

Towards automatic near real-time traffic monitoring with an airborne wide angle camera system

Dominik Rosenbaum · Franz Kurz ·
Ulrike Thomas · Sahil Suri ·
Peter Reinartz

Received: 19 September 2008 / Accepted: 25 November 2008 / Published online: 7 December 2008
© European Conference of Transport Research Institutes (ECTRI) 2008

Abstract

Purpose Large area traffic monitoring with high spatial and temporal resolution is a challenge that cannot be served by today available static infrastructure. Therefore, we present an automatic near real-time traffic monitoring approach using data of an airborne digital camera system with a frame rate of up to 3 fps.

Methods By performing direct georeferencing on the obtained aerial images with the use of GPS/IMU data we are able to conduct near real-time traffic data extraction. The traffic processor consists mainly of three steps which are road extraction supported by a priori knowledge of road axes obtained from a road database, vehicle detection by edge extraction, and vehicle tracking based on normalized cross correlation.

Results Traffic data is obtained with a correctness of up to 79% at a completeness of 68%.

Conclusions With this system we are able to perform area-wide traffic monitoring with high actuality independent from any stationed infrastructure which makes the system well suited for deployments on demand in case of disasters and mass events.

Keywords Traffic monitoring · Vehicle detection · Tracking

1 Introduction

A society that relies on individual mobility day to day requires sufficient methods for traffic monitoring and guidance. Especially daily commuters want to know travel times for their way to work. Moreover, relief forces are interested in precise travel times for their routing in case of emergencies, mass events, and disasters. However, precise travel time prediction on road networks is one of the most important concerns and challenges in modern transportation and traffic sciences. In order to determine traffic flow on different road types automatically, several approaches are possible. In general, traffic monitoring is mainly based on data from conventional stationary ground measurement systems such as inductive loops, radar sensors or terrestrial cameras. All ground measurement systems embedded in road infrastructure deliver precise traffic data punctually with high temporal resolution, but their spatial distribution is still limited to selected motorways and main roads. The low spatial resolution of these systems makes area-wide traffic monitoring difficult. New approaches collect data by means of mobile measurement units which flow with the traffic. The so called floating car data (FCD, [4, 17]) obtained from taxicabs can deliver useful traffic information within cities, but they are only available in few big cities today. Furthermore, the traffic information available from this source depends on the routes taxicabs drive, but taxi drivers tend to avoid busy roads during rush hours. Hence, only few or no data will be available on roads burdened with commuter traffic. In order to contribute to area-wide traffic monitoring by remote sensing, several projects, based on airborne optical and SAR sensors as well

D. Rosenbaum (✉) · F. Kurz · U. Thomas · S. Suri · P. Reinartz
Remote Sensing Technology Institute,
German Aerospace Center (DLR),
P.O. Box 1116, 82230 Weßling, Germany
e-mail: dominik.rosenbaum@dlr.de

F. Kurz
e-mail: franz.kurz@dlr.de

U. Thomas
e-mail: ulrike.thomas@dlr.de

S. Suri
e-mail: sahil.suri@dlr.de

P. Reinartz
e-mail: peter.reinartz@dlr.de

as SAR satellite sensors are currently running at DLR or have already been concluded. In Reinartz et al. [16] the general suitability of image time series from airborne cameras for traffic monitoring was shown. Tests with several camera systems and various airborne platforms, as well as the development of an airborne traffic monitoring system and thematic image processing software for traffic parameters were performed within the projects “LUMOS” and “Eye in the Sky” [3, 8].

One of the actual projects is called “ARGOS” (*A*irborne wide area *h*igh altitude *m*onitoring *S*ystem). It aims on traffic monitoring in case of mass events and disasters. It is intended to support security authorities and organisations as well as rescue forces during these occasions. Collected traffic data will be provided to the relief forces via a traffic portal called “DELPHI” (e.g. [1]). Within the ARGOS project we are currently developing a system that will be able to deliver area-wide traffic data in near real-time by using airborne remote sensing technologies. It is mainly based on our newly developed 3 head digital frame sensor system, namely the “3K camera”. This sensor is capable of wide-angle imagery at a high repetition rate (up to 3 fps). The big advantage of the remote sensing techniques presented here is that the measurements can be applied nearly everywhere (exception: tunnel segments) and there are no dependencies on any third party infrastructure. Restrictions due to clouds and fog are overcome by using airborne SAR data, which will be implemented in the ARGOS project in future. First results on traffic monitoring based on remote sensing SAR systems have been already shown in e.g. Bethke et al. [2], or Suchandt et al. [19].

Up to now, there was also a restriction that optical data were not used during nights, but in our approach we show the capability of optical camera data to monitor traffic during nights.

Airborne imagery provides a high spatial resolution combined with acceptable temporal resolution depending on the flight repetition rate. However, automatic traffic monitoring from airborne optical imagery requires complex image analysis methods and traffic models. Moreover, estimates for travel times through the area of aerial surveillance can directly be determined from extracted traffic parameters [11]. Although this prototype airborne traffic monitoring system is still deployed on demand during disasters and mass events, future continuous missions for traffic monitoring in congested urban areas may be possible based on future carriers like unmanned aerial vehicles (UAVs) or high altitude long endurance (HALE) aircrafts.

The publication is arranged as follows: Section 2 gives an overview of the sensor system and the obtained testing data, while Section 3 describes the developed algorithms for traffic monitoring in detail. In Section 4 the results from testing the algorithms are presented. Section 5 demonstrates

the night shot capabilities of the system and Section 6 gives conclusions in brief.

2 System and database

The near real-time monitoring system consists of two parts. One part is installed onboard the aircraft, consisting of the 3K camera system, a real-time GPS/IMU unit, one PC for each single camera processing image data, one PC for traffic monitoring tasks, a downlink-antenna with a band width of 30 Mbit/s automatically tracking the ground station, and a PC for steering the antenna. The ground station mainly consists of a parabolic receiving antenna, which is automatically aligned with the antenna at the aircraft, and a PC system for visualization of the down-linked images and traffic data. Given an internet access at the place of the ground station, the obtained traffic data will be directly transferred to the DELPHI traffic portal.

2.1 The 3K-camera

The 3K-camera system (3K: “3Kopf”=3 head) consists of three non-metric off-the-shelf cameras (Canon EOS 1Ds Mark II, 16 Mpix). The cameras are arranged in a fixture unit with one camera looking in nadir direction and two in oblique sideward direction (Fig. 1), which leads to an increased FOV of max 110°/31° in across track/flight direction. The camera system is coupled to a GPS/IMU navigation system, which enables the direct georeferencing of the 3K optical images. Boresight angle calibration of the system is done on-the-fly without ground control points based on automatically matched three-ray tie points in combination with GPS/IMU data [12].

Figure 2 illustrates the image acquisition geometry of the DLR 3K-camera system. Based on the use of 50 mm Canon lenses, the relation between airplane flight height, ground



Fig. 1 DLR 3K-camera system consisting of three Canon EOS 1Ds Mark II, integrated in a ZEISS aerial camera mount, and an IMU (red box)

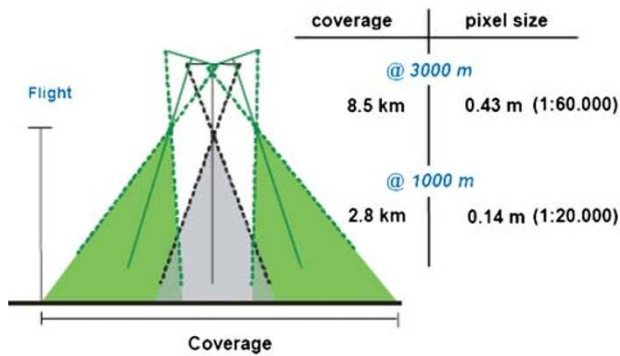


Fig. 2 Illustration of the image acquisition geometry. The tilt angle of the sideward looking cameras is approx. 35°

coverage, and pixel size is shown, e.g. the ground sampling distance (GSD) at a flight height of 1,000 m is 15 cm in nadir (20 cm in side-look) and the image array covers up 2.8 km in width.

2.2 The onboard system

For processing images acquired by the 3K-camera system in real time we are currently developing a distributed image processing system consisting of five PCs that will be on board of the plane. Each of the three cameras is connected via firewire to one PC. These PCs will be responsible for image acquisition, for orthorectification of images in real time (direct georeferencing) and for street segmentation. The fourth PC will perform vehicle detection and vehicle tracking. The fifth PC mosaikes images and sends them down via an S-Band microwave link. Thus, many image processing modules run concurrently on several PCs. Within the project ARGOS a new middleware called DANAOS¹ (*Distributed middleWare for a Near real-time mOnitoring System*) has been recently developed at DLR. In order to organize the real time modules this middleware is running on each PC. DANAOS handles inter-process communication over the network, provides name services, and synchronizes access to shared memory. The middleware also supports the integration of different time depending processes, which are distributed on a computer network. For direct georeferencing and traffic monitoring several image processing algorithms have been developed and have to be controlled in their time dependencies. This will be the main tasks of the middleware. For increased performance a shared memory access is implemented in DANAOS. This means, that modules are supported to exchange large data, especially image data without copying it explicitly. Thereby, the middleware administrates all shared memory access. For safety computation it monitors the running modules and is able to restart them.

¹ DANAOS was king of ARGOS in the fifteenth century before Chr.

2.3 Direct georeferencing

Direct georeferencing is performed by orthorectifying images using graphic processing units (GPUs) of the PCs. Orthorectification of images is the main process for all further processing steps like road segmentation and car tracking. Only if the subsequent images fit geometrically into the right coordinate system, the overlay with road databases can be achieved. Also it is necessary for integrating the image data into Geographic Information Systems (GIS). Onboard the GPS/IMU data are available in real time with 128 Hz, which are necessary for the orthorectification process. In order to rectify images Digital Surface Models (DSMs) are loaded from a database prior to flight. For holding the appropriate DSM available in memory, a Kalman-Filter is applied estimating the most probable area and triggering the DSM loading process. Then, the DSM covering this area is triangulated as fast as possible and loaded into the GPU. Beyond attitude and position, further parameters of interior and exterior orientation are required for orthorectification: The focal length, and the distortion parameters, as well as the distance from principle point to projection centre have been determined during a laboratory calibration. Up to now, we remove the radial distortion of images from the original image analytically, but we will accelerate the computation by adding an appropriate 3d-mesh to the triangulated DSM. The exterior parameters are estimated prior to traffic monitoring flight campaigns. This is done on-the-fly without ground control points based on automatically matched three-ray tie points in combination with GPS/IMU data [12].

2.4 Test site and 3K imagery

The processing chain was tested on data obtained at the motorways A95 and A96 near Munich, A 4 near Dresden, and the “Mittlere Ring” in Munich. The “Mittlere Ring” is a circular main road and serves as the backbone for the city traffic in Munich. It and the adjacent Motorways A95 and A96 are used to full capacity regularly on weekdays during rush hour, and are quite populated all day long. Therefore, these roads are good candidates to find a broad spectrum of traffic situations ranging from free flowing traffic to traffic jam. Hence, they are good targets for aerial images obtained for testing traffic monitoring applications. However, data were taken on 30 April 2007 at noon, which was not during rush hour at all. Data acquisition was performed on two flight strips, one flying ENE, covering the A96 and the western part of the “Mittlere Ring”, the other one flying WSW. Thereby, the southern part of the “Mittlere Ring” and the motorway A95 were imaged. The flight height was 1,000 m above ground for both strips which leads to a GSD of 15 cm in the nadir camera and up to 20 cm in the side-look cameras. After that, the flight track was repeated at a

flight level of 2,000 m above ground. The data obtained at the motorway in Dresden was recorded during a flight campaign on 4 August 2008 at a flight level of 1,500 m. This campaign was performed in order to validate traffic data extracted from SAR satellite “TerraSAR-X” images, which were recorded at the same time and place.

For further traffic analysis, 3K images were orthorectified using onboard GPS/IMU measurements with an absolute position error of 3 m in nadir images and around one pixel relative. The relative georeferencing error between successive images mainly influences the accuracy of the derived vehicle velocities. Based on simulations and real data, the accuracy of the measured velocity was around 5 km/h depending on the flight height [9].

2.5 Road database

Data from a road database will be used as a priori information for the automatic detection of road area and vehicles. One of these road databases has been produced by the NAVTEQ Company. The roads are given by polygons which consist of piecewise linear “edges,” grouped as “lines” if the attributes of connected edges are identical. Up to 204 attributes are assigned to each polygon, including the driving direction on motorways, which is important for automated tracking. Recent validations of position accuracy of NAVTEQ road lines resulted in 5 m accuracies for motorways.

3 Processing chain

On the data obtained as described before, the processing chain for traffic monitoring was tested. This experimental processing chain, consisting of several modules can be roughly divided into three major steps. These are road extraction, car detection, and car tracking (see also Fig. 4).

3.1 Road extraction

For an effective real time traffic analysis, the road surface needs to be clearly determined. The road extraction starts by forming a buffer zone around the roads surfaces using a road database as described above as a basis for the buffer formation process. In the next step, two different methods for further feature analysis can be applied. Both modules automatically delineate the roadsides by two linear features. One module works as follows: Within the marked buffer zone, edge detection and feature extraction techniques are used. The edge detection is based on an edge detector proposed by Phillippe Paillau for noisy SAR images [15]. Derived from Deriche filter [6] and proposed for noisy SAR images, we found this edge detector after ISEF filtering [18] efficient for our purpose of finding edges along the

roadsides and suppressing any other kind of surplus edges and noise present. With this method, mainly the edge between the tarry road area and the vegetation is found. The alternative module searches for the roadside markings by extracting lines on a dynamic threshold image. In this module, only the longest lines are kept representing the drawn through roadside marking lines. As a side effect, the dashed midline markings are detected in this module, too. These markings often cause confusion in the car detection, since they resemble white cars. However, these false alarms can be deleted from car detection, since the module for roadside marking detection finds the dashed midline markings and stores them in a separate class.

In a next step, the roadside identification module, again with the help of the road database tries to correct possible errors (gaps and bumps) that might have crept in during the feature extraction phase. Furthermore, it smoothes the sometimes curly road boundary detections from feature extraction (see Fig. 3). Gaps due to occlusion of the road surface by crossing bridges are closed, if gapping is not too large. This has the advantage that the course of the road is not lost, although the road itself is not seen at this place. However, it could lead to false alarms in the car detection. If cars are crossing the bridge, they might be assigned belonging to the occluded road below the bridge spuriously in car detection. However, we try to sort them out by alignment, since they are elongated perpendicular to the course on the occluded road.

3.2 Vehicle detection

With the information of the roadside obtained in the processing step described before, it is possible to restrict vehicle detections and tracking only to the well determined road areas. This increases performance and enhances the accuracy of vehicle detection. Based on this, we developed an algorithm for the detection of vehicles which is described in the following.

A Canny edge algorithm [5] is applied and a histogram on the edge steepness is calculated. Then, a k -means algorithm is used to split edge steepness statistics into three parts which represent three main classes. These three classes are namely edges belonging to vehicles, edges belonging to roads, and edges within road and vehicle edges, and therefore not yet classifiable.

Edges in the class with lowest steepness are ignored, while edges in the highest steepness class are directly assumed to be due to vehicles. For the histogram part with medium steepness a hysteresis threshold is applied examining neighbourhood in order to assign edges in this class either to the vehicle or the road class. In the next step, the edges belonging to the roadside markings still contaminating the vehicle class are eliminated from the histogram.



Fig. 3 Examples for road extraction (clipping from nadir images). *Upper panel* shows line detections at a flight height of 1,000 m, *middle panel* shows the resulting road area after smoothing/gap filling

(A 96 near exit Munich-Blumenau, GSD of 15 cm). *Lower panel* shows the resulting road extraction on an image obtained at a flight height of 1,500 m (motorway A 4 near Dresden, GSD of 21 cm)

As the roads are well determined by the road extraction, these roadside lines can be found easily. Thus, the algorithm erases all pixels with high edge steepness laying on a roadside position. These pixels are considered mostly belonging to the roadside markings. Thereby, the algorithm avoids erasing vehicles on the roadside by observing the width of the shape. Since vehicles are usually broader than roadside lines, this works well. Midline markings, which were detected by the roadside identification module based

on the dynamical threshold image, are erased, too. Then, potential vehicle pixels are grouped by selecting neighbored pixels. Each region is considered to be composed of potential vehicle pixels connected to each other. With the regions obtained a list of potential vehicles is produced. In order to mainly extract real vehicles from the potential vehicle list, a closing and filling of the regions is performed. Using closed shapes, the properties of vehicle shapes can be described by their direction, area, the length and width. Furthermore, it can be checked if their alignments follow the road direction, and its position on the road can be considered as well. Based on these observable parameters, we created a geometric vehicle model. The vehicles are assumed to have approximately rectangular shapes with a specific length and width oriented in the road direction. Since they are expected to be rectangular, their pixel area should be approximately equal to the product of measured length and width and vehicles must be located on the roads. In case of several detections with very low distances the algorithm assumes a detection of two shapes for the same vehicle. Then, it merges the two detections into one vehicle by calculating averages of the positions. Finally, based on this vehicle model, a quality factor for each potential vehicle is found and the best vehicles are chosen. For traffic monitoring, the camera system is in a recording mode, that we call “burst mode”. In this mode, the camera takes a series of four or five exposures with a frame rate of 3 fps, and then it pauses for several seconds. During this pause, the plane moves significantly over ground. Then, with an overlap of about 10% to 20%

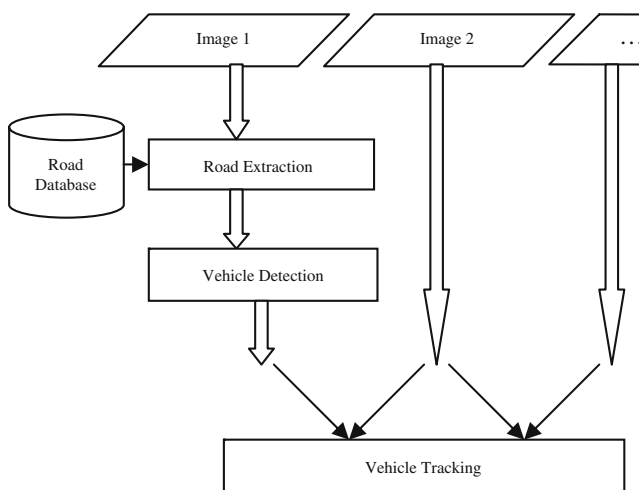


Fig. 4 Scheme of the implemented processing chain for a knowledge based road extraction, vehicle detection, and vehicle tracking on an image sequence. Mind that road extraction and vehicle detection is only performed on the first image of each exposure burst

the first exposure “burst”, the second exposure sequence is started. Continuing this periodical shift between exposure sequences and brakes, we are able to perform an area-wide traffic monitoring without producing an overwhelming amount of data. Our strategy for traffic monitoring from this exposures obtained in “burst mode” is to perform a car detection only in the first image of an image sequence and then to track the detected cars over the next images (Fig. 4).

3.3 Vehicle tracking

Vehicle tracking is based on matching by normalized cross correlation (e.g. [13]). Tracking is performed on each consecutive image pair within an exposure burst. With the vehicle detection done on the first image of the burst, vehicle tracking starts with the image pair consisting of the first and second image of an image sequence. For each vehicle detected in the first image, a circular template image of a certain radius (e.g. $r=3$ m for cars) is generated at the position of the vehicle detection in the first image. The vehicle position is transferred into the second image. There, a rectangular search window is opened aligned into driving direction starting at the vehicle position obtained from the detection in the first window. Thereby, driving direction is obtained from the road database.

The length of the search window depends on the maximum expected velocity for the road and the time difference between the two images. Then, the normalized cross correlation between the template image and second image is calculated while the template image is shifted all along the search window. The calculated correlation value gives a score for a possible hit. This value obtained lies between 0.0 and 1.0. We store the maximum score and the corresponding position in the second image. Furthermore we require the score to exceed a certain value for keeping it as a hit. We reached maximum correctness with an acceptable completeness in tracking by setting this score threshold to a value of 0.9. A vehicle detection that does not reach this threshold during correlation at any position in the search window is not tracked anymore. The program for tracking can be restarted with the second and the third image (and with further consecutive pairs of the exposure burst in succession) in order to track the vehicles through a whole image sequence. For vehicles that disappear at image borders or below bridges during an exposure of the sequence (but have been detected or tracked in the image before) the tracking algorithm does not dump a match. This means that disappeared vehicles are normally not confused with other vehicles or objects, because of the high matching threshold of 0.9. Vehicles occluded by bridges or other objects may be detected again after reappearance by a new vehicle detection performed on a further exposure sequence. However, they

appear as new detections and lose their identification relation, but this is irrelevant on our application. Due to illumination invariance vehicles normally can be tracked if they shift from fully illuminated regions into shadow regions. In order to increase correctness, cross correlation is performed as matching in RGB color space. Here, the average score obtained from cross correlation in each of the three channels is calculated and stored. This helps since vehicles are varicoloured objects. For vehicle tracking on motorways, rotations of the template vehicle image are neglected. This is valid, since the lane change angles on typical velocities obtained on motorway is quite low due to physical reasons, and hence the change in course in between two exposures (at a frame rate of 3 fps) can be neglected. However, for city regions, rotation of the template during correlation can be switched on, but this will rise in calculation time linear with the number of rotation steps during correlation. We accelerate normalized cross correlation by an estimation of the normalization, since calculating the full norm at each position in the search window costs quite a lot of calculation time. Assuming that the illumination situation does not change a lot between two images, an upper limit of the correlation score is estimated for each correlation position in the search window.

Only if this upper limit exceeds the score threshold the exact normalized cross correlation is calculated at that position. For the estimate of the score only the first channel of the RGB-image is used. These arrangements decrease calculation time by a factor of at least four.

Since vehicle tracking based on normalized cross correlation in RGB color space itself works fine at high resolutions, it is sensitive to false vehicle detections. Although several false vehicle detections can be eliminated during tracking as outliers in direction or velocity space, other false alarms still remain in tracking. Especially objects from the dashed lane markings that were detected as vehicles erroneously, may still remain in tracking. This is due to the fact, that the object shape of the dashed markings reappears periodically within a search window and the fact that all of these markings have almost exactly the same shape and intensity. Hence, the focus for improving our traffic monitoring algorithms will be placed in future on improving the vehicle detection module.

4 Results

We tested our processing chain based on the data take from 30 April 2007 as described in Section 2. For that, the completeness and correctness of vehicle detection and tracking are determined on data of several resolutions, obtained from different flight levels.

Table 1 Results on testing vehicle detection on data obtained at several test sites (from different flight heights)

Site	Correct	False	Missed	Correctness (%)	Completeness (%)
Motorway (1,000 m)	85	22	41	79	68
Motorway (1,500 m)	67	32	19	68	78
Motorway (2,000 m)	95	30	76	76	56
City (1,000 m)	47	16	25	75	65

Counts of correct vehicle detections, false alarms and missed detections, as well as correctness and completeness in percentage are given

4.1 Road detection

Road detection was performed using two different modules. It turned out, that detecting roadside markings for determining the road area is a good strategy on images taken at a lower flight height of 1,000 to 1,500 m resulting in a resolution of 15 to 21 cm GSD. Nevertheless, at higher flight levels (for instance at 2,000 m) road extraction works well with the module searching the edge between blacktop and vegetation. Figure 3 shows typical results of road extraction using roadside markings. Top image shows the line extraction, whereas in the image in the middle the finally extracted roadsides after smoothing and closing gaps are shown. Bottom image shows road extraction on a nadir image taken from a flight height of 1,500 m.

4.2 Vehicle detection

In order to quantify the vehicle detection efficiency, test data were processed and the results of the automatic vehicle detection were compared to manual car detection. Table 1 shows the results of the comparison between automatic and manual car detection. On a flight height of 1,000 m (15 cm GSD), vehicle detection performs well on motorways with a correctness of around 80% and a completeness of 68%. In

a complex scene like the city ring road we can proof that car detection delivers respectable results with a completeness of 65% and a correctness of 75%. However, at a flight height of 2,000 m (GSD=30 cm) performance drops down to 56% in completeness but correctness is still high with 76%. The testing data obtained at a flight height of 1,500 m had another illumination situation, since data were taken in the evening. This could explain the slightly reduced correctness with respect to the results obtained at other flight heights although the completeness of vehicle detection is quite high.

In Hinz [10] vehicle detection from aerial images at similar resolution (15 cm GSD) is based on matching of geometric 3D-wireframe vehicle models to the image. These models consider the viewing angle, shadow, color constancy, edge magnitude, and edge direction. A high correctness of 87% at a completeness of 60% was achieved. In case of 15 cm GSD, our completeness is slightly increased in comparison to the results of Hinz [10], whereas the correctness of our vehicle detection is marginal lower. Compared to the results of Moon et al. [14], who tested a rectangular (vehicle shaped) edge filter on aerial images of parking lots (correctness of 86%, completeness 82%) our methods have a deficit in completeness. The project ARGOS rather focus on building up a run-time optimized



Fig. 5 Examples for vehicle detection on motorways (*upper image*, A96 exit Munich–Blumenau, clipped nadir exposure) and in the city (*lower image*, Munich “Mittlerer Ring”, clipped side-look-left

exposure). *Rectangles* mark automatic vehicle detections, *triangles* point into direction of travel

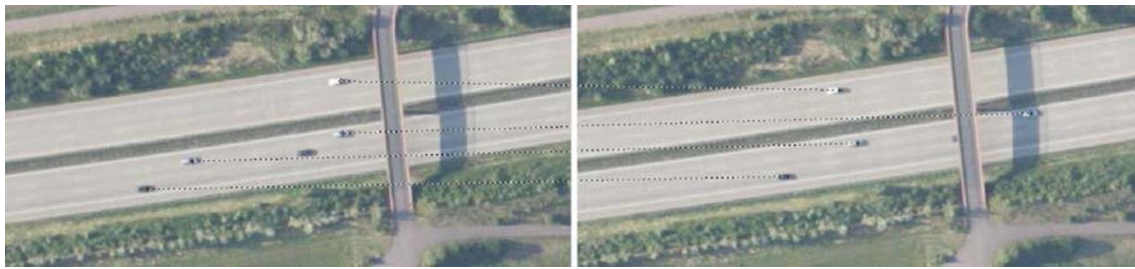


Fig. 6 Car tracking by normalized cross correlation of a group of three cars detected in the first image of a sequence (*left*) to the second image (*right*, timebase between exposures 0.7 s). Clipped images were

taken from the scene shown before at the motorway A 4 near Dresden (with a GSD of 21 cm)

complete system for online near real-time traffic monitoring than to develop new methods for highly increased detection performance. Nevertheless, we end up with sufficient detection and completeness rates. Figure 5 shows examples of vehicle detection performed on images obtained at a flight height of 1000 m. Upper image was taken on highway A96 near exit Munich–Blumenau, lower image shows part of the circular road “Mittlerer Ring” in Munich city. Only few false alarms were detected.

4.3 Vehicle tracking

Vehicle tracking was tested on the same data takes obtained at a flight height of 1,000 m (15 cm GSD), 1,500 m (21 cm GSD), and at a flight height of 2,000 m (30 cm GSD). Figure 6 shows a typical result on tracking vehicles from the first image of an image sequence into the second exposure of the sequence.

On images with a resolution of 15 cm GSD, vehicle tracking on motorways performs perfectly well, with a correctness of better than 95% and a completeness of almost 100% on each image pair. On images obtained from higher flight levels (≥ 30 cm GSD) tracking still works fine with a completeness of 90% while having a correctness of 75%. We attribute the good tracking performance on low flight heights to the fact that with a resolution of 15 cm GSD vehicle details like sunroof, windscreen and backlight, and body type go into the correlation which simplifies the search for the correct match. However, these details are not anymore seen at higher flight levels.

4.4 Performance

Traffic monitoring requires actual traffic parameters. Thus, we are planning to execute the extraction of traffic parameters on the 3×16 Mpix RGB-images in near real-time with high performance. Till now, tests on road and vehicle detection as well as vehicle tracking were performed on actual standard hardware consisting of a dual-core PC with a CPU frequency of 1.86 GHz and 2 GB RAM. The first generation of research programs for road extraction, vehicle detection, and vehicle tracking was developed within DLR in-house image processing software *XDibias* (*X*-Window *D*igital *B*avarian *I*mage *A*nalysis *S*ystem), based on C code. With this *XDibias* based prototype of the processing chain, computing times of about 2 min for images covering an area of 1 km^2 were achieved for a whole traffic extraction. In order to guarantee high actuality of traffic information and to enable near-real time traffic data extraction, the research modules were accelerated using the Machine Vision Library “HALCON” [7]. This library provides fast implementations of image processing operators due to the use of extended processor instruction sets like MMX and SSE(2), as well as due to parallel processing on multi-core CPUs. By replacing the operators used in the first generation of the traffic processor with the fast HALCON operators, we are now able to extract traffic data from images covering an area of 1 km^2 within less than 1 min.

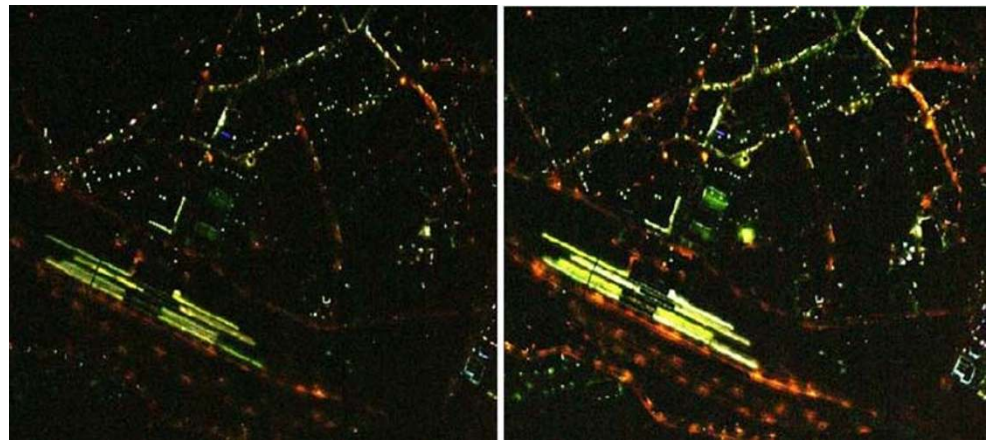
Road extraction on a typical motorway takes less than 10 s for one nadir and two side-look exposures in total. Vehicle detection on these three images needs 20 s of calculation time

Table 2 Gray values of vehicle head- and taillights and maximum total blurring

Strip	Δt (s)	Apert.	Vehicle headlights R/G/B	Vehicle taillight	Max. blurring ^a (m)
A-1	1/1,024	F1.8	\varnothing 50/-- Max $R=109$	<5	0.06+0.04
A-2	1/512	F1.8	\varnothing 102/80/61 Max $R=192$	Max $R=48$	0.13+0.08
B-1	1/512	F1.8	\varnothing 59/46/35 Max $R=67$	<5	0.13+0.08
B-2	1/800	F1.4	\varnothing 82/-- Max $R=182$	<5	0.08+0.05
B-3	1/800	F1.8	\varnothing 46/-- Max $R=82$	<5	0.08+0.05

^a Blurring caused by airplane movement (65 m/s) and vehicle movement (max. 150 km/h)

Fig. 7 Comparison of orthorectified night images from flight strips A-1 (left) with flight strip A-2 (right). Left image has an exposure time of 1/1,024 s, the exposure time of right image is 1/512 s at same aperture F 1.8 and ISO 1600



on the present system. In comparison, car tracking is quite fast, consuming only 15 s for a tracking of 3×15 cars over an image sequence consisting of 3×4 images. Moreover, the pure calculation time for cross correlation is 30 ms per vehicle for a tracking through the whole sequence. In total, it costs less than 60 s to analyse the traffic within one image sequence. However, the onboard computer system for traffic monitoring will possess a multi-core CPU with at least four cores. By sufficient parallelization of the processes that will be managed by the middleware DANAOS, we expect to be able reducing the processing time by a factor of 2. That means, assuming a break of 7 s between each image “burst” (which would result in an overlap of 10% between two image “bursts” at a flight speed of 60 m/s and a flight height of 1,000 m), we will have a time overhead in the processing chain for traffic monitoring of a factor of 4. However, the prototype of our processing chain is built up still modular, which means that each module in the chain reads an image from hard disk into memory, performs an operation, and at the end writes a new image to hard disk. We estimate to halve the overhead by reducing hard disk read/write. Nevertheless we are already able to perform automatic traffic data extraction on a large amount of data in near real-time. Therefore, the system already provides area wide traffic data with a high actuality, with capacities of increasing performance in near future.

5 Night shot capability

During a test flight on 6 May 2008 from 9:40 PM until 10:28 PM near the city Rosenheim and the motorway junction “Inntaldreieck” we were able to show the capabilities of the 3K camera system for traffic monitoring applications at night. Two strips which cover a part of the city (strip A) of Rosenheim and the motorway (strip B) were acquired repeatedly with different camera configurations. Flight height was 2,000 m above ground, flight speed was 65 m/s. For this test flight, the sensor was set into a special configuration called along-track modus. In order to

increase the chance for recording car headlights the camera platform was rotated with an angle of 90° azimuthally. Hence, one of the former side-looking cameras was aligned in flight direction, the other side-looking camera was now looking in backward direction. With an off nadir angle of 35° and a flight direction along the motorways we expected the forward camera being able to detect headlights of forthcoming vehicles, and the backward camera to record the headlights of cars travelling in flight direction. In this configuration, a nadir image covers an area of 1.4×1.0 km; the ground pixel size is around 29 cm.

Figure 7 shows two orthorectified images (A-1 and A-2) from the city of Rosenheim taken with different camera configuration. As the exposure time in A-2 with 1/512 s is double than in A-1, more lights from the city of Rosenheim are visible, but also the motion blurring is more visible.

With respect to traffic monitoring applications, the visibility of vehicle head- and taillights in the images is of great interest. An object is defined as visible with an absolute gray value more than five, as the image noise is around two to three gray values. In Table 2, the visibility of

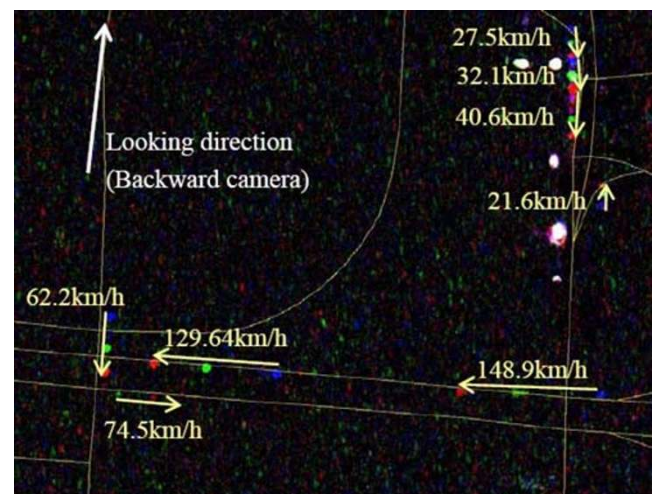


Fig. 8 RGB composition of image sequences from flight strip A-2 used for traffic monitoring

head- and taillights in the different data sets is listed. Taillights are only in strip A-2 visible with a maximal value of 48 in *R* band, headlights were visible in all strips. The average *R* values range from 46 in strip B-3 to 102 in strip A-2, the *B* and *G* values are in general lower. The total blurring consists for moving objects of the blurring caused by the airplane and of the moving objects. For a moving object with a ground speed of 150 km/h in opposite direction to the airplane movement, the total blurring is around 0.21 m in strip A-2 and B-1.

We propose RGB compositions of image sequences to visualize moving objects in the images. For this, the red channels of the orthorectified images from the sequence are overlaid and composed to a RGB image again. Figure 8 shows an example RGB composition of the motorway south of Rosenheim. A moving vehicle appears in the RGB composition as an array of a blue, a green, and a red point where the color blue/green/red corresponds to the first/second/third image in the sequence. Static objects like street lamps or illuminated traffic signs appear white.

Based on this point pattern, automatic vehicle detection could be applied and the moving direction of the vehicles and the speed could be derived. Since algorithms for automatic traffic extraction from night exposures have not yet been developed, manually measured vehicle directions and vehicle speeds are visualized in Fig. 8. Vehicle velocities were calculated by measuring manually the distance and using the time span between the acquisition times which can be derived with high accuracy from the GPS/IMU data.

The accuracy of speed determination is influenced not only by the accuracy of georeferencing but also from blurring effects caused by the exposure time, as the distance measurement is not so precise. In the examples in Fig. 8 it could be seen, that vehicles are detected by head- and taillights from the front as well as from the side, i.e. traffic in different directions can be detected. Information about completeness and correctness of vehicle visibility are not available as no ground truth data were acquired. Furthermore, no algorithms for automatic traffic data extraction on night shots are available at this time and have to be developed in future. Hence, we could show that using this optical sensor system for traffic monitoring under night conditions is basically possible, which might be an interesting field for future research and development.

6 Conclusions

Despite the large amount of incoming data from the wide angle camera system, we are able to perform traffic data extraction with high actuality in near real-time. This means that the processing chain is capable to perform a complete

traffic data extraction on an area of 1 km² within few 10 s. Thereby, high accuracies for velocities (5 km/h), good correctness in vehicle detection (79%) and in vehicle tracking (90% of detected vehicles) is reached. Furthermore, the system performs image orthorectification in real-time using GPU computing power. Although algorithms for automatic traffic monitoring at night have not yet been developed the capability of the system to provide traffic information at night has been demonstrated successfully during a test flight.

Hence, the investigations show the high potential using aerial wide angle image time series for traffic monitoring and similar applications, like the estimation of travel times or the derivation of other relevant traffic parameters. In future, the data processing speed will be further improved by converting the modules of the processing chain into tasks that share memory access to image data stored in the RAM instead of reading and writing them on hard disk, as done by our prototype. We further plan to evaluate the performance of the system in case of difficult scenes such as large cities with high buildings occluding parts of the roads and under various weather conditions (e.g. snow, wet roads) during three campaigns in 2009. Moreover it is planned to use an additional radar sensor providing traffic data in case of bad visibility conditions (e.g. clouds, fog) where remote sensing traffic monitoring based on optical sensors fails.

The whole system is thought to be a technology test bed for future traffic monitoring applications and it is in operation at a DLR research aircraft. This limits the operations at the moment only to campaigns on demand, like mass events or in case of disasters. However, this prototype of a traffic monitoring system or a successor version of this system could be mounted to any other carrier such as UAV or HALE in future. This would enable continuous and area-wide traffic monitoring in metropolitan areas at high actuality without the use of stationary infrastructure.

References

- Behrisch M, Bonert M, Brockfeld E, Krajzewicz D, Wagner P (2008) Event traffic forecast for metropolitan areas based on microscopic simulation. Third International Symposium of Transport Simulation 2008 (ISTS08), Queensland, Australia
- Bethke K-H, Baumgartner S, Gabele M (2007) Airborne road traffic monitoring with RADAR. World Congress on Intelligent Transport Systems (ITS), Beijing, China, pp 1–6
- Börner A, Ernst I, Ruhé M, Sujew S, Hetscher M (2004) Airborne camera experiments for traffic monitoring. XXth ISPRS Congress, 12–23 July 2004, Vol. XXXV, Part B, 6 p
- Busch F, Glas F, Bermann E (2004) Dispositionssysteme als FCD-Quellen für eine verbesserte Verkehrs-lagerekonstruktion in Städten-eine Überblick. Straßen-verkehrstechnik 09/04
- Canny JF (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8(6):679–698

6. Deriche R (1987) Using Canny's criteria to derive an optimal edge detector recursively implemented. *Int J Comput Vis* 1(2):167–187
7. Eckstein W, Steger C (1999) The Halcon vision system: an example for flexible software architecture. 3rd Japanese Conference on Practical Applications of Real-Time Image Processing, Technical Committee of Image Processing Applications, Japanese Society for Precision Engineering, pp 18–23
8. Ernst I, Sujew S, Thiessenhusen K-U, Hetscher M, Raßmann S, Ruhé M (2003) LUMOS-Airborne Traffic Monitoring System. Proceedings of 6th IEEE International Conference on Intelligent Transportation Systems, 12–15 October 2003, Shanghai, China
9. Hinz S, Kurz F, Weihing D, Suchandt S, Meyer F, Bamler R (2007) Evaluation of traffic monitoring based on spatio-temporal co-registration of SAR data and optical image sequences. *PFG—Photogrammetrie–Fernerkundung–Geoinformation*, 5/2007, pp 309–325
10. Hinz S (2004) Detection of vehicle queues in high resolution aerial images. *PFG—Photogrammetrie–Fernerkundung–Geoinformation*, 3/2004, pp 201–213
11. Kurz F, Charmette B, Suri S, Rosenbaum D, Spangler M, Leonhardt A, Bachleitner M, Stätter R, Reinartz P (2007a) Automatic traffic monitoring with an airborne wide-angle digital camera system for estimation of travel times. In: Stilla U, Mayer H, Rottensteiner F, Heipke C, Hinz S (eds) *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. 36 (3/W49B), Institute of Photogrammetry and Cartography, Technische Universität München, pp 83–86
12. Kurz F, Müller R, Stephani M, Reinartz P, Schroeder M (2007b) Calibration of a wide-angle digital camera system for near real time scenarios. In: Heipke C, Jacobsen K, Gerke M (eds) *ISPRS Hannover Workshop 2007, High Resolution Earth Imaging for Geospatial Information*, Hannover, 2007-05-29-2007-06-01, ISSN 1682-1777
13. Lewis JP (1995) Fast normalized cross correlation. *Vision Interface, Canadian Image Processing and Pattern Recognition Society*, pp 120–123
14. Moon H, Chellappa R, Rosenfeld A (2002) Performance analysis of a simple vehicle detection algorithm. *Image Vis Comput* 20/1:1–13
15. Paillau P (1997) Detecting step edges in noisy SAR images: a new linear operator. *IEEE Trans Geosci Remote Sens* 35(1):191–196
16. Reinartz P, Lachaise M, Schmeer E, Krauss T, Runge H (2006) Traffic monitoring with serial images from airborne cameras. *ISPRS J Photogramm Remote Sens* 61:149–158
17. Schaefer R-P, Thiessenhusen K-U, Wagner P (2002) A traffic information system by means of real-time floating-car data. Proceedings of ITS World Congress, October 2002, Chicago, USA
18. Shen J, Castan S (1992) An optimal linear operator for step edge detection. *CVGIP, Graph Models Image Process* 54(2):112–133
19. Suchandt S, Runge H, Breit H, Kotenkov A, Weihing D, Hinz S (2008) Traffic measurement with TerraSAR-X: processing system overview and first results. VDE: Proceedings of EUSAR 2008, Friedrichshafen, Germany, VDE Verlag GmbH, pp 55–58