Photogrammetrie • Fernerkundung • Geoinformation 5/2007, S. 309–325, 22 figs., 2 tabs.

Evaluation of Traffic Monitoring based on Spatio-Temporal Co-Registration of SAR Data and Optical Image Sequences*

STEFAN HINZ, München, FRANZ KURZ, Oberpfaffenhofen, DIANA WEIHING, München, STEFFEN SUCHANDT, Oberpfaffenhofen, FRANZ MEYER, Fairbanks, USA, & RICHARD BAMLER, München/Oberpfaffenhofen

Keywords: Synthetic Aperture Radar, aerial images, co-registration, moving objects, traffic monitoring

Summary: This article presents an approach for spatio-temporal co-registration of dynamic objects in Synthetic Aperture Radar (SAR) and optical imagery. Background of this work is the performance evaluation of vehicle detection and velocity estimation from SAR images by comparison with reference data derived from aerial image sequences. In terms of detection rates for individual vehicles, the results of evaluation show the challenges of traffic monitoring with SAR. However, it is also shown that general traffic *flow* parameters can be derived with high quality.

Zusammenfassung: *Bewertung der Verkehrsüberwachung mittels raumzeitlicher Ko-Registrierung von SAR-Daten und optischen Luftbildsequenzen.* Dieser Artikel erläutert einen Ansatz zur raumzeitlichen Ko-Registrierung von Synthetik Apertur Radar (SAR) Daten und optischen Luftbildern für Bewegtobjekte. Hintergrund dieser Arbeit ist die Qualitätsbewertung der Fahrzeugerkennung und Geschwindigkeitsschätzung mittels SAR Daten auf Grundlage von luftbildgestützten Referenzdaten. Die Ergebnisse zeigen, dass gute Erkennungsraten von Einzelfahrzeugen in SAR Daten immer noch eine große Herausforderung darstellen. Die Bestimmung von daraus abgeleiteten Verkehrs*fluss*parametern gelingt jedoch mit erstaunlich hoher Qualität.

1 Motivation

Traffic monitoring is a very important task in today's traffic control and flow management. The acquisition of traffic data in almost real-time is essential to swiftly react to current situations. Stationary data collectors such as induction loops and video cameras mounted on bridges or traffic lights are matured methods. However, they only deliver local data and are not able to observe the traffic situation in a large road network. Hence, traffic monitoring approaches relying on airborne and spaceborne remote sensing come into play. Especially spaceborne sensors do cover very large areas. Systems like IKONOS and Quickbird as well as the Synthetic Aperture Radar (SAR) sensors mounted on TerraSAR-X and RADARSAT-2 (late to be launched in 2007) deliver images in the (sub-)meter range. Thanks to this high resolution, significant steps forward towards spaceborne traffic data acquisition are currently made (see, e. g., LIVINGSTONE et al. 2002, CHIU & LIVINGSTONE 2005, BETHKE et al. 2006, MEYER et al. 2006, LEITLOFF et al. 2006). Because of their relatively short acquisition time and their long revisit period, such systems mainly contribute to the periodic collection of statistical traffic data to validate and improve traffic models. On the other

* A preliminary, condensed version of this article has been presented at the PIA07 workshop, 19.–21. 9. 07, Munich.

1432-8364/07/2007/0309 $ 4.25
© 2007 DGPF / E. Schweizerbart'sche Verlagsbuchhandlung, D-70176 Stuttgart

hand, flexible systems are desired when traffic monitoring on demand is necessary. Airborne optical and SAR sensors or future HALE (High Altitude Long Endurance) systems show great potential to meet the demands of flexibility and mobility. While they have the capability of covering large areas, they can deliver both temporally and spatially densely sampled data.

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 as well as layover and shadow effects in case of city areas or rugged 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 is 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 parameters of higher abstraction 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 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.

2 Motion Detection in Airborne Imagery

2.1 *Vehicle Velocities from Optical Image Sequences*

The general methodology to derive vehicle velocities from airborne image sequences is sketched in Fig. 1. After image acquisition, the images are co-registered and geo-referenced. This process is commonly supported by simultaneously recorded navigation data of an INS-/GPS-System. Usually, GIS road data is available, e. g. stemming from NAVTEQ or ATKIS data bases. These data are mapped onto the geo-referenced images and approximate regions of interest (RoI) are selected. Thus, the search area for the following vehicle detection can be reduced significantly. It is in addition helpful for further processing to extract the road as well as their lanes, since geo-referencing might not be accurate enough and GIS data rarely includes the position of individual lanes.

Within the RoI, a car detection algorithm is applied that is supposed to deliver positions and optionally additional attributes such as boundary, direction etc. of the cars. The detection approaches mostly incorporate some kind of image differencing or optical flow to better identify image regions where movements might have appeared. Such procedures, however, limit the applicability of an approach to strictly moving ob-

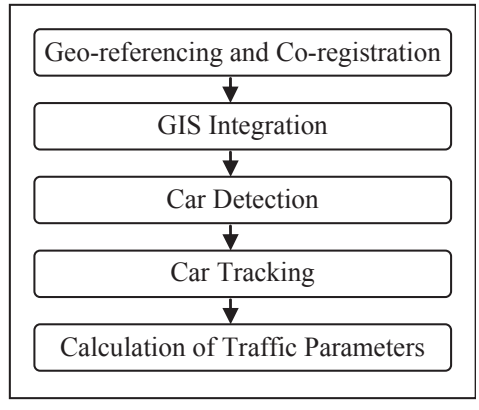

Fig. 1: Work flow of vehicle detection and tracking.

jects. After detection, the cars are tracked through matching them over the following images. In the simplest way, matching is performed sequentially between consecutive images, while advanced approaches include multiple images to make the matching more robust and to bridge temporal occlusions. The tracking procedures usually inhere general "smoothness" constraints regarding vehicle shape, vehicle radiometry, and vehicle trajectory, i. e., it is assumed that vehicles move smoothly and, despite of changing viewing and illumination geometry, their imaged geometry and radiometry do not vary very much from one image to the other. Typical approaches based on terrestrial and airborne images are found, for instance, in (Dreschler & Nagel 1982, Tan et al. 1998, Haag & Nagel 1999, Ernst et al. 2005, Kang et al. 2005, Yu et al. 2006, Kirchhof & Stilla 2006, Toth & Grejner-Brzezinska 2006).

From the results of car tracking various traffic parameters can be calculated, most importantly individual vehicle speeds and accelerations, vehicle density per road segment, and traffic flow, which is traffic density multiplied by average speed, eventually yielding the number of cars passing a given point on the road network within a certain time interval. Fig. 2 illustrates a tracking example over image triplets taken from (Lenhart & Hinz 2007). White rectangles indicate detected cars in the first image. Correctly tracked cars are marked green while incorrect track results are marked red. Black rectangles indicate cars which were correctly matched in the second image but moved out of the field-of-view of the third image. Blue marked vehicles are correctly matched in the second image but could not be tracked in the third image even though they were present. In this triplet, 16 out of 20 cars could be tracked correctly. One car moved out of sight in the third image, therefore the comparison with the third image failed. One car was incorrectly tracked. Two cars could not be found in the third image although they were present, one of those was at least correctly found in the second image.

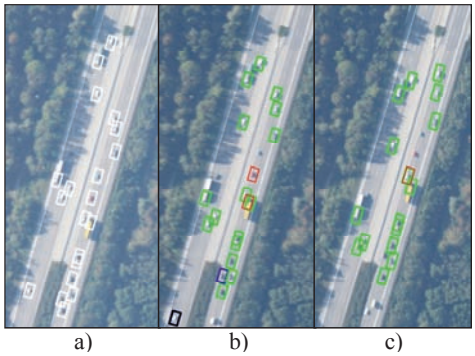

Fig. 2: Tracking of vehicles over image triplets.

2.2 *Vehicle Velocities from Dual-Channel SAR data*

In contrast to optical cameras, RADAR is an active sensor that emits frequency modulated signals – so-called chirps – with a predefined "pulse repetition rate" (PRF) in a side-looking, oblique imaging geometry and records the echoes scattered at the objects on the ground; see Fig. 3 for illustration of the RADAR imaging geometry. The echoed chirps are correlated with the sent (reference) chirp eventually yielding a compressed pulse-shaped signal whose width is mainly determined by the chirp's band width (see Fig. 4). The travelling time of the signals is proportional to the distance to the

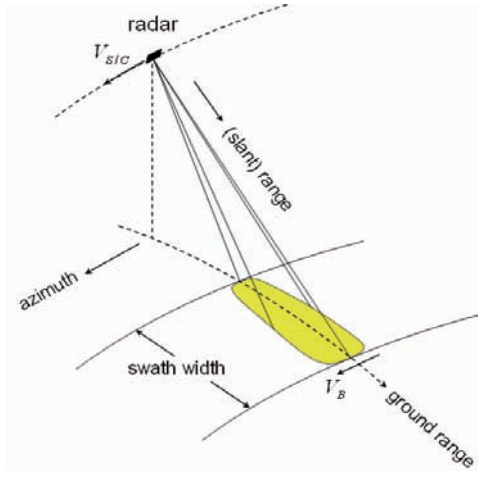

Fig. 3: RADAR imaging geometry.

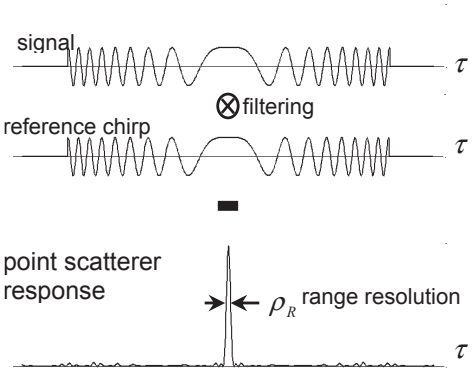

Fig. 4: Compression of sent chirp into pulse.

objects and defines the image dimension perpendicular to the flight direction, the so-called *range* or *across-track* co-ordinates. The second dimension, *azimuth* or *along-track*, is simply aligned with the flight direction. While the resolution in range direction ρ_R is determined by the chirp band width and is typically in the (sub-)meter area, the resolution in azimuth direction of the raw data depends on the antenna's real aperture characteristics (antenna length L, carrier wavelength λ, and range R) and is impractically coarse for geospatial applications. Hence, to enhance the azimuth resolution, the well-known Synthetic Aperture Radar (SAR) principle is applied, i. e., the motion of the real antenna is used to construct very long synthetic antenna by exploiting each point scatterer's range history recorded during a point's entire observation period. Since the length of the synthetic aperture increases proportional with the flying height, the resolution in azimuth direction ρ_{SA} is purely depending on the length of the physical antenna given a sufficiently large PRF to avoid aliasing. Fig. 5 compares real aperture and formation of synthetic aperture for a typical spaceborne SAR configuration.

To identify and quantify movements of objects on the ground, a thorough mathematical analysis of this so-called SAR focusing process is necessary:

The position of a Radar transmitter on board a satellite is given by $P_{sat}(t) = [x_{sat}(t); y_{sat}(t); z_{sat}(t)]$ with x being the along-track direction, y the across-track ground range direction and z being the vertical (see Fig. 6). A point scatterer is assumed to be at position $P(t) = [x(t); y(t); z(t)]$, and the range to an arbitrarily moving and accelerating point target from the radar platform is defined by $R(t) = P_{sat}(t) - P(t)$. Omitting pulse envelope, amplitude, and antenna pattern for simplicity reasons, and approximating the range history $R(t)$ by an parabola, the measured echo signal $u_{stat}(t)$ of this stationary point scatterer can be written as $u_{stat}(t) =$

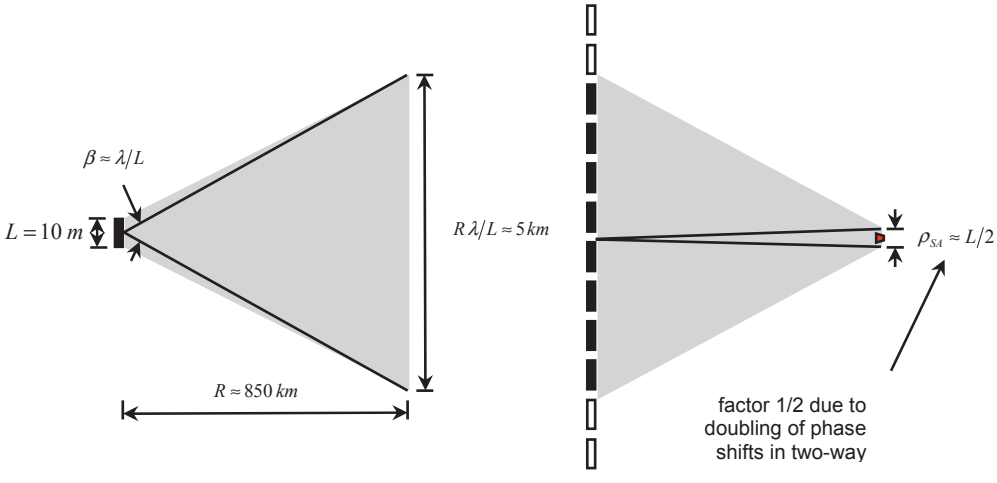

Fig. 5: Left – Example for resolution of real aperture. Right – Formation of synthetic aperture by exploiting antenna motion.

$\exp\{j\pi FM\, t^2\}$ with FM being the frequency modulation rate of the azimuth chirp:

$$FM = -\frac{2}{\lambda}\frac{d^2}{dt^2}R(t) = -\frac{2}{\lambda R}v_{sat}\, v_B$$

and v_{sat} and v_B being the platform velocity and the beam velocity on ground, respectively. Azimuth focussing of the SAR image is performed using the matched filter concept (Bamler & Schättler 1993, Cumming & Wong, 2005). According to this concept the filter must correspond to $s(t) = \exp\{-j\pi FM\, t^2\}$.

An optimally focused image is obtained by complex-valued correlation of $u_{stat}(t)$ and $s(t)$. To construct $s(t)$ correctly, the actual range or phase history of each target in the image must be known, which can be inferred from sensor and scatterer position. Usually, a time dependence of the scatterer position is ignored yielding $P(t) = P$. This concept is commonly referred to as *stationary-world matched filter* (SWMF). Because of this definition, a SWMF does not correctly represent the phase history of a significantly moving object.

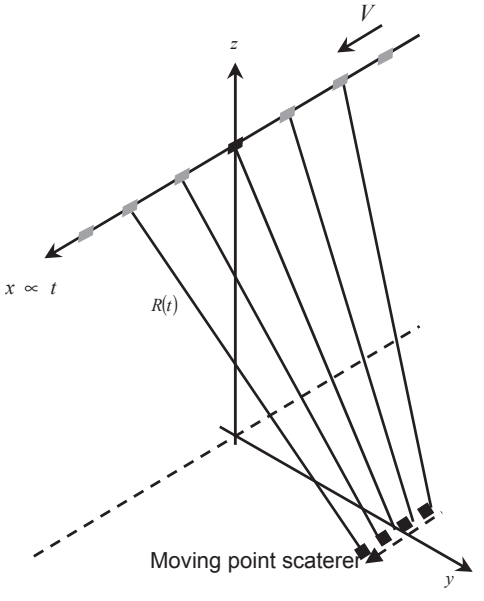

Fig. 6: Geometry of along track moving object while formation of synthetic aperture.

To quantify the impact of a significantly moving object we first assume the point to move with velocity v_{x0} in azimuth direction (along-track, see Fig. 6). The relative velocity of sensor and scatterer is different for the moving object and the surrounding stationary world. Thus, along track motion changes the frequency modulation rate FM of the received scatterer response. Focusing the signal with a SWMF consequently results in a blurring of the signal. It is unfortunately not possible to express the amount of defocusing exactly in closed form.

Yet, when considering the stationary phase approximation of the Fourier-Transform, the width Δt of the focused peak can be approximated by

$$\Delta t \approx 2T_A \frac{v_{x0}}{v_B}[s]$$

with T_A being the synthetic aperture time.

As can be seen, the amount of defocusing depends strongly on the sensor parameters. A car traveling with 80 km/h, for instance, will be blurred by approx. 30 m when inserting TerraSAR-X parameters. However, it has to be kept in mind that this approximation only holds if $v_{x0} \gg 0$. It is furthermore of interest, to which extent blurring causes a reduction of the amplitude h at position $t = 0$ (the position of the signal peak) depending on the point's along-track velocity. This can be calculated by integrating the signal spectrum and making again use of the stationary phase approximation:

$$h(t = 0, v_{x0}) \approx \frac{B}{T_A}\frac{v_B}{v_{sat}}$$

with B being the azimuth bandwidth.

When a point scatterer moves with velocity v_{y0} in across-track direction (see Fig. 7), this movement causes a change of the point's range history proportional to the projection of the motion vector into the line-of-sight direction of the sensor $v_{los} = v_{y0}\sin(\theta)$ with θ being the local elevation angle. In case of constant motion during illumination the change of range history is linear and causes an additional linear phase trend in the echo signal.

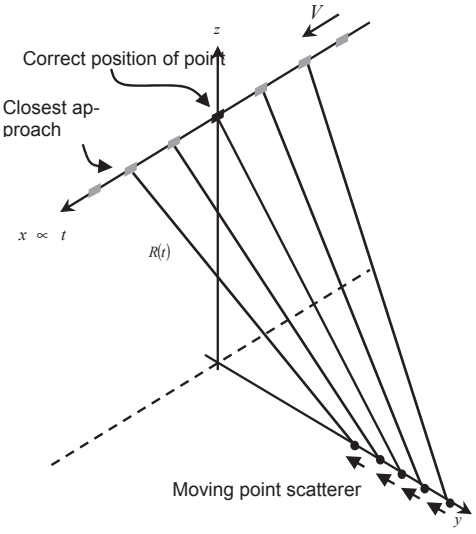

Fig. 7: Geometry of across track moving object while formation of synthetic aperture.

Correlating such a signal with a SWMF results in a focused point that is shifted in azimuth direction by

$$t_{shift} = \frac{2\, v_{los}}{\lambda \cdot FM}\,[s]$$

in time domain, respectively by

$$\Delta_{az} = -\,R\,\frac{v_{los}}{v_{sat}}\,[m]$$

in space domain.

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 1.5 km for a car traveling with 80 km/h in across-track direction. Fig. 8 shows an example of the combination

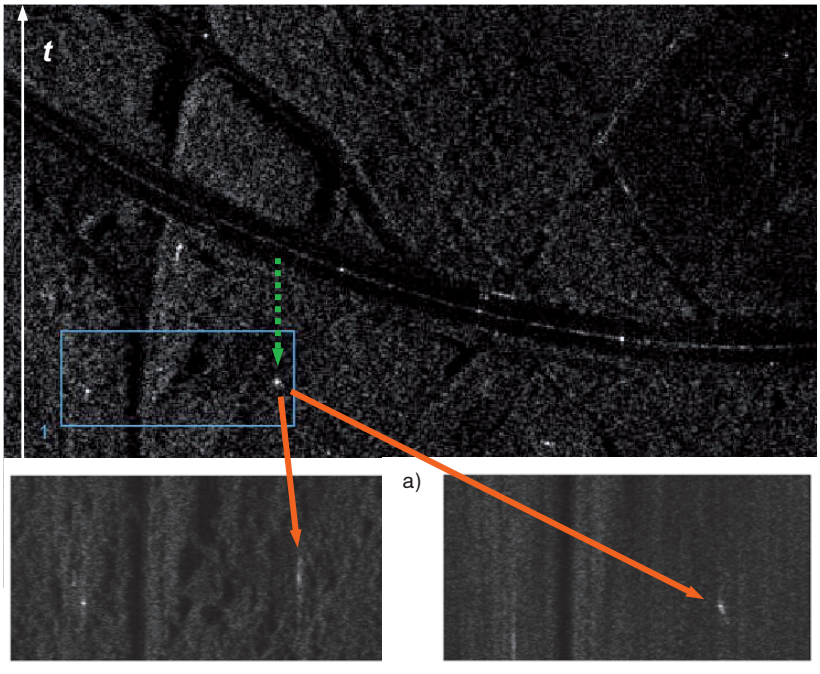

Fig. 8: a) SAR image of a highway section with displaced car due to across track motion (green arrow). b) Detail: Defocused car when processed with a SWMF. c) Same part processed with filter corresponding to the car's along track velocity. Now the car is imaged sharply while the background gets blurred.

of both effects. Due to across track motion a car is displaced from its real-word position on the road (green arrow in Fig. 8a). In addition, the car is defocused because of the along track motion when processed with a SWMF (Fig. 8b). If it was filtered with the correct reference signal, the point should be sharp as shown in Fig. 8c.

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 Δl aligned with the sensor trajectory. The interferometric phase is defined as the phase difference of the two co-registered SAR images $\psi = \varphi_1 - \varphi_2$ and is proportional to motions in line-of-sight direction. Hence, the interferometric phase can also be related to the displacement in space domain:

$$\Delta_{az} = -R\frac{v_{los}}{v_{sat}} = -R\psi\frac{\lambda}{4\pi\cdot\Delta l}\,[m]$$

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 velocity components vary on a curved trajectory during the Radar illumination. The effects caused by along-track or across-track acceleration have recently been studied in (Sharma et al. 2006, Meyer et al. 2006). Summarizing, along-track acceleration a_x results in an asymmetry of the focused point spread function, which leads to a small azimuth-displacement of the scatterer after focusing, whose influence can often be neglected. However, the acceleration in across-track direction a_y causes a spreading of the signal energy in time or space domain. The amount of this defocusing is significant and comparable with that caused by along-track motion. We refer the interested reader to (Meyer et al. 2006) where an in-depth study about all the above mentioned influences in TerraSAR-X data can be found.

The effects of moving objects hinder the detection of cars in conventionally processed SAR images. On the other hand, these effects are mainly deterministic and thus can be exploited to not only detect vehicles but also measure their velocity. Typical schemes for detection and velocity estimation of moving objects rely on (a) the interferometric phase and (b) the FM-rate variation for detecting the across-track and along-track velocities (Gierull 2004, Bethke et al. 2006). To make the estimation more robust (Meyer et al. 2006, Hinz et al. 2007) include also GIS data from road databases as a priori knowledge.

a) Detection using interferometric phase:

For classifying stationary and moving points in interferometric data, e. g. through a constant false alarm rate detection scheme, the probability density distributions of vehicles and background need to be known. For all stationary targets the interferometric phase values are assumed to be statistically distributed around the expectation value $E\{\psi\} = 0$. Using the underlying assumption of jointly Gaussian-distributed data in the two images, the joint probability density function (pdf) $f_c\{\eta, \psi\}$ of amplitude η and phase ψ of an interferogram is given by:

$$f_c(\eta,\psi) = \frac{2\,n^{n+1}\eta^n}{\pi(1-|\rho|^2)\cdot\Gamma(n)} \cdot K_{n-1}\left(\frac{2\,n\eta}{1-|\rho|^2}\right) \cdot \exp\left(\frac{2\,n\eta\,|\rho|\cos(\psi)}{1-|\rho|^2}\right)$$

where ρ is the coherence, n the number of looks (effectively the amount of averaging), $\Gamma(\cdot)$ is the gamma function and $K_n(\cdot)$ is the modified Bessel function of the n^{th} order. In case of public traffic, it can in addition be assumed that vehicles travel on a known

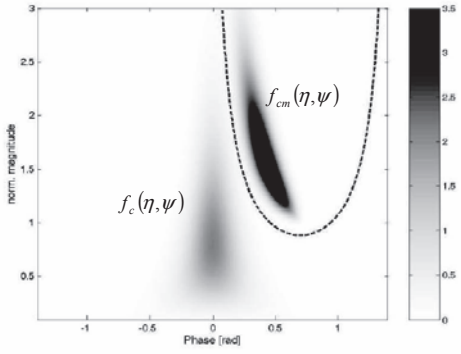

Fig. 9: Pdf for clutter $f_c(\eta, \psi)$, moving object $f_{cm}(\eta, \psi)$ and corresponding curve of separation.

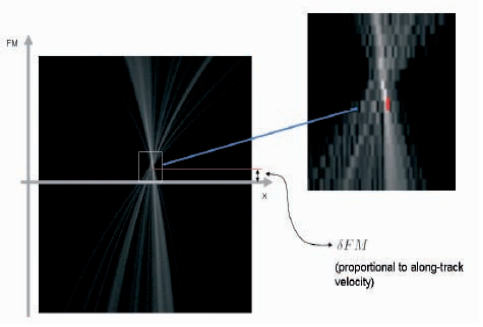

Fig. 10: Azimuth line *x* of moving object processed with varying FM rates and detected peak (right: detail). *δFM* corresponds to along-track velocity (w/o acceleration).

road network so that a similar pdf as above can be derived depending on the vehicle backscatter and velocity induced phase (Gierull 2001, 2002). An analytic pdf $f_{cm}\{\eta, \psi\}$ for vehicles has not been found yet. However an approximation valid for $n \gg 1$ is given by:

$$f_{cm}(\eta,\psi) = \frac{2\,n^{n+1}\,\eta\,[(\eta-\delta\cos(\psi-\vartheta))^2+\delta^2\sin(\psi-\vartheta)^2]^{\frac{n-1}{2}}}{\pi\,(1-|\rho|^2)\cdot\Gamma(n)}$$

$$\cdot\,K_{n-1}\left(\frac{2\,n\sqrt{(\eta-\delta\cos(\psi-\vartheta))^2+\delta^2\sin(\psi-\vartheta)^2}}{1-|\rho|^2}\right)$$

$$\cdot\,\exp\left(\frac{2\,n\,\rho\,(\eta\cos\psi-\delta\cos\vartheta)}{1-|\rho|^2}\right)$$

with $\delta = \beta/\eta$ while the moving target's signal is assumed to have a peak amplitude β.

Using this approximation as an alternative hypothesis, $f_c\{\eta, \psi\}$ and $f_{cm}\{\eta, \psi\}$ allow to defining a likelihood ratio to which a threshold can be applied. Fig. 9 shows an example and a corresponding curve of separation.

b) Detection using FM rate variation

The scheme outlined so far can only be applied if displacement or interferometric phase occurs at all. This does not happen for objects moving purely in along-track direction. A common strategy to estimate a vehicle's velocity in along-track direction relies on hypothesizing a series of FM rates corresponding to different vehicle velocities and analyzing each pixel's "sharpness function" over these FM rates (see Fig. 10). That particular FM rate, which produces the sharpest peak, corresponds to the correct estimate of the along-track velocity. Different schemes for analyzing the sharpness function have been proposed. They include feature based methods (Weihing et al. 2006, Hinz et al. 2007) as well as time-frequency analysis (Gierull 2004, Bethke et al. 2006).

Typical results for the detection of cars are depicted in Fig. 11. The red boxes show the detected (displaced) vehicles and the arrows indicate their corresponding "real-world" position on the road. Colors correspond to the vehicle velocity derived from displacement and orientation of the road.

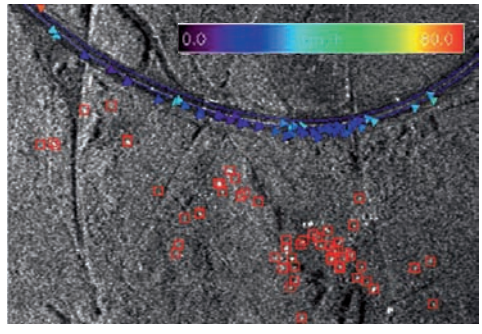

Fig. 11: Typical detection result of vehicles is SAR data.

3 Matching moving vehicles in optical and SAR data

To evaluate SAR based vehicle detection and velocity estimation, it is reasonable to make use of simultaneously acquired optical image sequences. As described above, these images allow to deriving vehicle velocities and, when choosing an appropriate focal length, they can also cover the same part of a scene as SAR images. In addition, optical images are rather easy to interpret for a human operator so that reliable reference data of moving objects can be achieved. Yet matching dynamic objects in SAR and optical data remains challenging since the two data sets do not only differ in geometric properties but also in temporal aspects of imaging. Hence, our approach for matching vehicles consists of a geometric part (Section 3.1) and a time-dependent part (Section 3.2).

3.1 *Geometric matching*

Digital frame images, as used in our approach, inhere the well-known radial perspective imaging geometry that defines the mapping $[X, Y, Z] \Rightarrow [x_{img}, y_{img}]$ from object to image co-ordinates. As sketched in Fig. 12, the spatial resolution on ground (ρ_X, ρ_Y) is mainly depending on the flying height H, the camera optics with focal length c and the size of the CCD elements (ρ_x, ρ_y). On the other side, the geometry of SAR results from time/distance measurements in range

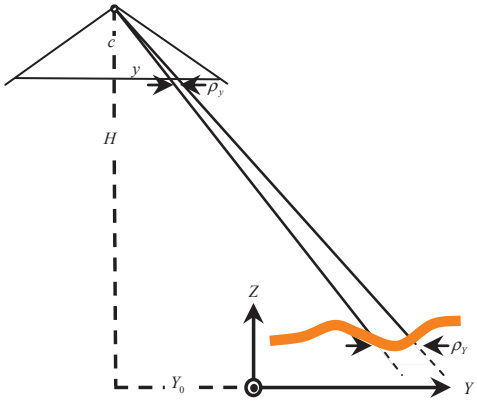

Fig. 12: Imaging geometry of optical images.

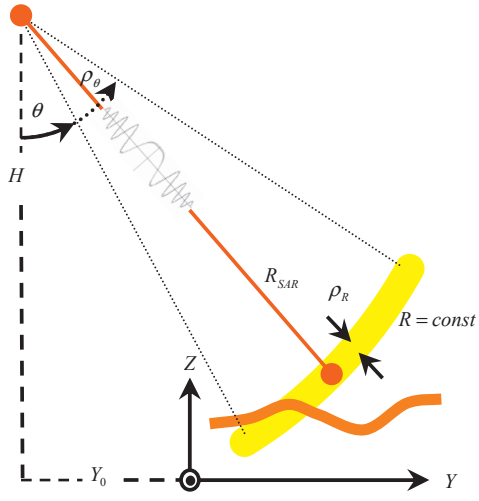

Fig. 13: Imaging geometry of SAR in range direction.

direction and parallel scanning in azimuth direction defining a mapping $[X, Y, Z] \Rightarrow [x_{SAR}, R_{SAR}]$. 3D object co-ordinates are thus mapped onto circles with radii R_{SAR} parallel aligned in azimuth direction x_{SAR} (see Fig. 13). As mentioned above, after SAR focusing, the spatial resolutions (ρ_R, ρ_{SA}) of range and azimuth dimension are mainly depending on the bandwidth of the range chirp and the length of the physical antenna. Please note that the resolution ρ_θ of the viewing angle θ is usually too coarse to derive the 3rd dimension so that SAR remains a 2D imaging system.

The different imaging geometries of frame imagery and SAR require the incorporation of differential rectification to assure highly accurate mapping of one data set onto the other. To this end, we employ a Digital Elevation Model (DEM), on which both data sets are projected*. Direct georeferencing the data sets is straightforward, if the exterior orientation of both sensors is known precisely. In case the exterior orientation lacks of high accuracy – which is especially commonplace for the sensor attitude – an alternative and effective approach is to

* We use an external DEM; though, it could be derived directly from the frame images.

transform an existing ortho-image into the approximate viewing geometry at sensor position C:

$$[x_C, y_C] = f(\boldsymbol{p}_{ortho}, X_{ortho}, Y_{ortho}, Z_{ortho})$$

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

$$[x_{img}, y_{img}] = f(\boldsymbol{p}_{rel}, x_C, y_C),$$

which is accomplished by matching interest points. Due to the large number of interest points, $\boldsymbol{p}_{rel}$ 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, $\boldsymbol{p}_{ortho}$ describe the transformation of the ortho-images into the SAR slant range geometry. The result of geometric matching consists of accurately geo-coded 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 *Temporal matching*

The different sensor principles of SAR and optical cameras lead to the fact that the time of imaging a moving object would differ for both sensors – even in the theoretical case of exactly coinciding trajectories of the SAR antenna's phase center and the camera's projection center. Frame cameras take snapshots of a scene at discrete time intervals with a frame rate of, e. g., 0.3–3 Hz. 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.

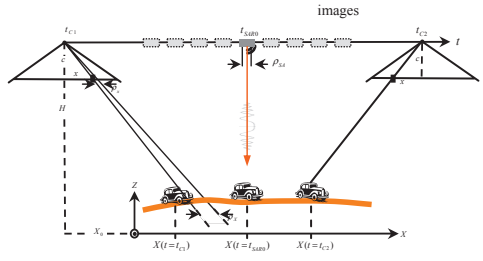

Fig. 14: Imaging moving objects in optical image sequences.

Fig. 14 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-depending object co-ordinates $X(t = t_{C1})$ 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. Fig. 14 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. Hence, our strategy for matching includes following steps:

a) 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 Fig. 14). Fig. 15 depicts an example of reconstructing the trajectory using Hermite splines. Alternatively, GIS road axes could be used if they were accurate enough.
b) Calculation of a time-continuous velocity profile along the trajectory, again using piecewise interpolation.
c) 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

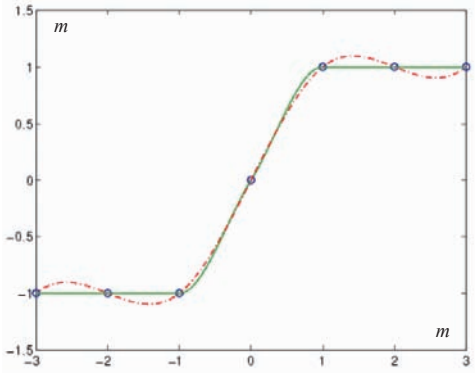

Fig. 15: Spatial interpolation between control points (circles): Cubic splines (dashed) vs. Hermite splines.

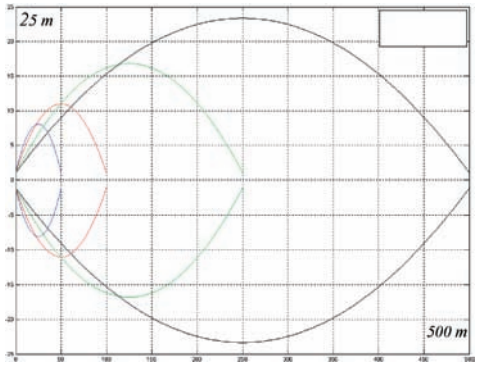

Fig. 16: Examples for evolution of velocity variance depending on distance between control points (on horizontal axis).

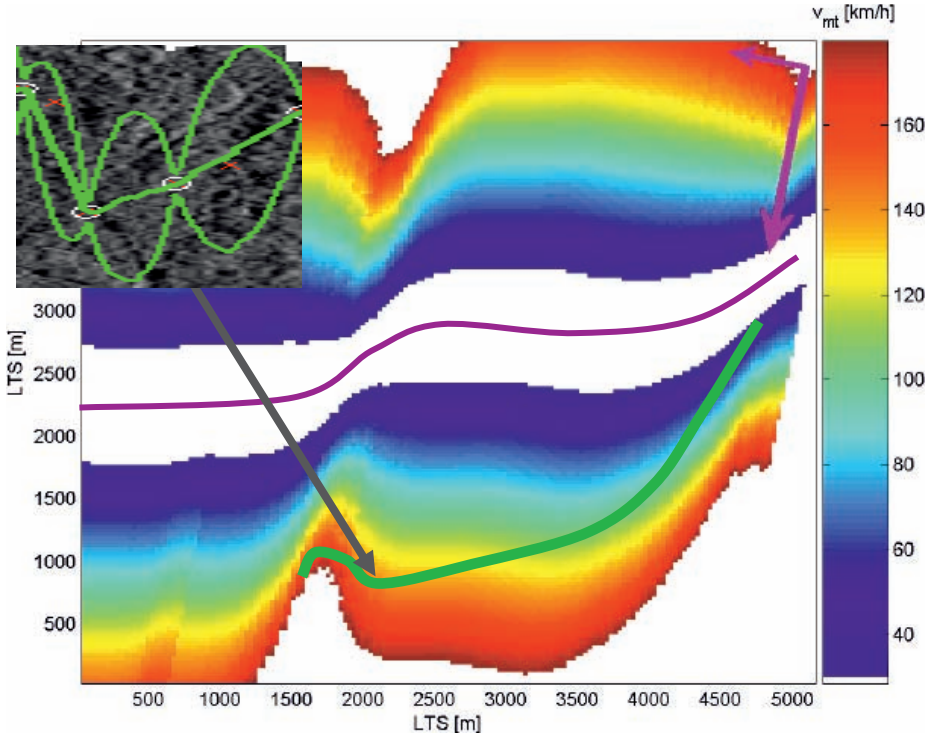

Fig. 17: Matching: highway section (magenta line), corresponding displacement area (color coded iso-velocity surface), displaced track of a decelerating car (green line), local RADAR coordinate system (magenta arrows). Cut-out shows detail of uncertainty buffer. Cars correctly detected in the SAR image are marked by red crosses.

a simple and empirically tested dynamic model defining that the variance between control points follows a parabolic shape as exemplified in Fig. 16. 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.

d) 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.
e) 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. Fig. 17 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 Aspects, Validation and Results

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 vehicle measurements in 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 $v_{I2\text{-}1}$ derived from two consecutive co-registered or geo-coded optical images *I1* and *I2* is simply calculated by the displacement Δs over the time elapsed Δt. Two ways of calculating the displacement are possible, first based on the transformed coordinate differences in the object space in geo-coded images and second based on the pixel differences multiplied with a scale factor m in co-registered images.

$$v_{I2\text{-}1} = \frac{\Delta s}{\Delta t} = \frac{\sqrt{(X_{I2} - X_{I1})^2 + (Y_{I2} - Y_{I1})^2}}{t_{I2} - t_{I1}} = m \frac{\sqrt{(r_{I2} - r_{I1})^2 + (c_{I2} - c_{I1})^2}}{t_{I2} - t_{I1}}$$

where X_{Ii} and Y_{Ii} are object coordinates, r_{Ii} and c_{Ii} the pixel coordinates of moving cars, and t_{Ii} the acquisition times of images $i = 1, 2$.

The advantage of the second way is the separation of the image geo-coding process (represented by factor m) from the process of car measurements, which simplifies the calculation of theoretical accuracies. Thus, three main error sources on the accuracy of car velocity are spotted: the measurement error σ_P in pixel units, the scale error σ_m assumed to be caused mainly by DEM error σ_H, and finally the time error σ_{dt} of the image acquisition time.

Fig. 18 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 σ_P is defined as *1.0* pixel including co-registration errors, the time distance error σ_{dt} as *0.02 s*, which corresponds to the registration frequency of the airplane navigation system, and finally a DEM error σ_H of 10 m is assumed. The simulation in Fig. 18 shows decreasing accu-

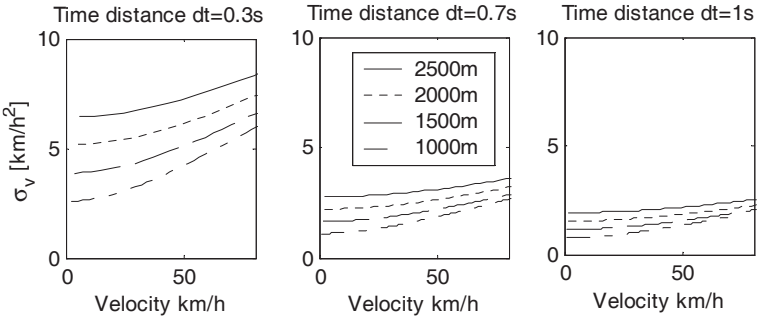

Fig. 18: Accuracy of vehicle velocities derived from positions in two consecutive acquired images for three time differences 0.3 s, 0.7 s, and 1.0 s. For each time distance, four airplane heights from 1000 m up to 2500 m and car velocities from 0 to 80 km/h were considered.

racy 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.

Given three consecutive images, the vehicle accelerations $a_{I3\text{-}2\text{-}1}$ can be calculated by the quotient of measured vehicle velocity difference and the time distance. For equidistant acquisition times, the time difference is half the time distance between image 3 and 1.

$$a_{I3\text{-}2\text{-}1} = \frac{v_{I3\text{-}2} - v_{I2\text{-}1}}{0.5(t_{I3} - t_{I1})}$$

The simulation was performed based on three accelerations $1.0\,m/s^2$, $2.0\,m/s^2$, and $3.0\,m/s^2$ and car velocities from 0 to 60 km/h. The rest of the parameters are the same as before. In general, the accuracy level of acceleration measurements is unsatisfying, as the standard deviation is over $1.0\,m/s^2$ for nearly all configurations.

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 0.7 s. 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 0.02 s. Based on onboard GPS/IMU measurements, the images were geo-coded and finally resampled to a ground pixel size of 30 cm. Fig. 19 illustrates the results of the validation for one car track. The empirically derived accuracies are slightly worse 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.

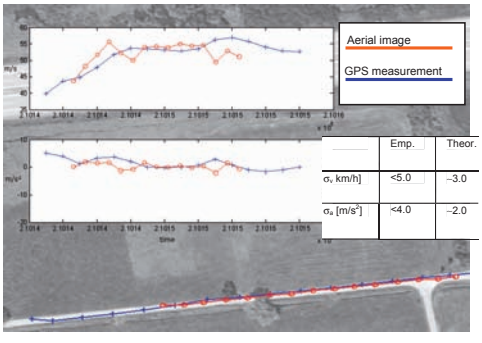

Fig. 19: Vehicle positions (projected tracks), vehicle velocities (top figure), and accelerations (bottom figure) derived from airborne images and GPS measurements. Empirically measured and theoretically calculated accuracies are listed in the table.

4.2 *Accuracy of vehicle 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 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_{Tn} that their appearance in the (airborne) SAR data approximately matches traffic scenarios as recorded by satellites (see Tab. 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. 20) 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.

Tab. 1: Comparison of velocities from GPS and SAR.

Target #	v_{Tn}^{GPS} [km/h]	v_{Tn}^{disp} [km/h]	Δv [km/h]
4	5.22	5.47	0.25
5	9.24	9.14	0.1
6	10.03	9.45	0.58
8	2.16	2.33	0.17
9	4.78	4.86	0.08
10	3.00	2.01	0.01
11	6.31	6.28	0.03

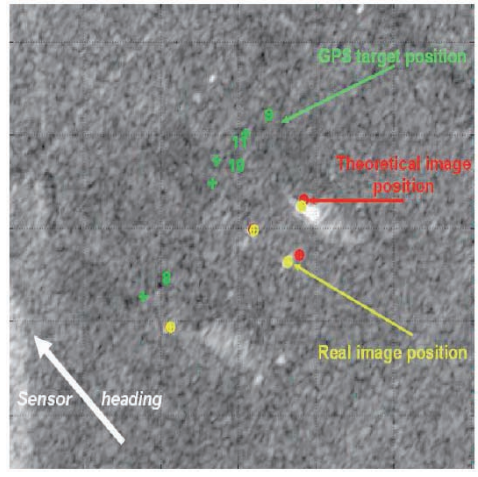

Fig. 20: True GPS positions (green) of cars, displaced positions derived from GPS velocity (red), displaced position measured in the image (yellow).

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_{Tn}^{disp} has been compared to the GPS velocity v_{Tn}^{GPS} (see Tab. 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 Fig. 20.

4.3 *Results from flight campaign*

The 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 Fig. 21.

To detect vehicles in the SAR data, an existing modular traffic processor has been applied, which processes SAR data specifically with the objective of moving vehicle detection, see (Suchandt et al. 2006, Weihing 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. Fig. 22 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 triangels represent the positions of these vehicles when back-projected to the assigned road. These triangles are color-coded regarding their 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.

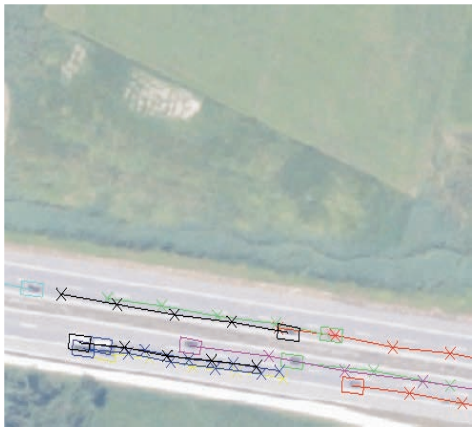

Fig. 21: Example of vehicles tracked in optical image sequence.

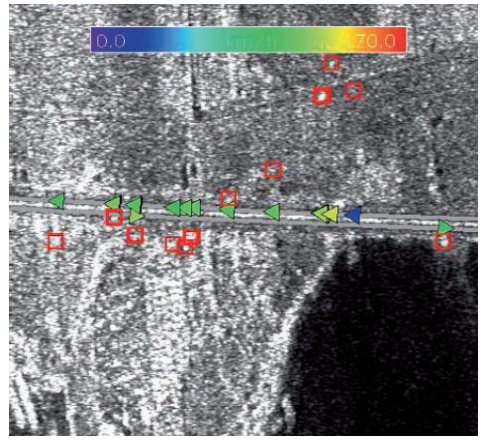

Fig. 22: Cars detected in SAR image. Displaced position of detection (rectangle), backprojection onto road (triangle), estimated velocity (color of triangle).

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 Tab. 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

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

Traffic parameters	SAR data	optical data
mean velocity	104 km/h	100 km/h
velocity range	29–129 km/h	81–135 km/h
number of vehicles	12	31
detection rate	39 %	100 %

only two vehicles have been detected which makes the derivation of reliable parameters impossible.

It can be seen from Tab. 2 that the detection rate is quite fair, as expected from other studies (e. g., MEYER et al. 2006). However, the results obtained for more generic traffic parameters are very encouraging, e. g. when comparing the values of the estimated mean of velocity, a good correspondence can be seen. Hence, even for a lower percent of detections in the SAR data, reliable parameters can be extracted. As has been shown in (SUCHANDT et al. 2006) one can derive, for instance, drive-through times for a road section from these data with high accuracy. Such information is highly useful for near-realtime traffic management since it allows to advising the drivers in choosing the best route.

Possible false detections, like the vehicle marked blue (estimated velocity is 29 km/h), could be eliminated, if considering additional information. In case of free traffic such a low velocity is very unlikely so that it could be rejected by an outlier test based on the velocity distribution in a certain neighborhood.

4 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

BAMLER, R. & SCHÄTTLER, B., 1993: SAR Data Acquisition and Image Formation. – In: SCHREIER, G. (Ed.): Geocoding: ERS-1 SAR Data and Systems. – Wichmann-Verlag.

BETHKE, K.-H., BAUMGARTNER, S., GABELE, M., HOUNAM, D., KEMPTNER, E., KLEMENT, D., KRIEGER, G. & ERXLEBEN, R., 2006: Air- and spaceborne monitoring of road traffic using SAR moving target indication. – Project TRAMRAD. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4): 243–259.

CHIU, S. & LIVINGSTONE, C., 2005: A comparison of displaced phase centre antenna and along-track interferometry techniques for RADARSAT-2 ground moving target indication. – Canadian Journal of Remote Sensing **31** (1): 37–51.

CUMMING, I. & WONG, F., 2005: Digital Processing of Synthetic Aperture Radar Data. – Artech House, Boston.

DRESCHLER, L. & NAGEL, H.-H., 1982: Volumetric model and trajectory of a moving car derived from monocular TV frame sequence of a street scene. – Computer Graphics and Image Processing **20:** 199–228.

ERNST, I., HETSCHER, M., THIESSENHUSEN, K., RUHÉ, M., BÖRNER, A. & ZUEV, S., 2005: New approaches for real time traffic data acquisition with airborne systems. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36 (3/W24): 69–73.

GIERULL, C., 2001: Statistics of SAR interferograms with application to moving target detection – Technical Report DREO-TR-2001–045, Defence R&D Canada.

GIERULL, C., 2002: Moving Target Detection with Along-Track SAR Interferometry: A Theoretical Analysis. – Technical Report DRDC-OTTAWA-TR-2002–084, Defence R&D Canada.

GIERULL, C., 2004: Statistical Analysis of Multilook SAR Interferograms for CFAR Detection of Ground Moving Targets – IEEE Transactions on Geoscience and Remote Sensing **42:** 691–701.

HAAG, M. & NAGEL, H.-H., 1999: Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Sequences. – International Journal of Computer Vision **35** (3): 295–319.

HINZ, S., BAMLER, R. & STILLA, U., 2006: Theme issue ''Airborne and spaceborne traffic monitoring''. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4).

Hinz, S., Meyer, F., Eineder, M. & Bamler, R., 2007: Traffic monitoring with spaceborne SAR – Theory, simulations, and experiments. – Computer Vision and Image Understanding, Vol. **106** (2/3): 231–244.

Kang, J., Cohen, I., Medioni, G. & Yuan, C., 2005: Detection and Tracking of Moving Objects from a Moving Platform in Presence of Strong Parallax. – International Conference on Computer Vision, Vol. I: 10–17.

Kirchhof, M. & Stilla, U., 2006: Detection of moving objects in airborne thermal videos. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4): 187–196.

Leitloff, J., Hinz, S. & Stilla, U., 2006: Detection of Vehicle Queues in Quickbird Images of City Areas. – Photogrammetrie – Fernerkundung – Geoinformation, **4** (6): 315–325.

Lenhart, D. & Hinz, S., 2006: Automatic Vehicle Tracking in Low Frame Rate Aerial Image Sequences. – International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, Vol. 36-3: 203–208.

Livingstone, C.-E., Sikaneta, I., Gierull, C., Chiu, S., Beaudoin, A., Campbell, J., Beaudoin, J., Gong, S. & Knight, T.-A, 2002: An Airborne Sythentic Apertur Radar (SAR) Experiment to Support RADARSAT-2 Ground Moving Target Indication (GMTI). – Canadian Journal of Remote Sensing **28** (6): 794–813.

Meyer, F., Hinz, S., Laika, A., Weihing, D. & Bamler, R., 2006: Performance Analysis of the TerraSAR-X Traffic Monitoring Concept – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4): 225–242.

Sharma, J., Gierull, C. & Collins, M., 2006: Compensating the effects of target acceleration in dual-channel SAR-GMTI. – IEE Radar, Sonar, and Navigation **153** (1): 53–62.

Suchandt, S., Eineder, M., Mueller, R., Laika, A., Hinz, S., Meyer, F. & Palubinskas, G., 2006: Development of a GMTI Processing System for the Extraction of Traffic Information from TerraSAR-X Data. – Proceedings of European Conference on Synthetic Aperture Radar: on CD.

Tan, T., Sullivan, G. & Baker, K., 1998: Model-Based Localisation and Recognition of Road Vehicles. – International Journal of Computer Vision **27** (1): 5–25.

Toth, C.K., & Grejner-Brzezinska, D., 2006: Extracting dynamic spatial data from airborne imaging sensors to support traffic flow estimation. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4): 137–148.

Weihing, D., Hinz, S., Meyer, F., Laika, A. & Bamler, R., 2006: Detection of Along-Track Ground Moving Targets in High Resolution Spaceborne SAR Images. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 36-7: on CD.

Weihing, D., Hinz, S., Meyer, F., Suchandt, S. & Bamler, R., 2007: An Integral Detection Scheme for Moving Object Indication in Dual-Channel High Resolution Spaceborne SAR Images. – Proceedings of IEEE-ISPRS Workshop URBAN 2007, Paris, France, on CD.

Yu, Q., Cohen, I., Medioni, G. & Wu, B., 2006: Boosted Markov Chain Monte Carlo Data Association for Multiple Target Detection and Tracking. – International Conference on Pattern Recognition, Vol. II: 675–678.

Addresses of the authors:

Dr.-Ing. Stefan Hinz, Dipl.-Ing. Diana Weihing, Prof. Dr.-Ing. Richard Bamler, Technische Universität München, Lehrstuhl für Methodik der Fernerkundung, D-80333 München, Tel.: 089-289 22671, Fax: 089-280 9573, e-mail: mail@RemoteSensing-TUM.de

Dr.-Ing. Franz Kurz, Dipl.-Ing. Steffen Suchandt, Prof. Dr.-Ing. Richard Bamler, Deutsches Zentrum für Luft- und Raumfahrt, Institut für Methodik der Fernerkundung, D-82230 Wessling/Oberpfaffenhofen, Tel.: 08153-282672, Fax: 08153-281446, URL: www.caf.dlr.de

Dr.-Ing. Franz Meyer, University of Alaska Fairbanks, Alaska Satellite Facility, Fairbanks, AK 99775, Tel.: +1-907-474 7211, e-mail: fmeyer@asf.alaska.edu

Manuskript eingereicht: April 2007
Angenommen: Juni 2007