REAL-TIME IMAGE PROCESSING FOR ROAD TRAFFIC DATA EXTRACTION FROM AERIAL IMAGES

D. Rosenbaum, J. Leitloff, F. Kurz, O. Meynberg, and T. Reize

DLR - German Aerospace Center, Remote Sensing Technology Institute,
Münchner Str. 20, 82234, Weßling, Germany

Commission VII Symposium 2010

KEY WORDS: Monitoring, Recognition, Orthorectification, Georeferencing, Image, Pattern, Sequences, Tracking

ABSTRACT:

A world with growing individual traffic requires sufficient solutions for traffic monitoring and guidance. The actual ground based approaches for traffic data collection may be barely sufficient for everyday life, but they will fail in case of disasters and mass events. Therefore, a road traffic monitoring solution based on an airborne wide area camera system has been currently developed by DLR. Here, we present a new image processing chain for real-time traffic data extraction from high resolution aerial image sequences with automatic methods. This processing chain is applied in a computer network as part of an operational sensor system for traffic monitoring onboard a DLR aircraft. It is capable of processing aerial images obtained with a frame rate of up to 3 Hz. The footprint area of the three viewing directions of an image exposure with three cameras is 4 x 1 km at a resolution of 20 cm (recorded at a flight height of 1500 m). The processing chain consists of a module for data readout from the cameras and for the synchronization of the images with the GPS/IMU navigation data (used for direct georeferencing) and a module for orthorectification of the images. Traffic data is extracted by a further module based on a priori knowledge from a road database of the approximate location of road axes in the georeferenced and orthorectified images. Vehicle detection is performed by a combination of Adaboost using Haar-like features for pixel wise classification and subsequent clustering by Support Vector Machine based on a set of statistical features of the classified pixel. In order to obtain velocities, vehicle tracking is applied to consecutive images after performing vehicle detection on the first image of the burst. This is done by template matching along a search space aligned to road axes based on normalized cross correlation in RGB color space. With this processing chain we are able to obtain accurate traffic data with completeness and correctness both higher than 80 % at high actuality for varying and complex image scenes. The proposed processing chain is evaluated on a huge number of images including inner city scenes of Cologne and Munich, demonstrating the robustness of our work in operational use.

1 INTRODUCTION

Mass events with a large attendance hold in big cities overload road infrastructure at regular intervals, since metropolis roads are used to full capacity in normal course of life. In case of disaster, maybe with parts of the road network being impassable, total collapse of road traffics menaces. Both cases require sufficient methods of traffic monitoring and guidance. Common infrastructure for traffic monitoring like induction loops and video cameras is ground based and mainly distributed on main roads. New monitoring approaches collect data by means of mobile measurement units which flow with the traffic as test particles for local traffic situations and travel times. The so called floating car data (FCD, e.g. Schaefer et al., 2002; Busch et al., 2004) obtained from taxicabs can deliver useful traffic information within cities, but they are only available in few big cities today.

These methods and sensors are suited for everyday life, but traffic monitoring and guidance based on these sensors may fail in case of mass events or disaster. In those cases a coverage as complete as possible of road level of service is mandatory. Such a complete coverage could not be provided by common sensor networks with their low spacial resolution data obtained by punctual traffic registration limited to main roads as provided by ground based sensor networks. In case of disaster, sensor networks based on ground infrastructure may fail completely. Not only publicity but especially security authorities and organizations which have to coordinate and route action and relief forces into and within affected areas require precise traffic information. Furthermore, a spacial complete area wide traffic surveillance at high actuality provides the possibility to generate precise predictions of traffic situation in near future by simulations (e.g. Behrisch et al., 2008).

For these purposes airborne and satellite based solutions for wide area traffic monitoring have been produced or are currently under development at DLR. In Reinartz et al. (2006) 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" (Ernst et al., 2003; Börner et al., 2004). The actual project for airborne traffic monitoring and simulation is called "VABENE" (German: *V*erkehrsm*a*nagement *be*i Großereignisse*n* und Katastroph*e*n, that means: traffic management under mass event and disaster conditions). Aim of the project among others is to develop operational airborne optical and radar systems for automatic traffic data extraction in real time. First results on traffic monitoring based on remote sensing synthetic aperture radar (SAR) systems were already shown in e.g. Bethke et al. (2007) or Suchandt et al. (2006). On the optical regime, a first proposal for a prototype processing chain capable of traffic data extraction from sequences of optical aerial images in near real time was shown in Rosenbaum et al. (2008). There, algorithms and methods for edge and line based roadside extraction, vehicle detection based on edge detection and geometry validation, and vehicle tracking by template matching using a normalized cross correlation operator were presented. That template matching for car tracking works sufficiently had already been proven in Lenhart et al. (2008). There, a similar algorithm using a shape based matching operator was introduced. Furthermore, different methods for validation of potential vehicle tracks and vehicle detections for outlier elimination were presented. Some of them are used in the present processing chain.

The present publication shows a vehicle detection approach based on the Viola-Jones detector (Viola and Jones, 2004) trained by Gentle AdaBoost (Friedman et al., 2000). Vehicle detectors based on Boosting had already been applied to aerial images successfully in literature (e.g. Nguyen et al., 2007; Grabner et al., 2008).

In the optical system for online traffic monitoring the DLR inhouse developed sensor called 3 K camera (e.g. Kurz et al., 2007) is included, which is capable of direct orthorectification / georeferencing in conjunction with a IGI IId Aerocontrol RT IMU/GPS navigation system and several PCs for image processing. Furthermore, the sensor provides a high image repetition rate of up to 3 Hz which offers a high overlap of sequential images and makes vehicle tracking possible. With a big footprint of 4 km across track at a typical flight level of 1500 m over ground and a high native resolution of 20 cm GSD (Nadir) it is well suited for recording road traffic data. After several years of development our system has reached an operational state and is ready for application in disasters and mass events.

The paper is structured as follows. Second section presents the system for airborne traffic monitoring and the data sets obtained. Then, the processing chain for automatic road traffic data extraction from aerial images is introduced in section 3. Section 4 deals with the validation of the processing chain, and the last section gives conclusions and presents the plans for future work.

2 SYSTEM AND DATA SETS

The real-time road traffic monitoring system consists of two parts. One part is installed onboard the aircraft, consisting of the 3 K 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 5 Mbit/s (actual, upgradeable to a bandwidth of about 20 Mbit/s) automatically tracking the ground station, and a PC for steering the antenna and feeding the downlink with data. 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 downlinked images and traffic data. Given an internet access at the place of the ground station, the obtained traffic data will be directly transferred to an internet traffic portal.

2.1 Onboard System

The system for traffic monitoring aboard the airplane uses aerial image sequences obtained with the so called 3 K camera system. All results shown in this publication are based on this sensor. In near future this optical sensor will be replaced by the successor, the 3 K+ camera system. Both wide area digital frame camera systems offer similar properties (Tab. 1). Each consists of three non-metric Canon EOS 1Ds cameras mounted on a ZEISS aerial platform. One look in nadir direction and two looks in oblique sideward direction result in an increased FOV of up to 104 degree or 31 degree in side resp. flight direction based on a focal length of 50 mm. The cameras acquire images with a frame rate of up to 5Hz; here the absolute number of images is limited due to an overflow of the internal memory. Based on the image size of up to 21 MPix and a colour depth of 24 bits the overall output data rate of the three camera system lay between 6 and 10 MByte/s for jpeg compressed images. The data rates at the cameras depend also on the flight and image acquisition mode, e.g. the overlap and the flight height. The image acquisition geometry of the DLR 3 K and 3 K+ camera system are similar except the smaller GSD of the 3 K+ system. Typical flight heights of the camera systems are between 500 m and 3000 m above ground.

	3K camera system	3K+ camera system
Cameras	3 × EOS 1Ds Mark II	3 × EOS 1Ds Mark III
Image size	4992 ×3328 (16,7MPix)	5616 × 3744 (21,0 MPix)
Max. frame rate	3 Hz (~50 images*)	5Hz (63images*)
File size	20MByte (RAW) 5,5MByte (JPEG level 8)	25MByte (RAW) 6,5 MByte (JPEG level 8)
ISO	100 - 1600	50 - 3200
Aperture	1.4 – 22	1.4 - 22
Lenses	Canon EF 1.4 50mm	Zeiss Makro-Planar 2/50mm
Tilt angle of sideward cameras	Max 32° /variable	Max 32° / variable
Data rates (3 cameras, JPEG): 60% overlap, 1000m a.g., 2s 3xbursts (2Hz), 1000m a.g., 6s pause (traffic modus)	 8,3MByte/s 6,6MByte/s	 9,8MByte/s 7,8MByte/s
Interface	Firewire IEEE 1394a	USB 2.0
Image sensor	full frame CMOS sensor	full frame CMOS sensor
Pixel size	7.21µm	6.41µm
Footprint / GSD, 1000m a.g.	2560m × 480m / 15cm nadir	2560m × 480m / 13cm nadir
Footprint / GSD, 3000m a.g.	7680m × 1440m / 45cm nadir	7680m × 1440m / 39cm nadir
FOV (side resp. in flight)	104°/31°	104°/31°
Calibration (interior orientation)	5 parameters: focal length, focal point, radial distortion A_1 and A_2	Not calibrated yet.
GPS/IMU accuracies postproc. (rp/yaw/xy/z) real time (rp/yaw/xy/z)	IGI IId 0.003°/0.007°/0.08m/0.05m 0.01°/0.05°/0.1m/0.1m	IGI IId 0.003°/0.007°/0.08m/0.05m 0.01°/0.05°/0.1m/0.1m
Georeferencing accuracies Direct georeferencing plus bundle adjustment	 2-4m <1m	Not tested yet.

*Depending on the motif and other configuration parameters

Table 1: Overview about 3K and 3K+ camera system.

The high input data rate on the one hand and the processing intensive modules on the other hand put high demands on the on-board image processing hardware, which consequently leads to a multihost solution with five PCs in total (Fig. 1). All of them run 32bit-Windows XP due to the fact that some of the third-party software we use in our processing system only supports Windows. Each camera is connected via Firewire IEEE 1394a to a dedicated host. It streams the images directly without memory-card buffering to the camera PCs (PC1 - PC3). The EOS Digital Camera Software Development Kit (EDSDK) is installed on each of these hosts and provides a C language interface for the control of the cameras and the download of images to the host PC. Supported operating systems are Microsoft Windows 2000 or higher and Mac OS X since version 10.4. Since the orthorectification and georeferencing process needs the exact position and orientation of the airplane, the IGI IId GPS/IMU system is connected via Ethernet to the onboard system. The fibre-optic gyro based Inertial Measurement Unit with its integrated 12-channel L1/L2 GPS receiver are triggered by the cameras external flash signal. Every time a flash signal is received, the IGI IId sends coordinates and orientation via a TCP connection to one of the camera PCs. A module runs a TCP client and matches the received geo-data with the image received from the camera. The image is written to disk where it can be read by the orthorectification module. The geo data is sent to this module via message passing. After the orthorectification process (Müller et al., 2002) has completed, the orthorectified image is written to disk, copied to PC 4 and could be processed for automatic traffic data extraction or sent down to ground station.

The onboard system was ready to fly in mid 2009. Unfortunately the assignation of the permit to fly certificate of this system onboard the DLR aircraft Cessna 208 B was delayed. We got the certificate in April 2010 and we are currently carrying out series of test flights. During first flights, several problems have been exposed and fixed as well as an optimal start-up procedure for hardware and software of this complex system had to be found. Meanwhile the sensor and system components for orthorectification and automatic traffic data extraction are running properly during test flights and we recorded road traffic data aboard the aircraft during flights successfully. Due to a hardware defect of the S-band downlink, we have not yet transmitted traffic data to the ground, but we are sure to catch up on this very soon. Although the onboard traffic data processing is already working, re-

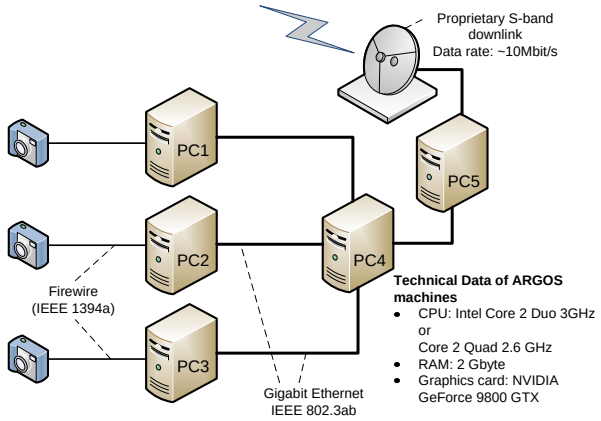

Figure 1: Topology of the onboard network. Each camera is directly connected to PC1,2,3 which perform georeferencing and orthorectification. Vehicle detection and tracking takes place on PC4. PC5 sends the results via a proprietary S-band transmitter to the ground.

sults presented in this publication are based on data processed offline after flight at the ground for evaluating the processing chain.

2.2 Test Sites and Imagery

In the previous and current project the metropolitan areas Munich and Cologne have been selected for system flight tests and to demonstrate the airborne traffic monitoring system to the stakeholders. Therefore the NAVTEQ road database installed on the system is limited to these regions. Hence, all data and results presented here are obtained during four campaigns in Munich and Cologne.

First campaign was performed in Munich on 30.4.2007 with two flight strips across the city core and the main city ring road at sunny weather conditions and a flight height of 1000 m. Second campaign was flown in Cologne on 2.06.2009 at a flight height of 1500 m. Third and fourth campaign were performed in Munich on 22.04.2010 and 24.04.2010 during BAUMA exhibition, both during fair to nice weather conditions at a flight height of 1000 m above ground.

All images were resampled to a resolution of 20 cm GSD independent from flight height, since vehicle detection is trained to that resolution. For evaluation in section 4 traffic data of the campaign three and four were analysed. Due to the huge amount of images, not every flight strip of the campaigns was probed for manual reference. However, the analyzed flight strips give a good mix of scenarios ranging from downtown over suburban to rural areas including road categories from lowest category back road over inner city main road to interstate motorway. Results from the first and fourth campaign are shown in Figures 3, 4, and 6.

During all campaigns images were taken in burst mode with 3 images per burst and a repetion rate of 2 Hz. During flight strips the bursts were activated every 7 seconds. Image recording in burst mode makes vehicle tracking and measuring vehicle speed possible within burst image pairs but reduces amount of data produced significantly in comparison to a continuous mode with a high frame rate. Images were resampled to a resolution of 20 cm GSD

3 PROCESSING CHAIN

Automatic traffic data extraction works on the georeferenced images produced by the orthorectification process in the step before.

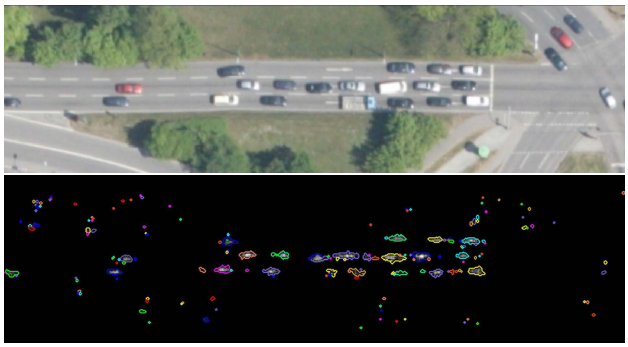

Figure 2: Upper: Horizontal aligned original image.
Lower: Confidence image with zero crossing.

Preprocessing consists of a first module reading images from the camera ring buffer and receiving the corresponding exterior orientation from the GPS/IMU navigation system in real time. Second preprocessing module performs direct georeferencing / orthorectification (section 2.1). Each image burst is processed separately. Processing for traffic data extraction starts with vehicle detection on the first image of each image sequence (burst) followed by a vehicle tracking based on template matching between the image pair consisting of the first and second image of the sequence.

With this strategy the traffic parameters flux and density can be derived from image sequences. Vehicle detection performed on the first georeferenced image of the sequence gives the vehicle density whereas vehicle tracking applied on all consecutive (georeferenced) image pairs gives the velocity of the vehicles so that the flux of the traffic can be calculated.

3.1 Vehicle Detection

Prior to vehicle detection, image buffers along road axes obtained from NAVTEQ road database are cut out of the georeferenced images. This reduces the search space for vehicles in the images. Then all roads are straightened along road axes, so that vehicles travelling along each road in the image are aligned in the same direction. This reduces calculation time, since the classifiers of vehicle detection do not have to be rotated.

In this work the vehicle detection is performed in three stages. The first step is based on the Gentle AdaBoost algorithm as introduced in Friedman et al. (2000). Compared to other boosting variants, e.g. Real AdaBoost (Schapire and Singer, 1999) or Discrete AdaBoost (Freund and Schapire, 1997), Gentle AdaBoost is less sensitive to errors in the training set as reported by Friedman et al. (2000) and Lienhart et al. (2003). Furthermore, AdaBoost is the only classification algorithm with can handle the huge amount of features used in this work. Our training database contains 4662 sample of vehicles under different lightning condition and with various appearance. Furthermore, 15,000 negative samples are generated from designated areas. The complete set of standard and rotated Haar-like features as presented in Lienhart et al. (2003) are calculated for each sample in a 30 by 30 search window. The total number of 657,510 features for the horizontal aligned gray value image of each training sample is reduced to the 200 most significant features during the training. Only these few features are used for later classification, which can be fast calculated by use of integral images as shown in Viola and Jones (2004). The application of the learned classifier results in confidence rated image reflecting the probability of each pixel belonging to a vehicle or the background respectively. Afterwards the confidence image is clustered by means of zero crossings as

Figure 3: Upper: Confidence image with detected blobs (red circles) and final SVM detection (green crosses).
Lower: Original image with final detection.

shown in Fig. 2. Since the confidence image is normalized, grey values range from -1 to 1 with zero crossings typically closuring regions of pixels potentially belonging to vehicles.

As it can be seen in the confidence image, vehicle areas exhibit a blob-like structure. Thus, in the second processing stage an interest point operator for blobs was implemented based on the work of Lepetit and Fua (2006). The parameters of the algorithm have been tuned for the used image resolution of 20 cm by 5-fold cross validation. These parameters mostly reflect geometrical properties of the vehicle clusters. Thus 80 % the non-vehicle areas can be rejected from further processing. Nearly all remaining wrong hypotheses are classified in the last processing stage. Therefore, a number of statistical values are calculated from geometric and radiometric properties of the remaining clusters in the confidence image and in all channels of the RGB image. Due to the partially high correlation between those channels the total number of more 100 statistical features is reduced by principal component analysis (PCA) transformation to the first 40 components which contain 99 % of the descriptive information. This reduced feature set is used to train a Support Vector Machine (SVM). The slack variables and kernel type of the SVM are also optimized for the specific resolution by cross validation leading to an average False-Positive-Rate of approximately 12 %. As it will be shown 4 section, this accuracy is reflected in the correctness of the numerical evaluation. Figure 3 shows the results of the interest point operator (marked by circles) and the final vehicle detection (marked by crosses).

3.2 Vehicle Tracking

Vehicle tracking between two consecutive images of the burst is done by template matching based on normalized cross correlation. At each position of a detected vehicle in the first image of the image sequence a template image is created. Then, in the second image of the sequence a search space for this vehicle is generated. Its size and position depends on the position of the vehicle in the first image, driving direction obtained from NAVTEQ road database, and the expected maximum speed for the road plus a certain tolerance. Within that search space, the template is correlated and the maximum correlation score is stored in connection with the template position within the maximum appeared. This normally represents the found match of each vehicle in generally. The correlation is done in RGB-color space. Fig 4 shows a typical result of the tracking algorithm obtained on the motorway A96 near Munich. Left image was taken $0.5\ s$ before right image. The dashed lines show corresponding matching pairs from normalized cross correlation. Since all images are stored with their recording time, vehicle speed can directly be calculated from both the

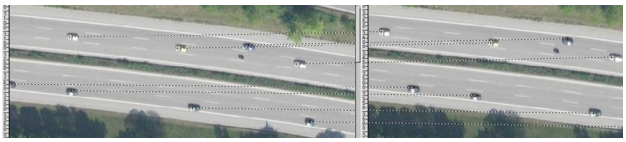

Figure 4: Tracking of a group of cars on motorway A96 near Munich. Corresponding matches are marked by dashed lines. Mind, that the motorbike was not tracked, because it was not detected (the classifier of detection was not trained to two-wheeled vehicles).

position of the vehicle detected in the first image and the position of the corresponding match in the second image. Then, vehicle tracking is applied to the following image pair of the sequence. Several measures to chase mismatches are imlemented mainly based on plausibility of velocity and driving direction, constance of velocity and driving direction within a burst, and plausibility of individual speed and direction with respect to local average values. Several potential mismatches as well as matches based on false positive vehicle detection can be eliminated that way.

After traffic data extraction the results are immediately copied to PC 5 (Fig. 1) and directly sent to the ground via S-band downlink. There, data can be used in a traffic internet portal for road level of service visualization and for traffic simulation.

3.3 Performance

Actuality of road traffic data is a general concern. For the use of aerial recorded data in the traffic simulation an actuality of less than five minutes is required. This means between exposure and receiving traffic data on the ground a maximum delay of five minutes is permitted. Hence, the processing chain must be optimized for processing speed. If traffic data extraction is limited to main roads, the bottle neck of the chain is produced by the orthorectification process that takes 10 to 12 s for each nadir image. The actuality criterion is fulfilled for the first bursts of each flight strip easily, but a stack of unprocessed images is built up that leads to a critical length of the flight path. If the critical length is exceeded, the actuality criterion of the simulation will be overrun. This critical length of the flight path can be estimated. Taking into account a typical flight speed of 70 m/s, 3 images recorded per burst, a break of 7 s between each burst, the critical length is around 5 km. In case of full traffic data extraction in urban cores the bottle neck moves to the traffic processor that slightly cannot keep orthorectification performance (Fig. 5). Nevertheless, for road level of service visualization, the performance of the present processing chain is sufficient, since the hard actuality criterion of the simulation does not apply in that case. However, the present processing chain holds potential for improvement of calculation time, even in the orthorectification process (section 5).

Scene	Suburban & Motorways	Urban Core	Total
Images evaluated	73	6	79
True Positives	5545	2911	8456
False Positives	613	429	1042
False Negatives	424	317	741
Correctness	90 %	87 %	89 %
Completeness	93 %	90 %	92 %
Quality	84 %	80 %	83 %

Table 2: Evaluation of vehicle detection quality.

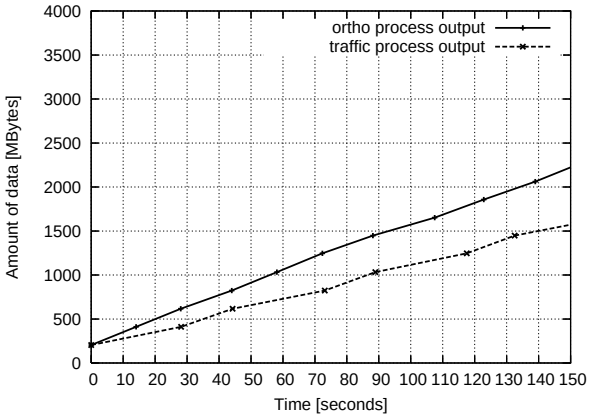

Figure 5: The diagram shows the amount of data which is processed by the Ortho Module but not by the Traffic Processor in case of full traffic data extraction including all road categories and urban core scenarios.

Campaign	BAUMA 1	BAUMA 2	Total
Bursts evaluated	29	19	48
True Positives	1002	724	1726
False Positives	75	47	122
False Negatives	36	23	59
Correctness	93 %	94 %	93 %
Completeness	97 %	97 %	97 %
Quality	90%	91%	91 %

Table 3: Evaluation of vehicle tracking quality.

4 RESULTS AND ACCURACY

Quality of traffic data obtained from the processing chain was evaluated on image sequences obtained on Cologne and "BAUMA" Munich campaigns. The Cologne image data were used for validating detection accuracy, whereas Munich data were used for evaluating tracking. The data represent a mix of scenarios ranging from urban core and suburban roads to motorways. Due to the higher complexity of the scenes, the quality of traffic data extracted in metropolitan cores is somewhat lower than on motorways and suburban areas. This deposits in vehicle detection notably. Therefore, the quality of vehicle detection is examined while distinguishing between these different scenes. However, each campaign contains a mix of both scenes, since German city cores are typically limited to a diameter of a few kilometers, even in German major cities. Therefore, a total quality containing both scene types is listed as well, giving the typical quality of the detection to be expected in operational use. The evaluation of tracking distinguishes between the two Munich "BAUMA" campaigns, since the weather and illumination conditions were better at the flight on 22.04. The definitions for completeness, correctness and quality are:

$$Correctness = \frac{true\ positives}{true\ positives + false\ positives}$$
$$Completeness = \frac{true\ positives}{true\ positives + false\ negatives}$$
$$Quality = \frac{true\ pos}{true\ pos + false\ pos + false\ neg}$$

with *true positives* being the number of vehicles detected, *false positives* the number of non-vehicle detections, and *false negatives* the number of vehicles missed. In tracking evaluation *true positives* is the number of vehicles tracked correctly, *false positives* the number of mistracked vehicles, and *false negatives* the number of vehicles detected, but not tracked (vehicles that are obscured during image bursts and cannot be tracked are not counted as *false positives*).

Table 2 shows the results of detection algorithm. With a value of around 90 % correctness of vehicle detection is at high level in total and in suburban and motorway situations. In urban cores the correctness drops slightly to 87 %. The completeness with 93 % in suburban and motorway scenes and 92 % in total is at high level again providing precise estimates of local traffic density for traffic simulations or road level of service visualization. In urban core regions it drops a little to a value of 90 %. This all results in a total quality of 83 %. Table 3 presents the results of evaluating vehicle tracking algorithm. The underlying image database represents a mix of urban core and suburban/motorway scenes. Tracking performs well with a correctness of about 93 % and a completeness of 97 %. There is nearly no difference between the results obtained from two flights at different dates indicating that tracking algorithm might be robust against slight changes in weather and illumination conditions.

The system accuracy is the product of detection and tracking quality. In the present case it is about 75 %. This makes the system accuracy competitive to that of ground based sensor networks. The quality of a typical induction loop is better than 90 % (e.g. Leonhardt, 2004) at time of production. During years of continuous operation it decreases slowly. In a complete sensor network of a metropolitan area there is a mix of new and older sensors, and some of them have even failed completely. This drops down the average quality or system accuracy of the sensor network. In Munich, the system accuracy of the complete ground based sensor network for traffic data acquisition is at a value of 80 %, as stated by Munich integrated traffic management center. The quality of traffic data obtained by the presented remote sensing system being competitive to that of road sensor networks in connection with the good spacial resolution (including minor roads) and the reliability in case of ground damage makes the system well suited for its operation during mass events and disasters.

5 CONCLUSIONS AND FUTURE WORK

The quality of traffic data obtained by the presented remote sensing system being competitive to that of road sensor networks in connection with the good spacial resolution (including minor roads) and the reliability in case of ground damage makes the system well suited for its operation during mass events and disasters. In the evaluated case with a mixture of urban core, suburban roads, and (local) motorways, the overall system accuracy is approximately 75 %. This accuracy value is comparable to that of typical metropolis ground based sensor networks with the advantage of the remote sensing system, that traffic data on all road categories (including smallest minor roads) can be recorded. E.g. in case of traffic congestion on main roads and motorways not only the traffic density of the main road but also densities on alternative routes (that are not covered by induction loops or stationary traffic cameras) can be determined. However, the operating costs of this prototype system are quite high which limits its application to temporally limited scenarios like mass events or disasters. With the ongoing development of UAV or HALE unmanned aircrafts, further applicability of the system may occur in future.

The system is sufficient for real-time traffic data extraction from aerial image time series on main roads, even in metropolitan core scenarios. In case of full traffic data extraction including all roads in urban core scenes with several thousands of vehicles per scene, a slight lack of performance in the traffic data algorithms is present. In order to close this gap, two measures are taken. Firstly, Adaboost classifier will be cascaded. Secondly,

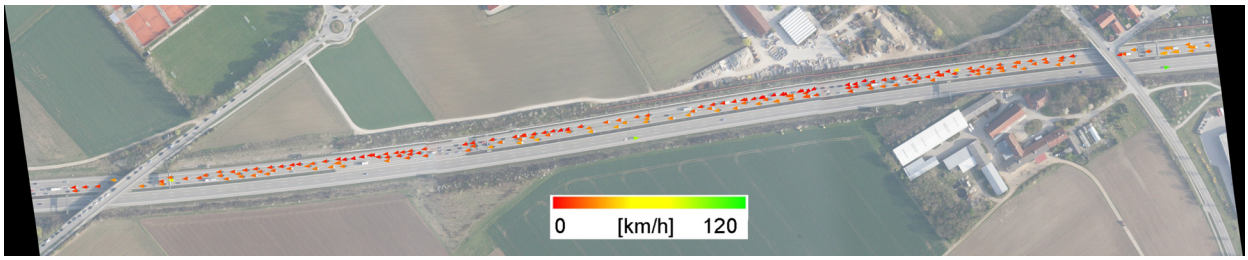

Figure 6: Traffic data extracted automatically by the processing chain on BAUMA campaign 2010 near exit "Munich exhibition center" of motorway A 94. Quality of traffic data is 80 % in case of that campaign. Periodic gaps in traffic data are due to missing overlap between image bursts (which can be corrected easily by flight planning in future campaigns).

car tracking algorithm by template matching will be implemented as extreme parallelization on actual high-performance graphics adapters. In spite of scope for increasing performance we can conclude that our system has reached an operational state and is ready for application in crisis and events.

References

Behrisch, M., Bonert, M., Brockfeld, E., Krajzewicz, D. and Wagner, P., 2008. Event traffic forecast for metropolitan areas based on microscopic simulation. In: Third International Symposium of Transport Simulation 2008 (ISTS08), Queensland, Australia.

Bethke, K.-H., Baumgartner, S. and Gabele, M., 2007. Airborne road traffic monitoring with radar. In: World Congress on Intelligent Transport Systems (ITS), Beijing, China, pp 16.

Börner, A., Ernst, I., Ruhe, M., Sujew, S. and Hetscher, M., 2004. Airborne camera experiments for traffic monitoring. In: XXth ISPRS Congress, 1223 July 2004, Vol. XXXV, Part B, 6 p.

Busch, F., Glas, F. and Bermann, E., 2004. Dispositionssysteme als fcdquellen fr eine verbesserte verkehrs-lagerekonstruktion in stdten-eine berblick. Straen-verkehrstechnik.

Ernst, I., Sujew, S., Thiessenhusen, K.-U., Hetscher, M., Rassmann, S. and Ruhe, M., 2003. Lumos - airborne traffic monitoring system. In: Proc. IEEE Intelligent Transportation Systems, Vol. 1, pp. 753–759.

Freund, Y. and Schapire, R. E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting,. Journal of Computer and System Sciences 55(1), pp. 119–139.

Friedman, J., Hastie, T. and Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. Annals of Satistics 28(28), pp. 337–407.

Grabner, H., Nguyen, T., Gruber, B. and Bischof, H., 2008. On-line boosting-based car detection from aerial images. 63(3), pp. 382–396.

Kurz, F., Müller, R., Stephani, M., Reinartz, P. and Schroeder, M., 2007. Calibration of a wide-angle digital camera system for near real time scenarios. In: ISPRS Hannover Workshop 2007, High Resolution Earth Imaging for Geospatial Information, Hannover, 2007-05-29-2007-06-01, ISSN 1682-1777.

Lenhart, D., Hinz, S., Leitloff, J. and Stilla, U., 2008. Automatic traffic monitoring based on aerial image sequences. Pattern Recognition and Image Analysis 18(3), pp. 400–405.

Leonhardt, A., 2004. Detektortest siemens ld4. Technical report, TU München http://www.siemens.nl/ITS/getfile.asp?id=52.

Lepetit, V. and Fua, P., 2006. Keypoint recognition using randomized trees. Pattern Analysis and Machine Intelligence, IEEE Transactions on 28(9), pp. 1465–1479.

Lienhart, R., Kuranov, A. and Pisarevsky, V., 2003. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. Pattern Recognition 2781, pp. 297–304.

Müller, R., Lehner, M., Müller, R., Reinartz, P., Schroeder, M. and Vollmer, B., 2002. A program for direct georeferencing of airborne and spaceborne line scanner images.

Nguyen, T. T., Grabner, H., Bischof, H. and Gruber, B., 2007. On-line boosting for car detection from aerial images. In: Proc. IEEE Int Research, Innovation and Vision for the Future Conf, pp. 87–95.

Reinartz, P., Lachaise, M., Schmeer, E., Krauss, T. and Runge, H., 2006. Traffic monitoring with serial images from airborne cameras. 61(3-4), pp. 149–158.

Rosenbaum, D., Charmette, B., Kurz, F., Suri, S., Thomas, U. and Reinartz, P., 2008. Automatic traffic monitoring from an airborne wide angle camera system. In: ISPRS08, p. B3b: 557 ff.

Schaefer, R.-P., Thiessenhusen, K.-U. and Wagner, P., 2002. A traffic information system by means of real-time floating-car data. In: Proceedings of ITS World Congress, October 2002, Chicago, USA.

Schapire, R. E. and Singer, Y., 1999. Improved boosting algorithms using confidence-rated predictions. Machine Learning 37(3), pp. 297–336.

Suchandt, S., Eineder, M., Breit, H. and Runge, H., 2006. Analysis of ground moving objects using srtm/x-sar data. 61(3-4), pp. 209–224.

Viola, P. and Jones, M., 2004. Robust real-time face detection. International Journal of Computer Vision 57, pp. 137–154.