SPATIO-TEMPORAL MATCHING OF MOVING OBJECTS IN OPTICAL AND SAR DATA

S. Hinz¹, F. Kurz², D. Weiheing¹, S. Suchandt³
¹Remote Sensing Technology, Technische Universität München, 80290 München, Germany
²Remote Sensing Technology Institute, German Aerospace Center, 82235 Wessling
{Stefan.Hinz | Diana.Weihing}@bv.tu-muenchen.de; {franz.kurz | steffen.suchandt}@dlr.de

KEY WORDS: Vehicle Detection, Vehicle Tracking, Traffic Monitoring, Traffic Parameters, Optical images, SAR images

ABSTRACT:
We present an approach for spatio-temporal co-registration of dynamic objects in Synthetic Aperture Radar (SAR) and optical imagery. Goal of this work is the performance evaluation of vehicle detection and velocity estimation from SAR images when comparing it with reference data derived from aerial image sequences. The results of evaluation show the challenges of traffic monitoring with SAR in terms of detection rates for individual vehicles.

1. INTRODUCTION

Increasing traffic has major influence on urban and suburban planning. Usually traffic models are utilized to predict traffic and forecast transportation. To derive statistical parameters of traffic for these models, data of large areas acquired at any time is desirable. Therefore, spaceborne SAR missions can be a solution for this aim. With the upcoming TerraSAR-X or RADARSAT-2 mission, SAR images up to 1 m resolution will be available. Additionally, the Dual Receive Antenna (DRA) mode enables the reception of two SAR images of the same scene within a small timeframe, which can be utilized for along-track interferometry.

In preparation of these missions, a variety of algorithms for vehicle detection and velocity estimation from SAR has been developed; see e.g. (LIVINGSTONE et al., 2002, GIERULL, 2004, MEYER et al, 2006). An extensive overview on current developments and potentials of airborne and spaceborne traffic monitoring systems is given in the compilation of (HINZ et al., 2006). It shows that civilian SAR is currently not competitive with optical images in terms of detection and false alarm rates, since the SAR image quality is negatively influenced by Speckle noise as well as layover and shadow effects in case of city areas or ragged terrain. However, in contrast to optical systems, SAR is an active and coherent sensor enabling interferometric and polarimetric analyzes making data acquisition independent from weather and illumination conditions. While the superiority of optical systems for traffic monitoring are in particular evident when illumination conditions are acceptable, SAR has the advantage of being illumination and weather independent, which makes it to an attractive alternative for data acquisition in case of natural hazards and crisis situations. Hence, validating the quality of SAR traffic data acquisition is crucial to estimate the benefits of using SAR in such situations. It is of particular importance to observe in which way fair detection results influence more generic parameters like mean velocity per road segment.

In this paper, an approach for evaluating the performance of detection and velocity estimation of vehicles in SAR images is presented, which utilizes reference traffic data derived from simultaneously acquired optical image sequences. While the underlying idea of this approach is naturally straightforward, the different sensor concepts imply a number of methodological challenges that need to be solved in order to compare the dynamics of objects in both types of imagery.
and is proportional to $\phi$ in space domain where $\lambda$ is the carrier frequency, $v_{sar}$ the satellite velocity and $v_{sar}$ the object velocity projected into the sensor’s line of sight. In other words, across-track motion leads to the fact that moving objects do not appear at their “real-world” position in the SAR image but are displaced in azimuth direction – the so-called “train-off-the-track” effect. Again, when inserting typical TerraSAR-X parameters, the displacement reaches an amount of 1.5km for a car traveling with 80km/h in across-track direction. Figure 1 shows an example of the combination of both effects. Due to across track motion a car is displaced from its real-word position on the road (green arrow in Figure 1a). In addition, the car is defocused because of along track motion when processed with a SWMF (Figure 1b). If it was filtered with the correct reference signal, the point should be sharp as in Figure 1c. Across-track motions not only influence the position of an object in the SAR image but also the interferometric phase in case of an along-track interferometric data acquisition, i.e., the acquisition of two SAR images within a short time frame with baseline $\Delta t$ aligned with the sensor trajectory. The interferometric phase is defined as the phase difference of the two co-registered SAR images $\psi = \phi - \phi'$ and is proportional to motions in line-of-sight direction. Hence, the interferometric phase can also be related to the displacement in space domain:

$$\Delta_{v_{sar}} = -R \frac{v_{sar}}{v_{sar}} \lambda [m]$$

### 2.3 Accelerations

In the majority of the literature, it is assumed that vehicles travel with constant velocity and along a straight path. If vehicle traffic on roads and highways is monitored, target acceleration is commonplace and should be considered in any processor or realistic simulation. Acceleration effects do not only appear when drivers physically accelerate or brake but also due to curved roads, since the object’s along-track and across-track motion a car is displaced from its real-world position (green arrow in Figure 1a). In addition, the car is defocused because of along track motion when processed with a SWMF (Figure 1b). If it was filtered with the correct reference signal, the point should be sharp as in Figure 1c. Across-track motions not only influence the position of an object in the SAR image but also the interferometric phase in case of an along-track interferometric data acquisition, i.e., the acquisition of two SAR images within a short time frame with baseline $\Delta t$ aligned with the sensor trajectory. The interferometric phase is defined as the phase difference of the two co-registered SAR images $\psi = \phi - \phi'$ and is proportional to motions in line-of-sight direction. Hence, the interferometric phase can also be related to the displacement in space domain:

$$\Delta_{v_{sar}} = -R \frac{v_{sar}}{v_{sar}} \lambda [m]$$

### 3.1 Geometric co-registration

Digital frame images, as used in our approach, inheres the well-known radial perspective imaging geometry that defines the mapping $[X, Y, Z] = [x_{img}, y_{img}]$ from object to image coordinates. The spatial resolution on ground ($\rho_\Phi, \rho_\Psi$, cf. Figure 2) is mainly depending on the flying height $H$, the camera optics with focal length $c$ and the size of the CCD elements ($\rho_\Phi, \rho_\Psi$). Whereas, SAR images result from time/distance measurements in range direction and parallel scanning in azimuth direction defining a mapping $[X, Y, Z] = [x_{SAR}, y_{SAR}]$. 3D object coordinates are thus mapped onto circles with radii $R_{sat}$ parallel aligned in azimuth direction $x_{SAR}$. The spatial resolutions ($\rho_\Phi, \rho_\Psi$) of range and azimuth dimension are mainly depending on the bandwidth of the range chirp and the length of the physical antenna after SAR focusing.

To accommodate for the different imaging geometries of frame imagery and SAR, we employ a Digital Elevation Model (DEM), on which both data sets are projected. Differential rectification can then be conducted by direct georeferencing of both data sets, if the exterior orientation of both sensors is precisely known. In case the exterior orientation lacks of high accuracy – which is especially commonplace for the sensor attitude – an alternative and effective approach is to transform an existing ortho-image into the approximate viewing geometry at sensor position $C$:

$$[x_{C}, y_{C}] = f(p_{ortho}, X_{ortho}, Y_{ortho}, Z_{ortho})$$

where $p_{ortho}$ is the vector of approximate transformation parameters. Refining the exterior orientation reduces then to finding the relative transformation parameters $p_{rel}$ between the given image and the transformed ortho-image, i.e.

$$[x_{img}, y_{img}] = f(p_{rel}, x_{C}, y_{C})$$

which is accomplished by matching interest points. Due to the large number of interest points, $p_{rel}$ can be determined in a robust manner in most cases. This procedure can be applied to SAR images in a very similar way – with the only modification that, now, $p_{rel}$ describe the transformation of the ortho-images into the SAR slant range geometry.

The result of geometric matching consists of accurately geocoded optical and SAR images, so that for each point in the one data set a conjugate point in the other data set can be assigned. However, geometrically conjugate points may have been imaged at different times. This is crucial for matching moving vehicles and has not been considered in the approach outlined so far.

### 3.2 Time-dependent matching

Frame cameras take snapshots of a scene at discrete time intervals with a frame rate of, e.g., 0.3 – 3Hz. Due to overlapping images, most moving objects are imaged at multiple times. SAR, in contrast, scans the scene in a quasi-continuous mode with a PRF of 1000 – 6000 Hz, i.e. each line in range direction gets a different time stamp. Due to the parallel scanning principle, a moving vehicle is imaged only once, however, as outlined above, possibly defocused and at a displaced position. Consequently, the two complementary sensor principles of SAR and optical cameras lead to the fact that the time of imaging a moving object differs for both sensors.
Figure 2 compares the two principles: It shows the overlapping area of two frame images taken at position $C_1$ at time $t_{C1}$ and position $C_2$ at $t_{C2}$, respectively. A car travelling along the sensor trajectory is thus imaged at the time-dependent object coordinates $X(t = t_{C1})$ and $X(t = t_{C2})$. On the other hand, this car is imaged by the SAR at Doppler-zero position $X(t = t_{SAR0})$, i.e. when the antenna is closest to the object. It illustrates that exact matching the car in both data sets is not possible because of the differing acquisition times. Therefore, a temporal interpolation along the trajectory is mandatory and the specific SAR imaging effects must be considered.

Temporal matching includes thus following steps:

- Reconstruction of a continuous car trajectory from the optical data by piecewise interpolation (e.g. between control points $X(t = t_{C1})$ and $X(t = t_{C2})$ in Figure 2).
- Alternatively, GIS road axes could be used if they were accurate enough.
- Calculation of a time-continuous velocity profile along the trajectory, again using piecewise interpolation.

Figure 1. a) SAR image of a highway section with displaced car due to across track motion (green arrow). b) Detail of a): Defocused car when processed with a SWMF due to along track motion. c) Same part, however, processed with a filter corresponding to the car’s along track velocity. Now the car is imaged sharply while the background gets blurred.

Figure 2: Moving objects in optical image sequence and SAR image in azimuth direction
Derivation of a maximum velocity-variance profile. The velocity variance at the control points depends purely on the imaging and measurement accuracy (see Section 4.1). To propagate the variance into the interpolated regions, we employ a simple and empirically tested dynamic model defining that the variance between control points follows a parabolic shape as exemplified in the cut-out of Figure 3. This model accommodates the fact that velocity interpolation gets less accurate with greater distance to the adjacent control points. Together with the velocity profile, it defines an uncertainty buffer, i.e. a minimum and maximum velocity for each point along the trajectory.

Transforming the trajectory into the SAR image geometry and adding the displacement due to the across track velocity component. In the same way, the uncertainty buffer is transformed.

Intersection/matching of cars detected in the SAR image with the trajectory by applying nearest neighbour matching. Cars not being matched are defined as false alarms.

As result, each car detected in the SAR data and not labeled as false alarm is assigned to a trajectory and, thereby, uniquely matched to a car found in the optical data. Figure 3 visualizes intermediate steps of matching: a given highway section (magenta line); the corresponding displacement area color coded by an iso-velocity surface; a displaced track of a smoothly decelerating car (green line); and a cut-out of the displaced uncertainty buffer. Two cars correctly detected in the SAR image are marked by red crosses in the cut-out. The local RADAR co-ordinate axes are indicated by magenta arrows.

4. ACCURACY AND VALIDATION

In order to validate the matching and estimate the accuracy, localization and velocity determination have been independently evaluated for optical and SAR imagery.

4.1 Accuracy of velocity estimation from optical images

The basic concept of determining the accuracy of vehicle measurements in optical images is the comparison of theoretically derived accuracies with empirical accuracies measured with airborne images of reference cars.

Vehicle velocity \( v_{t_2-t_1} \) derived from two consecutive co-registered or geo-coded optical images \( I_1 \) and \( I_2 \) is simply calculated by the displacement \( \Delta x \) over the time elapsed \( \Delta t \). The displacement can be calculated through the transformed coordinate differences in the object space or by the pixel differences multiplied with a scale factor \( m \) in co-registered images.

\[
v_{t_2-t_1} = \frac{\Delta x}{\Delta t} = \frac{\sqrt{(x_{t_2} - x_{t_1})^2 + (y_{t_2} - y_{t_1})^2}}{t_{t_2} - t_{t_1}} = m \frac{\sqrt{(r_{t_2} - r_{t_1})^2 + (c_{t_2} - c_{t_1})^2}}{t_{t_2} - t_{t_1}}
\]

where \( X_0 \) and \( Y_0 \) are object coordinates, \( r_0 \) and \( c_0 \) the pixel coordinates of moving cars, and \( t_i \) the acquisition times of images \( i=1,2 \).

Using factor \( m \) simplifies the calculation of theoretical accuracies, since the calculation is separated from the geocoding process. Thus, three main error sources on the accuracy of car velocity are of interest: the measurement error \( \sigma_m \) in pixel units, the scale error \( \sigma_s \) assumed to be caused mainly by DEM error \( \sigma_{DH} \) and finally the time error \( \sigma_t \) of the image acquisition time.

Figure 4 shows accuracies of vehicle velocities derived from positions in two consecutive acquired images based on calculation of error propagation. For this, different assumptions about the error sources must be made. The measurement error \( \sigma_m \) is defined as 1.0 pixel including co-registration errors, the time distance error \( \sigma_t \) as 0.02s, which corresponds to the registration frequency of the airplane navigation system, and finally a DEM error \( \sigma_{DH} \) of 10m is assumed. The simulation in Figure 4 shows decreasing accuracy at higher car velocities and shorter time distances, as the influence of the time distance error gets stronger. On the other hand, the accuracies decrease with higher flight heights as the influence of measurement errors increases. Last is converse to the effect, that with lower flight heights the influence of the DEM error gets stronger.

![Image](image-url)

Figure 3. Intermediate steps of matching: highway section (magenta line), corresponding displacement area color coded by an iso-velocity surface, a displaced track of a smoothly decelerating car (green line); and a cut-out of the displaced uncertainty buffer. Cars correctly detected in the SAR image are marked by red crosses.

Figure 4. Accuracy of vehicle velocities derived from positions in two consecutive acquired images for three time differences 0.3s, 0.7s, and 1.0s. For each time distance, four airplane heights from 1000m up to 2500m and car velocities from 0 to 80 km/h were considered.

The theoretically calculated accuracies were validated with measurements in real airborne images and with data from a reference vehicle equipped with GPS receivers. The time distance between consecutive images was 0.7s. Exact assignment of the image acquisition time to GPS track times was a prerequisite for this validation and was achieved by connecting the camera flash interface with the flight control unit. Thus, each shoot could be registered with a time error less than 0.02s. Based on onboard GPS/IMU measurements, the images were geo-coded and finally resampled to a ground pixel size of 30cm.
Figure 5 illustrates the results of the validation for one car track. The empirically derived accuracies are slightly higher than theoretical values due to inaccuracies in the GPS/IMU data processing. Yet, it also shows that the empirical standard deviation is below 5 km/h which provides a reasonable hint for defining the velocity uncertainty buffer in Section 3.2. The validation exemplifies on the other hand that vehicle accelerations cannot be derived from these image sequences with sufficient accuracy.

4.2 Accuracy of velocity measurements in SAR images

Several flight campaigns have been conducted to estimate the accuracy of velocity determination from SAR images, thereby also verifying the validity of the above derived theory. An additional goal of the flight campaigns is to simulate TerraSAR-X data for predicting the performance of the extraction procedures. To this end, an airborne Radar system has been used with a number of modifications, so that the resulting raw data is comparable with the future satellite data. During the campaign 8 controlled vehicles moved along the runway of an airfield. All vehicles were equipped with a GPS system with a 10 Hz logging frequency for measuring their position and velocity. Some small vehicles were equipped with corner reflectors to make them better visible in the image. The experiments have been flown with varying angles between the heading of the aircraft and the vehicles. The vehicles have been driven with such velocities \( v_T \) that they approximately match traffic scenarios as recorded by satellites (see Table 1).

To estimate the accuracy, the predicted image position of a moving object is derived from the object’s GPS position and its measured velocity and compared with the position measured in the image. The positions of displaced vehicles detected in the image (yellow dots in Fig 6) are compared with their true GPS-position (green dots) and the theoretical displacement computed from the GPS-velocities (red dots). As can be seen, yellow and red dots match very well, so that the theoretical background of detection and velocity estimation seems justified. Although there might be some inaccuracies included in the measurements (varying local incidence angle, GPS-time synchronization, etc.) the results show a very good match of theory and real measurements. As expected, target 7 is not visible in the image. This is due to the low processed band width (PBW) of only 1/10 of the PRF and the targets velocity. The across-track velocity of target 7 shifts the spectrum of the target outside of the PBW.

![Image](image_url)

Figure 5 True GPS positions (green) of cars, displaced positions derived from GPS velocity (red), displaced position measured in the image (yellow).

To obtain a quantitative estimate of the quality of velocity determination SAR images, the velocity corresponding to the along-track displacement in the SAR images \( v_T^{disp} \) has been compared to the GPS velocity \( v_T^{GPS} \) (see Table 1). The numerical results show that the average difference between the velocity measurements is significantly below 1 km/h. When expressing the accuracy of velocity in form of a positional uncertainty, this implies that the displacement effect influences a vehicle’s position in the SAR image only up to a few pixels depending on the respective sensor parameters, as can be seen from Figure 6.

<table>
<thead>
<tr>
<th>Target #</th>
<th>( v_T^{GPS} ) [km/h]</th>
<th>( v_T^{disp} ) [km/h]</th>
<th>( \Delta v ) [km/h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>5.22</td>
<td>5.47</td>
<td>0.25</td>
</tr>
<tr>
<td>5</td>
<td>9.24</td>
<td>9.14</td>
<td>0.1</td>
</tr>
<tr>
<td>6</td>
<td>10.03</td>
<td>9.45</td>
<td>0.58</td>
</tr>
<tr>
<td>7</td>
<td>36.92</td>
<td>not visible</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.16</td>
<td>2.33</td>
<td>0.17</td>
</tr>
<tr>
<td>9</td>
<td>4.78</td>
<td>4.86</td>
<td>0.08</td>
</tr>
<tr>
<td>10</td>
<td>3.00</td>
<td>2.01</td>
<td>0.01</td>
</tr>
<tr>
<td>11</td>
<td>6.31</td>
<td>6.28</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Table 1: Comparison of velocities from GPS and SAR

4.3 Matching results with real data

The matching approach has been tested on real data stemming from DLR’s E-SAR and 3K optical system. The flight campaign aimed at monitoring a freeway nearby Lake Chiemsee, approx. 80 km in the south-east of Munich. The freeway is heading nearly in across-track leading to large displacements of the cars in the SAR image. During the flight, also optical images of the same scene have been acquired to enable the verification of the detection results. For ensuring error-free reference data, vehicle detection and tracking has been carried out manually. Some track sections are exemplified in Figure 7.

An existing modular traffic processor has been applied to detect vehicles in the SAR data automatically, see (Suchandt et al., 2006; Wehning et al. 2007) for details. Different detectors (ATI, DPCA, likelihood ratio detector) are integrated for finding...
vehicles and can be selected individually or can be combined. Figure 8 shows an example of vehicle detection with the likelihood ratio detector (WeiHing et al. 2007). Detected vehicles are marked with red rectangles at their displaced positions. The triangles represent the positions of these vehicles when backprojected to the assigned road, whereby their color indicates the estimated velocity ranging from blue to red (0 to 170 km/h). Having these detections projected back onto the road axis, it is possible to derive parameters describing the situation on the road and feeding them into traffic simulations and traffic prediction models.

![Figure 7. Example of vehicles tracked in optical image sequence](image)

![Figure 8. Cars detected in SAR image. Displaced position of detection (rectangle), backprojection onto road (triangle), estimated velocity (color of triangle).](image)

The traffic data from the optical and the SAR system have been co-registered as described above to evaluate the performance of vehicle detection and velocity estimation. In Table 2 the traffic flow parameters derived from the detections with the likelihood ratio detector are compared to those estimated from the reference data. The vehicles moving on the upper lane from right to left are considered in this case. On the opposite lane only two vehicles have been detected which makes the derivation of reliable parameters impossible.

<table>
<thead>
<tr>
<th>Traffic parameters</th>
<th>SAR data</th>
<th>optical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean velocity</td>
<td>104 km/h</td>
<td>100 km/h</td>
</tr>
<tr>
<td>velocity range</td>
<td>29-129 km/h</td>
<td>81-135 km/h</td>
</tr>
<tr>
<td>number of vehicles</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>detection rate</td>
<td>39 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Table 2: Traffic parameters for vehicles moving on the upper lane from right to left

5. SUMMARY AND CONCLUSION

In this article, an approach for spatio-temporal co-registration of dynamic objects in SAR and optical imagery has been presented. It was used to evaluate the performance of vehicle detection and velocity estimation from SAR images compared to reference data derived from aerial image sequences. The evaluation shows the challenges of traffic monitoring with SAR in terms of detection rate. However, the traffic flow parameters derived from these results show a good correspondence with the reference data, even for a low detection rate. Hence, traffic models can make use of such data to simulate and predict traffic or to even verify certain parameters of models.

REFERENCES


