in a wide variety of environments such as indoors [10], in urban [11-14] and rural areas [13,15], underwater [16,17] and in the air [18] and the platform is usually equipped with a multitude of sensors such as lasers, cameras, inertial measurement units, wheel encoders, etc. In this work, we will use only the radar sensor of a naval vessel to perform SLAM in a maritime environment. The data used were recorded in the Stockholm archipelago by Saab Bofors Dynamics [19].

Radars have been used for a long time to estimate movements, for example in the early experiments by Clark and Durrant-Whyte [20]. Radar reflecting beacons in known positions were tracked using a millimeter

radar, and this was shown to improve the movement estimates. Shortly after, Clark and Dissanayake [21] extended the work by tracking natural features instead of beacons.

Thereafter, laser range sensors became more popular since they are more reliable, giving a range measurement in all directions. The problem of estimating the vehicle movements became a problem of range scan alignment. This was studied among others in [22-25].

The advantages of the radar, such as its ability to function in all weather conditions, have though resulted in it making a comeback. Lately, microwave radars have been used in SLAM experiments but now using a landmark free approach. In [13], SLAM was performed in both urban and rural areas by aligning the latest radar scan with the radar map using 3D correlations to estimate the relative movements of the vehicle. The radar map was constructed by consecutively adding the latest aligned radar scan to the previous scans. Checchin et al. [14] performed SLAM in an urban scenario by estimating the rotation and translation of the robot over a sequence of scans using the Fourier-Mellin Transform. It can match images that are translated, rotated and scaled and can therefore be used to align radar scans [26]. Chandran and Newman [27] jointly estimated the radar map and the vehicle trajectory by maximizing the quality of the map as a function of a motion parametrization.

Millimeter wave radars have also become more commonplace in some segments of the automotive industry, and the number of applications for them are growing. For example, the road curvature has been estimated using the radar reflections that will be used in future systems in collision warning and collision avoidance [28,29].

The problem of radar alignment is also present in meteorology where space radar and ground radar observations are aligned to get a more complete picture of the weather in [30]. The scans are aligned by dividing them into smaller volumes that are matched by their respective precipitation intensities.

Visual features like SIFT or SURF have been used in camera-based SLAM many times before. Sometimes, the features were used to estimate relative movements [16,31,32], and other times, they were used to detect loop closures [10,17,33,34].

The combination of radar and SIFT has previously been explored by Li et al. in [35], where Synthetic Aperture Radar measurements were coregistered using matched SIFT features. Radar scan matching using SIFT was also suggested in the short papers [36,37]. A system with parallel stationary ground radars is discussed and SIFT feature matching is suggested as a way to estimate the constant overlaps between the scans. No radar scans ever seem to be matched in those papers though. To the best of the authors knowledge, this is the first time

visual features have been used to estimate the rotational and translational differences between radar images.

III. Theoretical framework

All vessel movements are estimated relative a global position. The positions of the tracked landmarks are not measured globally but relative to the vessel. Therefore, two coordinate systems are used, one global for positioning the vessel and all the landmarks and one local relating the measured feature positions to the vessel. Figure 3 shows the local and global coordinate systems, the vessel and a landmark m .

The variables needed for visual odometry are summarized in Table 1.

A. Detection model

Each radar scan has a radius of about 5 km with a range resolution of 5 meters, and the antenna revolution takes about 1.5 s.

If a landmark is detected at time t , the radar provides a range, r_t , and bearing, θ_t , measurement to the island landmark i as

$$y_t^i = \begin{pmatrix} r_t^i \\ \theta_t^i \end{pmatrix} + e_t^i \quad (1)$$

where e_t^i is independent Gaussian noise. These echos are transformed into a radar image using polar to rectangular coordinates conversion, and the result is shown in Figure 2. Figures 1 and 2 also show that the forward and sideways facing parts of the scans are the most useful ones for feature tracking. This is due to the

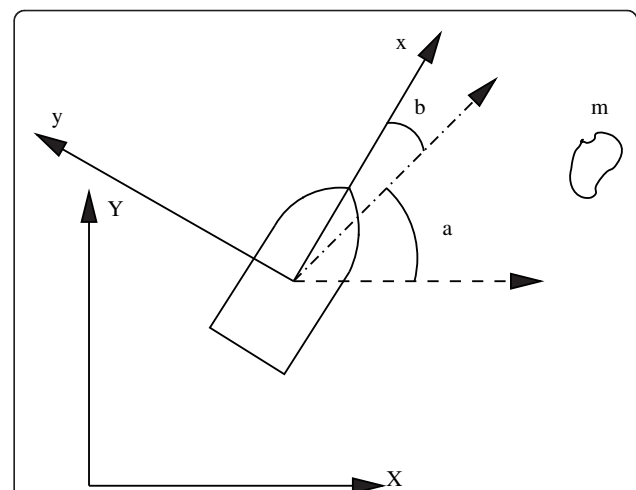


Figure 3 The global (X, Y) and the local boat fixed (x, y) coordinate systems, the course ψ and the crab angle ϕ giving difference between the course and velocity vector. The vessel and an island landmark m are also depicted as is the measured bearing θ and range r of the landmark relative to the vessel.

Table 1 Summary of notation

Parameter	Description
(X, Y)	Global position
(x, y)	Local coordinate system with x aligned with the stem
v	Velocity
ψ	Course, defined as the angle between the global X axis and the local x axis, as would be shown by a compass
ω	Turn rate, defined as $\dot{\psi}$
ϕ	Difference between course and velocity vector ("crab" angle), which is mainly due to streams and wind
(m_X^i, m_Y^i)	Global position of landmark i
r^i	Range from the strap-down body-fixed radar to landmark i
θ^i	Bearing from the radar to landmark i

significant backwash created by the jet propulsion system of the vessel, which is observed along the vessel trajectory in Figure 1. This backwash disturbs the radar measurements by reflecting the radar pulse, resulting in the stripe-shaped disturbances behind the vessel in Figure 2.

SIFT is today a well-established standard method to extract and match features from one image to features extracted from a different image covering the same scene. It is a rotation and affine invariant Harris point extractor that uses a difference-of-Gaussian function to determine scale. Harris points are in turn regions in the image where the gradients of the image are large, making them prone to stand out also in other images of the same area. For region description, SIFT uses gradient histograms in 16 subspaces around the point of interest.

In this work, SIFT is used to extract and match features from radar images. By tracking the SIFT features over a sequence of radar images, information about how the vessel is moving is obtained.

B. Measurement model

Once a feature has been matched between two scans, the position of the feature is used as a measurement to update the filter.

Since the features are matched in Cartesian image coordinates, the straightforward way would be to use the pixel coordinates themselves as a measurement. After having first converted the pixel coordinates of the landmark to coordinates in the local coordinate system, the Cartesian feature coordinates are now related to the vessel states as

$$\bar{y}_t^i = \begin{pmatrix} y_{x,t}^i \\ y_{y,t}^i \end{pmatrix} + \bar{e}_t^i = R(\psi_t) \begin{pmatrix} m_{X,t}^i - X_t \\ m_{Y,t}^i - Y_t \end{pmatrix} + \begin{pmatrix} \bar{e}_{X,t}^i \\ \bar{e}_{Y,t}^i \end{pmatrix} \quad (2)$$

where $y_{x,t}^i$ is the measured x -coordinate of feature i in the local coordinate frame at time t and $R(\psi_t)$ is the

rotation matrix between the ship orientation and the global coordinate system. (X, Y) and (m_X^i, m_Y^i) are global vessel position and global position of landmark i , respectively.

The problem with this approach is that $\bar{e}_{X,t}$ and $\bar{e}_{Y,t}$ in (2) are dependent since they are both mixtures of the range and bearing uncertainties of the radar sensor. These dependencies are also time dependent since the mixtures depend on the bearing of the radar sensor. Simply assuming them to be independent will introduce estimation errors.

A better approach is to convert the Cartesian landmark coordinates back to polar coordinates and use these as a measurement

$$y_t^i = \begin{pmatrix} r_t^i \\ \theta_t^i \end{pmatrix} = \begin{pmatrix} \sqrt{(m_{X,t}^i - X_t)^2 + (m_{Y,t}^i - Y_t)^2} \\ \arctan\left(\frac{m_{Y,t}^i - Y_t}{m_{X,t}^i - X_t}\right) - \psi_t \end{pmatrix} + \begin{pmatrix} e_{r,t}^i \\ e_{\theta,t}^i \end{pmatrix}. \quad (3)$$

This approach results in independent noise parameters $e_r \sim N(0, \sigma_r^2)$ and $e_\theta \sim N(0, \sigma_\theta^2)$, which better reflect the true range and bearing uncertainties of the range sensor.

C. Motion model

The system states describing the vessel movements at time instant t are

$$z_t = (X_t \ Y_t \ v_t \ \psi_t \ \omega_t \ \phi_t)^T \quad (4)$$

where v is the velocity, ψ is the course, ω is the angular velocity and ϕ_t is the crab angle, i.e. the wind and stream induced difference between course and velocity vector (normally small). Due to the size and the speed of the vessel, Figure 1, the crab angle is assumed to be very small throughout the experiments. The system states are more extensively described in Table 1 and are also shown in Figure 3. We will be using a coordinated turn model, though there are many possible motion models available.

When landmarks at unknown positions are tracked to estimate the movements of the vessel, these should be kept in the state vector. If the same landmarks are tracked over a sequence of radar scans, a better estimate of the vessel movement is acquired than if they are tracked between just two.

The system states are therefore expanded to also include all landmarks within the field of view to create a visual odometry framework. The new state vector becomes

$$z_t = (X_t \ Y_t \ v_t \ \psi_t \ \omega_t \ \phi_t \ m_{X,t}^k \ m_{Y,t}^k \ \dots \ m_{Y,t}^l)^T. \quad (5)$$

Only the $l - k + 1$ latest landmarks are within the field of view why only these are kept in the state vector. As the vessel travels on, the landmarks will one by one

leave the field of view why they will be removed from the state vector and subsequently replaced by new ones.

When all old landmarks are kept in the state vector even after they have left the field of view, it is a SLAM framework. If an old landmark that left the field of view long ago was rediscovered, this would allow for the whole vessel trajectory to be updated. This is called a loop closure and is one of the key features in SLAM. The SLAM state vector is therefore

$$z_t = (X_t \ Y_t \ v_t \ \psi_t \ \omega_t \ \phi_t \ m_{X,t}^1 \ m_{Y,t}^1 \dots)^T. \quad (6)$$

A discretized linearization of the coordinated turn model using the SLAM landmark augmentation gives

$$\begin{pmatrix} X_{t+\Delta t} \\ Y_{t+\Delta t} \\ v_{t+\Delta t} \\ \psi_{t+\Delta t} \\ \omega_{t+\Delta t} \\ \phi_{t+\Delta t} \\ m_{X,t+\Delta t}^1 \\ m_{Y,t+\Delta t}^1 \\ \vdots \end{pmatrix} = \begin{pmatrix} X_t + \frac{2v_t}{\omega_t} \sin\left(\frac{\omega_t \Delta t}{2}\right) \cos\left(\psi_t + \phi_t + \frac{\omega_t \Delta t}{2}\right) \\ Y_t + \frac{2v_t}{\omega_t} \sin\left(\frac{\omega_t \Delta t}{2}\right) \sin\left(\psi_t + \phi_t + \frac{\omega_t \Delta t}{2}\right) \\ v_t + v_{v,t} \\ \psi_t + \omega_t \Delta t \\ \omega_t + v_{\omega,t} \\ \phi_t + v_{\phi,t} \\ m_{X,t}^1 \\ m_{Y,t}^1 \\ \vdots \end{pmatrix} \quad (7)$$

where Δt is the difference in acquisition time between the two latest matched features. v_v , v_ω and v_ϕ are independent Gaussian process noises reflecting the movement uncertainties of the vessel.

D. Multi-rate issues

Having defined a motion model and a measurement model, state estimation is usually straightforwardly implemented using standard algorithms such as the extended Kalman filter (EKF), the unscented Kalman filter (UKF) or the particle filter (PF), see [38]. These filters iterate between a prediction step based on the motion model and a correction step based on the measurement model and the current measurement. The most natural approach is to stack all landmarks from one radar revolution into a large measurement vector and then run the filter with $\Delta t = T = 1.5$, where T is the radar revolution time (1.5 in our application). There are, however, two nonstandard problems here.

The first problem is that all matched features should not be used to update the filter at the same time. Since the vessel is moving while the radar is revolving, a shift in the radar image is introduced. A vessel speed of 10 m/s will result in a difference in landmark position of about 15 m from the beginning to the end of the scan. If the vessel is traveling with a constant velocity and course, then all relative changes in feature positions between two scans would be equal, but if the vessel is turning or accelerating, a shift in the relative position change is introduced. This results in different features

having different relative movements that will introduce estimation errors if all measurements are used at the same time. We have found the error caused by this batch approach to be quite large. Therefore, the filter should be updated by each matched feature independently.

If independent course and velocity measurements were available, they could be used to correct the skewness in the radar image. The filter estimates of velocity and course should though not be used for scan correction, since this would create a feedback loop from the estimates to the measurements that can cause filter instability.

Second, if there was one landmark detected in each scan direction, the filter could be updated with the rate $\Delta t = T/N = \frac{1.5}{2000}$ where N is the number of measurements per rotation (2,000 in our application). This is not the case though and we are facing a multi-rate problem with irregularly sampled measurements. The measurement model can now conveniently be written as

$$y_t = \begin{cases} y_t^i & \text{if landmark } i \text{ is detected at time } t \\ \text{NaN} & \text{otherwise.} \end{cases} \quad (8)$$

Now, any of the filters (EKF, UKF, PF) can be applied at rate $\Delta t = T/N$ using (8), with the understanding that a measurement being NaN simply means that the measurement update is skipped.

E. Alternative landmark free odometric framework

The landmark-based framework derived above suffers from one apparent shortcoming: the number of features will grow very fast. After only a short time period, thousands of potential landmarks will have been found, causing large overhead computations in the implementation. Either a lot of restrictions must be made on which of the new landmarks to track, or a different approach is needed.

If the map is not central, an approach based on differential landmark processing could be taken. Instead of tracking the same feature over a sequence of scans, features are only matched between two scans to compute the relative movement between the sweeps.

1) *Relative movement estimation*: As described in Section III-D, all features from the entire scan should not be used to estimate the relative movement at the same time. If the vessel is accelerating or turning, the scan will be scewed causing estimation errors. The idea is therefore to use subsets of features, measured over a short time interval, to compute the relative changes in course $\Delta\psi_t$ and position ΔX_t and ΔY_t between two scans. The scans are simply divided into multiple slices where each segment covers a time interval τ_t . The

relative movement and course estimates are therefore calculated multiple times per scan pair.

The relative change in position and course can be described as a relationship between the landmark positions measured in the local coordinate frame at time t and $t - T$. These landmark positions are related as

$$\begin{pmatrix} y_{x,t-T}^i \\ y_{y,t-T}^i \end{pmatrix} = \begin{pmatrix} \cos(\Delta\psi_t) & -\sin(\Delta\psi_t) \\ \sin(\Delta\psi_t) & \cos(\Delta\psi_t) \end{pmatrix} \begin{pmatrix} y_{x,t}^i \\ y_{y,t}^i \end{pmatrix} + \begin{pmatrix} \Delta X_t \\ \Delta Y_t \end{pmatrix} \quad (9)$$

where $y_{x,t}^i$ is the measured x - coordinate of landmark i at time instant t in the local coordinate system. $y_{x,t-T}^i$ is the measured x - coordinate in the previous scan.

If (9) was used to estimate the changes in course and position between two scans using each segment independently, quite large course changes could be experienced. Since each scan pair is used multiple times because it divided into segments, practically the same course and position change would be calculated over and over again. For example, the change in course registered between the scans will be similar for two adjacent segments. The only truly new information in the next segment are the changes experienced over that specific segment, not the changes experienced over the rest of the full scan because that has already been studied. To avoid calculating the same course change multiple times, the changes in course and position can be calculated recursively as

$$\Delta\psi_t = \Delta\psi_{t-T} + \delta\psi_t \quad (10a)$$

$$\Delta X_t = \Delta X_{t-T} + \delta X_t \quad (10b)$$

$$\Delta Y_t = \Delta Y_{t-T} + \delta Y_t. \quad (10c)$$

The change in course is subsequently divided into two parts: the estimated change in course $\Delta\psi_{t-T}$ using the previous segment, which is known, and a small change in course $\delta\psi_t$ experienced during the segment, which is unknown. Even though the vessel used for data acquisition is very maneuverable, $\delta\psi_t$ can be assumed small.

The sine and cosines of (9) can now be rewritten using (10a) and simplified using the small angle approximation $\cos(\delta\psi_t) \approx 1$ and $\sin(\delta\psi_t) \approx \delta\psi_t$

$$\begin{aligned} \cos(\Delta\psi) &= \cos(\Delta\psi_{t-T} + \delta\psi_t) \\ &= \underbrace{\cos(\Delta\psi_{t-T})}_{c_\Delta} \underbrace{\cos(\delta\psi_t)}_{\approx 1} - \underbrace{\sin(\Delta\psi_{t-T})}_{s_\Delta} \underbrace{\sin(\delta\psi_t)}_{\approx \delta\psi_t} \quad (11) \\ &\approx c_\Delta - s_\Delta \delta\psi_t \end{aligned}$$

$$\begin{aligned} \sin(\Delta\psi) &= \sin(\Delta\psi_{t-T} + \delta\psi_t) \\ &= \underbrace{\sin(\Delta\psi_{t-T})}_{s_\Delta} \underbrace{\cos(\delta\psi_t)}_{\approx 1} + \underbrace{\cos(\Delta\psi_{t-T})}_{c_\Delta} \underbrace{\sin(\delta\psi_t)}_{\approx \delta\psi_t} \quad (12) \\ &\approx s_\Delta + c_\Delta \delta\psi_t \end{aligned}$$

where c_Δ and s_Δ are known.

Dividing the change in position into two parts as in (10b) and (10c) does not change the equation system (9) in practice, why ΔX_t and ΔY_t are left as before.

The equation system becomes

$$\begin{aligned} \begin{pmatrix} y_{x,t-T}^i \\ y_{y,t-T}^i \end{pmatrix} &= \begin{pmatrix} c_\Delta - s_\Delta \delta\psi_t & -s_\Delta - c_\Delta \delta\psi_t \\ s_\Delta + c_\Delta \delta\psi_t & c_\Delta - s_\Delta \delta\psi_t \end{pmatrix} \begin{pmatrix} y_{x,t}^i \\ y_{y,t}^i \end{pmatrix} \\ &+ \begin{pmatrix} \Delta X_t \\ \Delta Y_t \end{pmatrix} \Leftrightarrow \\ \begin{pmatrix} y_{x,t-T}^i \\ y_{y,t-T}^i \end{pmatrix} &= \left(\begin{pmatrix} c_\Delta & -s_\Delta \\ s_\Delta & c_\Delta \end{pmatrix} + \delta\psi_t \begin{pmatrix} -s_\Delta & -c_\Delta \\ c_\Delta & -s_\Delta \end{pmatrix} \right) \begin{pmatrix} y_{x,t}^i \\ y_{y,t}^i \end{pmatrix} \\ &+ \begin{pmatrix} \Delta X_t \\ \Delta Y_t \end{pmatrix} \quad (13) \end{aligned}$$

which can be rewritten as

$$\begin{pmatrix} y_{x,t-T}^i - c_\Delta y_{x,t}^i + s_\Delta y_{y,t}^i \\ y_{y,t-T}^i - s_\Delta y_{x,t}^i - c_\Delta y_{y,t}^i \end{pmatrix} = \begin{pmatrix} -s_\Delta y_{x,t}^i - c_\Delta y_{y,t}^i & 1 & 0 \\ c_\Delta y_{x,t}^i - s_\Delta y_{y,t}^i & 0 & 1 \end{pmatrix} \begin{pmatrix} \delta\psi_t \\ \Delta X_t \\ \Delta Y_t \end{pmatrix} \quad (14)$$

The equation system is now approximately linear and by stacking multiple landmarks in one equation system

$$\begin{pmatrix} y_{x,t-T}^i - c_\Delta y_{x,t}^i + s_\Delta y_{y,t}^i \\ y_{y,t-T}^i - s_\Delta y_{x,t}^i - c_\Delta y_{y,t}^i \\ y_{x,t-T}^j - c_\Delta y_{x,t}^j + s_\Delta y_{y,t}^j \\ y_{y,t-T}^j - s_\Delta y_{x,t}^j - c_\Delta y_{y,t}^j \\ \vdots \end{pmatrix} = \begin{pmatrix} -s_\Delta y_{x,t}^i - c_\Delta y_{y,t}^i & 1 & 0 \\ c_\Delta y_{x,t}^i - s_\Delta y_{y,t}^i & 0 & 1 \\ -s_\Delta y_{x,t}^j - c_\Delta y_{y,t}^j & 1 & 0 \\ c_\Delta y_{x,t}^j - s_\Delta y_{y,t}^j & 0 & 1 \\ \vdots \end{pmatrix} \begin{pmatrix} \delta\psi_t \\ \Delta X_t \\ \Delta Y_t \end{pmatrix} \quad (15)$$

an overdetermined system is acquired and $\delta\psi_t$, ΔX_t and ΔY_t can be determined using a least squares solver.

This estimated change in position and course can in turn be used to calculate a velocity and an angular velocity measurement as

$$\bar{y}_t = \begin{pmatrix} \bar{v}_t \\ \bar{\omega}_t \end{pmatrix} = \begin{pmatrix} \sqrt{(\Delta X_t)^2 + (\Delta Y_t)^2} \\ \frac{\delta\psi_t}{\tau_t} \end{pmatrix} + \begin{pmatrix} \bar{e}_{v,t} \\ \bar{e}_{\omega,t} \end{pmatrix}. \quad (16)$$

The measurement noises \bar{e}_v and \bar{e}_ω are assumed to be independent Gaussian noises. This transformation that provides direct measurements of speed and course change gives what is usually referred to as odometry.

Although this approach simplifies the implementation a lot, it comes with certain drawbacks. First, the landmarks correctly associated between two images are used only pairwise, and this is sub-optimal since one loses both the averaging effects that occur when the same landmark is detected many times and also the correlation structure between landmarks. Second, assuming no cross-correlation between \bar{e}_v and \bar{e}_ω is a simplification since \bar{v}_t and $\bar{\omega}_t$ are based on ΔX_t , ΔY_t and $\delta\psi_t$ which are not calculated independently. Therefore, the measurements \bar{v}_t and $\bar{\omega}_t$ are actually dependent making the noise independence assumption incorrect. And third, in order to estimate the relative movements, the time interval

used to detect the landmarks must be informative enough to calculate $\delta\psi_t$, ΔX_t and ΔY_t , but not long enough to allow a significant scan skewedness to appear. This trade-off is vessel specific and must be balanced. By ensuring that the vessel cannot be expected to turn more than for example 10 degrees during each time interval, the small angle approximation holds.

2) *ESDF*: The simplified odometric model above can still be used for mapping if a trajectory-based filtering algorithm is used. One such framework is known in the SLAM literature as the Exactly Sparse Delayed-state Filter (ESDF) [39]. It has a state vector that consists of augmented vehicle states as

$$z_{1:t} = \begin{pmatrix} z_{1:t-1} \\ z_t \end{pmatrix}, \quad (17)$$

where the state z_t is given in (4) and $z_{1:t-1}$ are all previous poses. If no loop closures are detected, then the ESDF is simply stacking all pose estimates, but once a loop closure is detected and the relative pose between the two time instances is calculated, the ESDF allows for the whole trajectory to be updated using this new information.

Once the trajectory has been estimated, all radar scans can be mapped to world coordinates. By overlaying the scans on the estimated trajectory, a radar map is obtained. Each pixel now describes how many radar detections that have occurred in that coordinate.

IV. SIFT performance on radar images

SIFT is used to extract visual island features from the radar images. Figure 4 shows the features that are extracted from the upper right quadrant of a radar scan example. Two types of features are detected; island features and vessel-related features. The latter originate from radar disturbances caused by the vessel and the waves and are visible in the bottom left corner of Figure 4. Unfortunately, this section of the image cannot just be removed since the vessel commonly travels very close to land making island features in the area sometimes crucial for navigation.

The total number of features detected is of course depending on the number of islands in the area, but also on where these islands are situated. A large island close to the vessel will block a large section of the radar scan, resulting in few features. In these experiments, an average of 650 features was extracted per full radar scan.

A. Matching for movement estimation

The SIFT features are matched to estimate the relative movement of the vessel between multiple consecutive scans. Figure 5a,b shows examples of how well these features actually match. In Figure 5a, a dense island distribution results in a lot of matches that provide a good

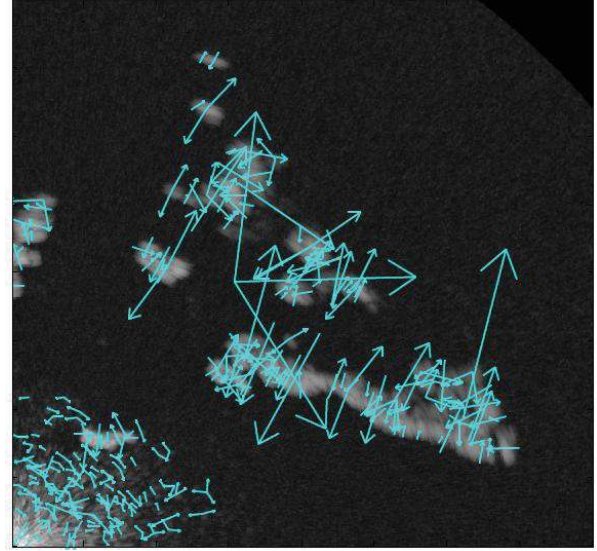


Figure 4 Radar image with extracted features marked with arrows. This is a typical example of how a radar image from which SIFT extracts a large number of features may look.

movement estimation. In Figure 5b, there are very few islands making it difficult to estimate the movements accurately.

There are two situations that can cause few matches. One is when there are few islands, and the other is

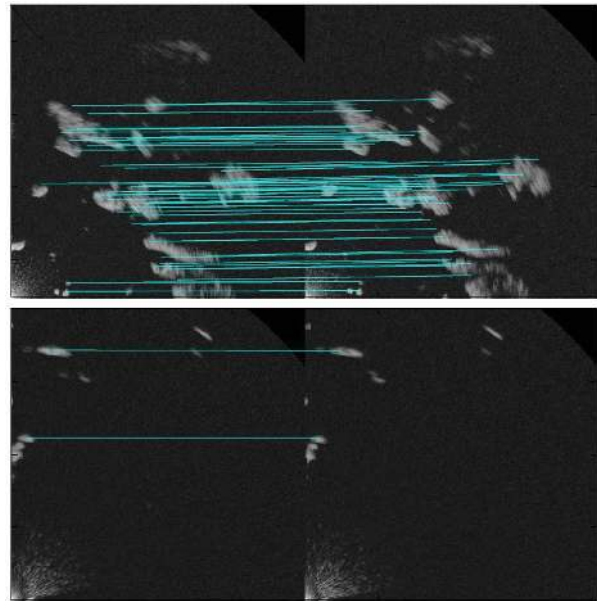


Figure 5 Two pairs of consecutive radar images with detected matches marked with lines. **a** Plenty of islands makes it easy to find good matching features. **b** Few islands result in few features to match

when a large island is very close to the vessel, blocking the view of all the other islands. When the vessel passes close to an island at high speed, the radar scans can differ quite significantly between two revolutions. This results not only in few features to match but also in features that can be significantly more difficult to match causing the relative movement estimates to degrade. On average though, about 100 features are matched in each full scan.

B. Loop closure matching

Radar features can also be used to detect loop closures that would enable the filter to update the entire vessel trajectory. The rotation invariance of the SIFT features makes radar scans acquired from different headings straightforward to match. Quite a large difference in position is also manageable due to the range of the radar sensor. This requires of course that no island is blocking or disturbing the view. Figure 6a shows example locations *a*, *b* and *c* that were used to investigate the matching performance of the visual features.

In area *a*, Figure 6b, and 6b, Figure 6c, the features are easy to match despite the rather long translational difference over open water in *b*. In both cases, a 180° difference in course is easily overcome by the visual features. This shows the strength of both the radar sensor and of the visual features. The long range of the sensor makes loop closures over a wide passage of open water

possible. These scans would be used in a full-scale SLAM experiment to update the trajectory.

In area *c*, Figure 6d, only two features are matched correctly, and there are also two false positives. If the scans are compared ocularly, it is quite challenging to find islands that are clearly matching, mostly due to blurring and to blocking islands. It is also noticeable that the radar reflections from the islands differ due to differences in radar positions which of course alters the SIFT features. The poor matching result is therefore natural in this case.

C. Feature preprocessing

Two problems remain that have not been addressed. First is the problem of false feature matches. In Figure 6b, a false feature match is clearly visible, and it would introduce estimation errors if not handled properly. An initial approach would be to use an algorithm like RANSAC[40] to remove matches that are inconsistent with all other matches. One could also use the filtering framework to produce an estimate of the probable feature position and perform a significance test on the features based on for example Mahalanobis distance. Only the features that pass this test would be allowed to update the filter estimates.

The other problem is accidental vessel matching. In heavily trafficked areas like ports, other vessels will be detected on the radar scans. If a moving vessel is

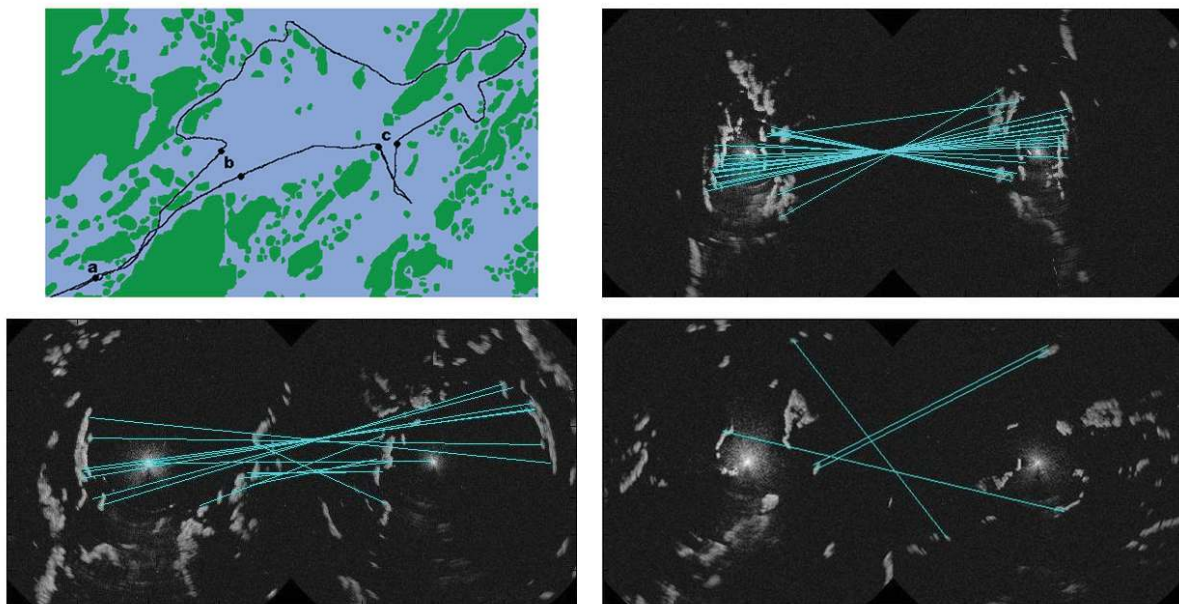


Figure 6 Examples of loop closure experiments using visual SIFT features. **a** Example locations used to examine the loop closure detection performance. **b** Multiple matched features from a loop closure in area *a* in (a). A 180° difference in course is easily handled by the SIFT features. **c** Location *b* has both a 180° course difference and a translative change but a couple of features are still matched. **d** 2 correct and 2 false matches in area *c* makes it unsuitable for loop closure updates.

deemed stationary and is used to update the filter, errors will be introduced. Two approaches could be taken to handle this problem. Again, a significance test could be used to rule out features from fast moving vessels. Alternatively, a target tracking approach could be used. By tracking the features over multiple scans, the vessels can be detected and ruled out based on their position inconsistency compared to the stationary features. Describing such a system is though beyond the scope of this paper. The joint SLAM and target tracking problem has previously been studied in [41].

V. Experimental results

The experimental results in this section come from the master thesis by Henrik Svensson [42]. The implemented framework is the one described in Sections III-E1 and III-E2.

A. Results

The trajectory used in the SLAM experiment is shown in bold in Figure 7. The track is about 3,000 s long (50 min) and covers roughly 32 km. The entire round trip was unfortunately never used in one single experiment since it was constituted of multiple data sets.

The estimated trajectory with covariance is compared to the GPS data in Figure 8. The first two quarters of the trajectory consist of an island rich environment, see Figure 6a, resulting in a very good estimate. The third quarter covers an area with fewer islands causing the performance to degrade. This results in an initial misalignment of the final quarter that makes the estimated trajectory of this segment seem worse than it actually is.

Both velocity and course estimates, Figures 9 and 10, are quite good when compared to GPS data. There is though a positive bias on the velocity estimate, probably due to the simplifications mentioned in Section III-E. The course error grows in time since the estimate is the

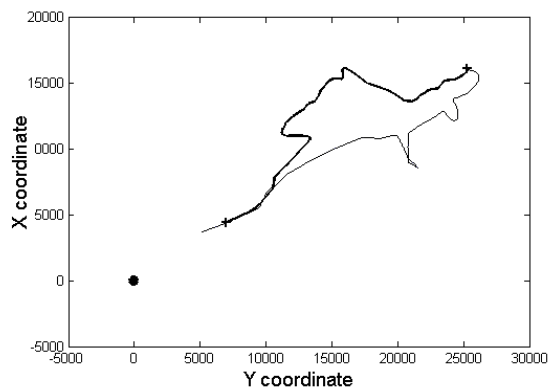


Figure 7 The whole trajectory where the studied interval is marked by bold line.

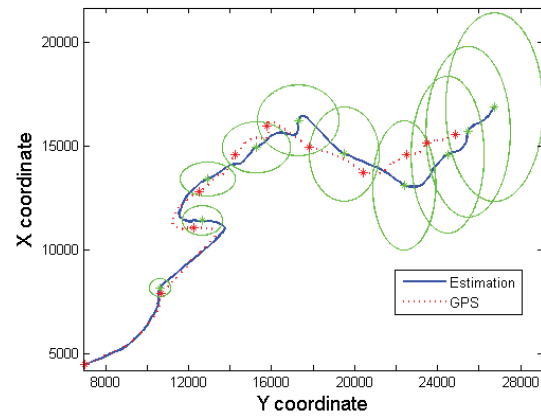


Figure 8 ESDF trajectory estimate with covariance ellipses compared to GPS measurements.

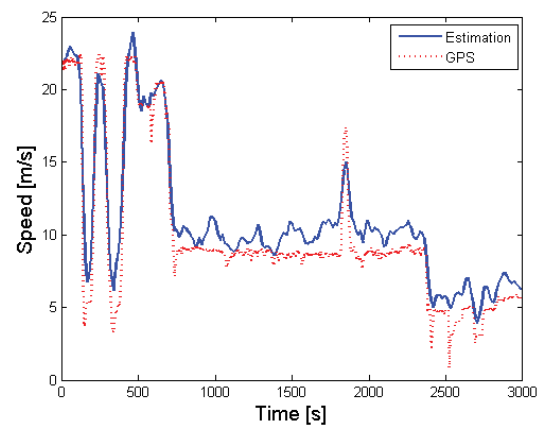


Figure 9 Velocity estimate compared to GPS measurements. A slight positive bias is present, probably due to the simplifications mentioned in Section III-E.

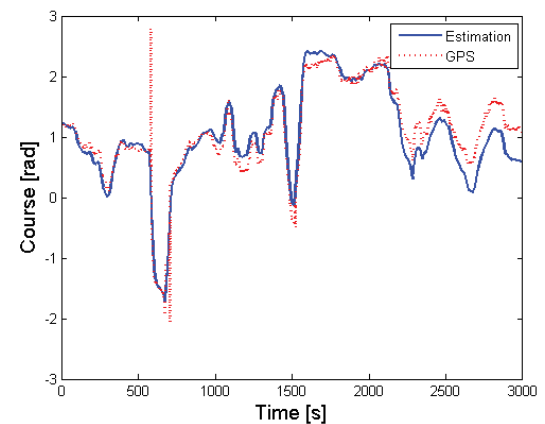


Figure 10 Course estimate compared to GPS measurements. The course estimate is the sum of all estimated changes in course, causing the error to grow in time.

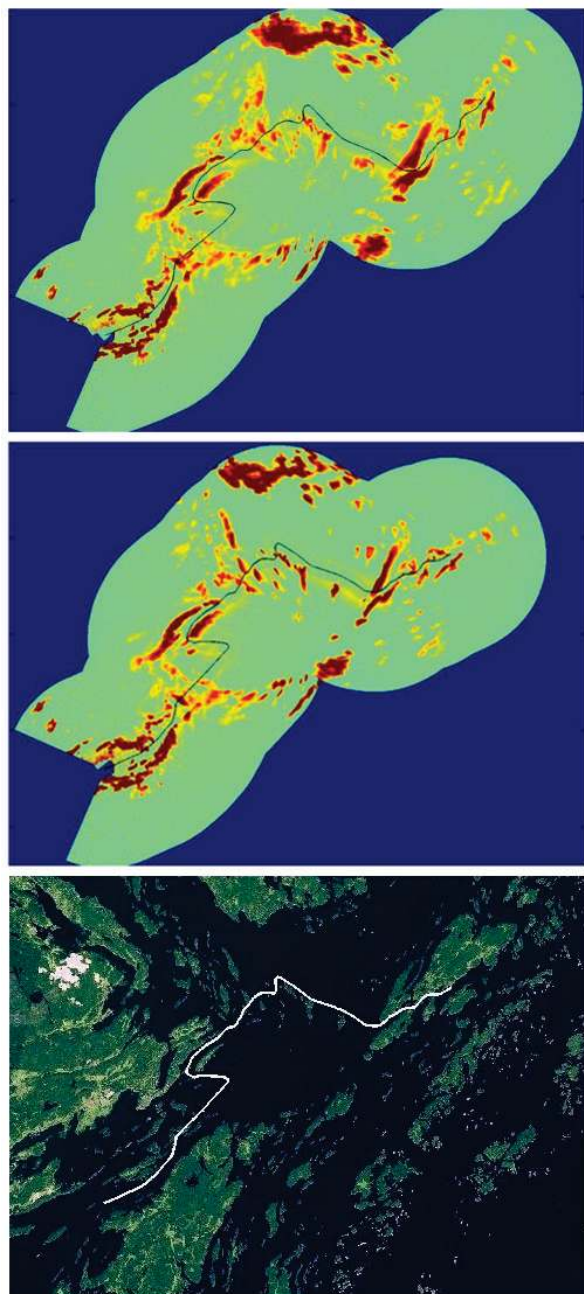


Figure 11 From top to bottom: the estimated radar map of the area, the true radar map generated using GPS and a satellite photo of the archipelago. **a** The estimated map created by overlaying the radar scans on the estimated trajectory. The estimated trajectory is marked with a black line. **b** A radar map created using the GPS measurements of the trajectory. The true trajectory is marked with a black line. **c** A satellite photo of the area. The true trajectory is marked with a white line. Many islands in (a) are easily identified in this photo.

sum of a long sequence of estimated changes in course, see (16), and there are no course measurements available. The final course error is about 30 degrees.

B. Map estimate

A radar map of the area was generated by overlaying the radar scans on the estimated trajectory. Figure 11a shows the estimated radar map that should be compared to the map created using the GPS trajectory, Figure 11b. They are very similar although small errors in the estimated trajectory are visible in Figure 11a as blurriness. Some islands appear a bit larger in the estimated map because of this. Overall, the map estimate is good.

The estimated map should also be compared to the satellite photo of the area with the true trajectory marked in white as shown in Figure 11c. When compared, many islands in the estimated map are easily identified. This shows that the rather simple approach of using visual features on radar images can provide good mapping results.

VI. Conclusions

We have presented a new approach to robust navigation for surface vessels based on radar measurements only. No infrastructure or maps are needed. The basic idea is to treat the radar scans as images and apply the SIFT algorithm for tracking landmarks of opportunity. We presented two related frameworks, one based on the SLAM idea where the trajectory is estimated jointly with a map, and the other one based on odometry. We have evaluated the SIFT performance and the odometric framework on data from a high-speed patrol boat and obtained a very accurate trajectory and map.

An interesting application of this work would be to apply this method to underwater vessels equipped with synthetic aperture sonar as the imagery sensor, since there are very few low-cost solutions for underwater navigation.

Acknowledgements

This work was supported by the Strategic Research Center MOVIII, funded by the Swedish Foundation for Strategic Research, SSF and CADICS, a Linnaeus center funded by the Swedish Research Council. The authors declare that they have no competing interests.

Author details

¹Division of Automatic Control, Linköping University Linköping, Sweden

²Nira Dynamics Linköping, Sweden ³Saab Dynamics Linköping, Sweden

Received: 10 December 2010 Accepted: 23 September 2011

Published: 23 September 2011

References

1. J Volpe, Vulnerability assessment of the transportation infrastructure relying on global positioning system final report. National Transportation Systems Center, Tech Rep (2001)
2. GNSS vulnerability and mitigation measures, (rev. 6). European Maritime Radionavigation Forum, Tech Rep (2001)
3. A Grant, P Williams, N Ward, S Basker, GPS jamming and the impact on maritime navigation. *J Navig.* **62**(2), 173–187 (2009). doi:10.1017/S0373463308005213
4. M Dalin, S Mahl, Radar Map Matching, Master's thesis, (Linköping University, Sweden, 2007)
5. R Karlsson, F Gustafsson, Bayesian surface and underwater navigation. *IEEE Trans Signal Process.* **54**(11), 4204–4213 (2006)
6. D Lowe, Object recognition from local scale-invariant features, in *ICCV '99: Proceedings of the International Conference on Computer Vision*. **2** (1999)
7. H Bay, T Tuytelaars, L Van Gool, Surf: speeded up robust features, in *9th European Conference on Computer Vision (Graz Austria, May 2006)*
8. H Durrant-Whyte, T Bailey, Simultaneous localization and mapping (SLAM): part I. *Robot. Autom Mag IEEE.* **13**(2), 99–110 (2006)
9. T Bailey, H Durrant-Whyte, Simultaneous localization and mapping (SLAM): part II. *Robot. Autom Mag IEEE.* **13**(3), 108–117 (Sept. 2006)
10. P Newman, K Ho, SLAM-loop closing with visually salient features, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (2005)
11. MC Bosse, R Zlot, Map matching and data association for large-scale two-dimensional laser scan-based SLAM. *Int J Robot Res.* **27**(6), 667–691 (2008). doi:10.1177/0278364908091366
12. K Granström, J Callmer, F Ramos, J Nieto, Learning to detect loop closure from range data, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (2009)
13. R Rouveure, M Monod, P Faure, High resolution mapping of the environment with a ground-based radar imager, in *Proceedings of the International Radar Conference* (2009)
14. P Checchin, F Grossier, C Blanc, R Chapuis, L Trassoudaine, Radar scan matching SLAM using the Fourier-Mellin transform, in *IEEE International Conference on Field and Service Robotics* (2009)
15. F Ramos, J Nieto, H Durrant-Whyte, Recognising and modelling landmarks to close loops in outdoor SLAM, in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)* (2007)
16. R Eustice, H Singh, J Leonard, M Walter, R Ballard, Visually navigating the RMS Titanic with SLAM information filters, in *Proceedings of Robotics: science and Systems* (2005)
17. I Mahon, S Williams, O Pizarro, M Johnson-Roberson, Efficient view-based SLAM using visual loop closures. *IEEE Trans Robot.* **24**(5), 1002–1014 (2008)
18. M Bryson, S Sukkarieh, Bearing-only SLAM for an airborne vehicle, in *Proceedings of the Australasian Conference on Robotics and Automation (ACRA)* (2005)
19. P Carlbom, Radar map matching. (2005) Technical Report
20. S Clark, H Durrant-Whyte, Autonomous land vehicle navigation using millimeter wave radar, in *Proceedings of the IEEE International Conference on Robotics and Automation* (1998)
21. S Clark, G Dissanayake, Simultaneous localisation and map building using millimetre wave radar to extract natural features, in *Proceedings of the IEEE International Conference on Robotics and Automation* (1999)
22. L Feng, E Milios, Robot pose estimation in unknown environments by matching 2D range scans, in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on* (1994)
23. F Lu, E Milios, Globally consistent range scan alignment for environment mapping. *Auton Robot.* **4**(4), 333–349 (1997). doi:10.1023/A:1008854305733
24. Y Chen, G Medioni, Object modeling by registration of multiple range images. *Image Vis Comput.* **10**(3), 145–155 (1992). doi:10.1016/0262-8856(92)90066-C
25. F Ramos, D Fox, H Durrant-Whyte, CRF-matching: Conditional random fields for feature based scan matching, in *Proceedings of Robotics: science and Systems* (2007)
26. Q Chen, M DeFrise, F Deconinck, Symmetric phase-only matched filtering of fourier-mellin transforms for image registration and recognition. *IEEE Trans Pattern Analysis Mach Intell.* **16**(12), 1156–1168 (1994). doi:10.1109/34.387491
27. M Chandran, P Newman, Motion estimation from map quality with millimeter wave radar, in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems* (2006)
28. SH Tsang, PS Hall, EG Hoare, NJ Clarke, Advance path measurement for automotive radar applications. *IEEE Trans Intell Transp Syst.* **7**(3), 273–281 (2006). doi:10.1109/TITS.2006.880614
29. C Lundquist, Automotive Sensor Fusion for Situation Awareness, Licentiate Thesis No 1422, (Linköping University, Sweden, 2009)
30. SM Bolen, V Chandrasekar, Methodology for aligning and comparing spaceborne radar and ground-based radar observations. *J Atmos Ocean Technol.* **20**, 647–659 (2003). doi:10.1175/1520-0426(2003)20<0.CO;2
31. S Se, D Lowe, J Little, Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *Int J Robot Res.* **21**, 735–758 (2002). doi:10.1177/027836402761412467
32. P Jensfelt, D Kragic, J Folkesson, M Bjorkman, A framework for vision based bearing only 3-D SLAM, in *Proceeding IEEE International Conference on Robotics and Automation (ICRA)* (2006)
33. M Cummins, P Newman, Probabilistic appearance based navigation and loop closing, in *IEEE International Conference on Robotics and Automation* (2007)
34. J Callmer, K Granström, J Nieto, F Ramos, Tree of words for visual loop closure detection in urban slam, in *Proceedings of the 2008 Australasian Conference on Robotics and Automation (ACRA)* (2008)
35. F Li, G Zhang, J Yan, Coregistration based on SIFT algorithm for synthetic aperture radar interferometry, in *Proceedings of ISPRS Congress* (2008)
36. M Schikora, B Romba, A framework for multiple radar and multiple 2D/3D camera fusion, in *4th German Workshop Sensor Data Fusion: trends, Solutions, Applications (SDF)* (2009)
37. H Essen, G Luedtke, P Warok, W Koch, K Wild, M Schikora, Millimeter wave radar network for foreign object detection, in *2nd International Workshop on Cognitive Information Processing (CIP)* (2010)
38. F Gustafsson, *Statistical Sensor Fusion* (Studentlitteratur, Lund, Sweden, 2010)
39. R Eustice, H Singh, J Leonard, Exactly sparse delayed-state filters for view-based SLAM. *IEEE Trans Robot.* **22**(6), 1100–1114 (2006)
40. MA Fischler, RC Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM.* **24**, 381–395 (1981). doi:10.1145/358669.358692
41. CC Wang, C Thorpe, S Thrun, M Hebert, H Durrant-Whyte, Simultaneous localization, mapping and moving object tracking. *Int J Robot Res.* **26**(9), 889–916 (2007). doi:10.1177/0278364907081229
42. H Svensson, *Simultaneous Localization and Mapping in a Marine Environment using Radar Images*, Master's thesis, (Linköping University, Sweden, 2009)

doi:10.1186/1687-6180-2011-71

Cite this article as: Callmer et al.: Radar SLAM using visual features. *EURASIP Journal on Advances in Signal Processing* 2011 **2011**:71.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com