AUTOMATIC TRAFFIC MONITORING WITH AN AIRBORNE WIDE-ANGLE DIGITAL CAMERA SYSTEM FOR ESTIMATION OF TRAVEL TIMES

F. Kurz[a*], B. Charmette[a], S. Suri[a], D. Rosenbaum[a], M. Spangler[b], A. Leonhardt[b], M. Bachleitner[c], R. Stätter[a], P. Reinartz[a]

[a]German Aerospace Center (DLR), Remote Sensing Technology Institute, PO Box 1116, D-82230 Weßling, Germany
[b]Munich University of Technology (TUM), Traffic Engineering and Control, Arcisstr. 21, D-80290 München, Germany
[c]General German Automobile Association (ADAC), Developement of traffic information, Am Westpark 8, D-81373 München, Germany
franz.kurz@dlr.de

Commission 3 WG III/5

KEY WORDS: Aerial cameras, image series, traffic parameters, road detection, vehicle detection

ABSTRACT:
Knowledge of accurate travel times between various origins and destinations is a valuable information for daily commuters as well as for security related organizations (BOS) during emergencies, disasters, or big events. In this paper, we present a method for automatic estimation of travel times based on image series acquired from the recently developed optical wide angle frame sensor system (3K = "3-Kopf"), which consists of three non-metric off-the-shelf cameras (Canon EOS 1Ds Mark II, 16 MPixel). For the calculation of overall travel times, we sum up averaged travel times derived from individual vehicle velocities to pass defined road segments. The vehicle velocities are derived from vehicle positions in two consecutive geocoded images by calculating its distance covered over time elapsed. In this context, we present an automatic image analysis method to derive vehicle positions and vehicle distances involving knowledge based road detection algorithm followed by vehicle detection and vehicle tracking algorithms. For road detection, we combine an edge detector based on Deriche filters with information from a road database. The extracted edges combined with the road database information have been used for road surface masking. Within these masked segments, we extract vehicle edges to obtain small vehicle shapes and we select those lying on the road. For the vehicle tracking, we consider the detected vehicle positions and the movement direction from the road database which leads to many possible matching pairs on consecutive images. To find correct vehicle pairs, a matching in the frequency domain (phase correlation) is used and those pairs with the highest correlation are accepted. For the validation of the proposed methods, a flight and ground truth campaign along a *16 km* motorway segment in the south of Munich was conducted in September *2006* during rush hour.

1. INTRODUCTION

Near real time monitoring of natural disasters, mass events, and large traffic disasters with airborne SAR and optical sensors will be the focus of several projects in research and development at the German Aerospace Center (DLR) in the next years. In this overall frame, knowledge of accurate travel times between various origins and destinations is a valuable information not only for traffic management and information purposes but also for security related organizations (BOS) during emergencies, disasters, or big events. One application currently under development is the generation of isochronal maps derived from airborne imagery, which shows the up-to-date travel times from each map position to the accident scene. This may support BOS coordinators for fleet disposition and adequate alerting in times of adversities.

In general, travel time estimation is based on data from conventional stationary measurement systems such as inductive loops, radar sensors or terrestrial cameras. One handicap of these methods is the low spatial resolution depending on the ground distribution. New approaches include data by means of mobile measurement units which flow with the traffic (floating car data, FCD, (Schaefer et al., 2002), (Busch et al. 2004)).

The big advantage of the remote sensing techniques presented here is that the measurements can be applied nearly everywhere (exception: tunnel segments) and do not depend on any third party infrastructure. Besides, airborne imagery provides a high spatial resolution combined with acceptable temporal resolution depending on the flight repetition rate, but require complex image analysis methods and traffic models to derive the desired traffic parameters (Ernst, 2005).

In this paper, we present a method for automatic estimation of travel times based on geocoded image series acquired from a recently developed DLR wide angle frame sensor system. For the calculation of the point-to-point travel times, we apply a simple semi-empirical travel time method which is based on averaged travel times derived from individual vehicle velocities to pass defined road segments. For this, individual vehicle positions and velocities are required which are provided by automatic image analysis tools.

In this context, we present an automatic image analysis method to derive vehicle positions and vehicle distance covered based on a road detection algorithm using a road database followed by vehicle detection, and vehicle tracking algorithms.

2. ESTIMATION OF TRAVEL TIMES

While only few detection methods offer the possibility to directly measure travel times of single cars (such as vehicle re-identification by licence plate recognition, floating car tracking), certain model assumptions are needed to derive travel times from local detector data.

Different methods for the estimation of travel times are applied depending on the sensors used and on the type and availability of sensor data. Simple methods derive travel times by constructing pseudo vehicle trajectories from time mean speed measured locally. Those methods are easy to implement and need no calibration but have difficulties to capture traffic

dynamics such as traffic jams at high densities. Better results can be generated by applying traffic flow models that are widely used in traffic engineering and control. For simulation or traffic prediction, data from inductive loops, terrestrial cameras, variable speed limits and other context data are fed into these traffic models.

Macroscopic traffic models like METANET (Kotsialos, 2002) reproduce the spatial and temporal traffic flow based on averaged microscopic car dynamics and are used to estimate or predict traffic state information. A large group of macroscopic models is based on the fluid-dynamic theory, interpreting the traffic flow as a compressible fluid (e.g. Daganzo, 1997).

Microscopic traffic models (e.g. Krauß, 1997) describe the vehicles with detailed behaviour. Each vehicle's reactions to the actual traffic situations are modelled by assumptions concerning car following (e.g. Gipps, 1981) and overtaking procedures (e.g. Gipps, 1986). Microscopic models are more detailed than macroscopic models, but they require significantly higher computational efforts, what limits their use for real-time applications. A class of microscopic traffic models based cellular automat technology tries to overcome those problems (e.g. Nagel, 1992).

New approaches aim to fuse all relevant data sources such as floating car data, stationary sensor data, modelled data and historic information..

In case of airborne optical data, travel times for a platoon of vehicles can be recorded from a helicopter, rather than a single car. The helicopter simply follows the vehicles in a traffic jam (Angel, 2003).

In case of wide-angle airborne data, like the DLR 3K camera, we propose an instantaneous method to derive overall travel times T for a road section. For this, we divide the road section in N equally spaced segments with length L and calculate mean vehicle velocities $\overline{v}_i$ for segment i. The overall travel time T is then the sum of travel times for each segment.

$$T = \sum_{i=1}^{N} \min\left\{\frac{L}{\overline{v}_i}, \frac{L}{v_{\min}}\right\} \quad (1)$$

In case of very low mean vehicle velocities, the travel times tend to be very high (standing cars would lead to an infinite value). Thus, a minimum vehicle velocity v_{min} is introduced to avoid too high travel times. The minimum vehicle velocity parameter is empirical and must be adjusted to the traffic type as well as to the road segment monitored as e.g. in case of blocked motorway lanes the estimated travel times will be erroneous using a constant minimum vehicle velocity for all traffic types.

In this paper, we focus on congested traffic with traffic densities higher than the critical density without additional disturbances like blocked lanes and we set *a priori* the minimum vehicle velocity to *7.2 km/h* following the METANET settings.

The mean vehicle velocities $\overline{v}_i$ will be derived from the 3K image sequences based on the vehicle positions in at least two consecutive geocoded images and based on the time elapsed (Hinz, 2007).

In the following chapter, automatic image analysis tools to derive vehicle positions and velocities will be presented.

3. AUTOMATIC ROAD AND VEHICLE DETECTION

For an effective real time traffic analysis, the road surface needs to be clearly demarcated. Thus, we automatically delineate the roadsides by two linear features using the processing chain illustrated in Fig 1. Automated road extraction has been developed as an independent module in our ongoing research on real time traffic analysis. For brevity, we are not explaining the road extraction procedure in details. The road extraction starts by forming a buffer zone around the roads surfaces using a road database as basis for the buffer formation process. In the marked buffer zone, we use edge detection and feature extraction techniques. The critical step of edge detection is based on an edge detector proposed by Phillipe Paillau for noisy SAR images (Paillou, 1997). Derived from Deriche filter (Deriche, 1989) and proposed for noisy SAR images, we found this edge detector after ISEF filtering (Shen and Caston, 1992) extremely efficient for our purpose of finding edges along the roadsides and suppressing any other kind of surplus edges and noise present. The roadside identification module, again with the help of the road database tries to correct possible errors (gaps and bumps) that might have creped in during the feature extraction phase.

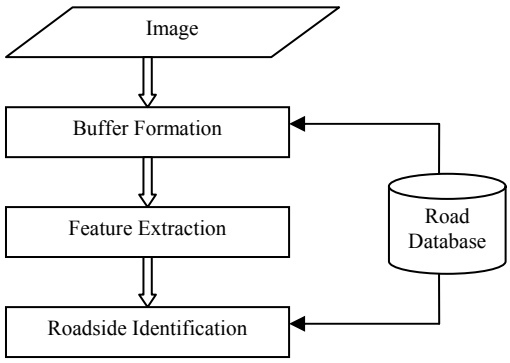

Fig 1 Implemented processing chain for a knowledge based road extraction

With the information of the roadside obtained in the processing step described before, it is possible to perform vehicle detections and tracking limited only to the roads. For this, we developed an algorithm for the detection and tracking of vehicles which is described in the following (see Fig 2).

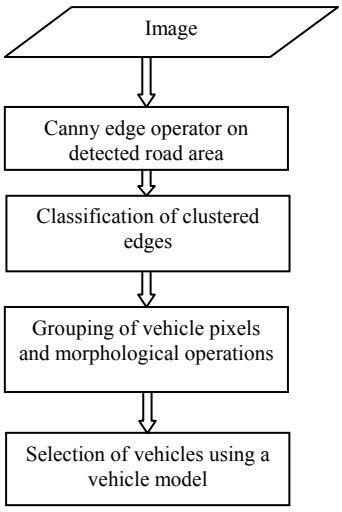

Fig 2 Vehicle detection algorithm

Based on information about the alignment and direction of the roadside, all pixels of the road including the road direction are marked. For the vehicle detection, a Canny edge operator 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 the three main classes, vehicles, roads, and not classifiable.

We consider the part with the lowest steepness being mainly pixels of the road background, since its intensity is quite uniform. Besides, we assume that the part with the highest steepness is due to the high discontinuity in the intensity probably populated by vehicles. This part of the statistic is also contaminated by the paintings on the roadside, shadows, sign boards, trees, etc.

In the part which is not classifiable, it must be determined which pixels belong to the road background or to potential vehicles. For the decision, the pixel neighbourhood is examined. Pixels directly connected with a potential vehicle pixel are moved into the vehicle class. Remaining pixels are finally considered as road background.

In case of coloured images, the processes described above are realized on each channel separately in order to obtain a multi channel edge image.

In the next step, the roadside pixels are eliminated from the part with higher steepness. As the roads are well determined by the road extraction, the roadside lines can be found easily. Thus, the algorithm erases all pixels with high edge steepness which are laying on a roadside position. Thereby, it avoids erasing vehicles on the roadside by considering the width of the shape. Mid-line markings are erased using a dynamic threshold detector. This is done in order to reduce false detections, since these mid-line markings may look like white cars.

Then, potential vehicle pixels are grouped by selecting neighboured pixels. Each group is considered to be composed of potential vehicle pixels connected to each other. With the groups obtained a list of potential vehicles is produced.

In order to extract real vehicles from the potential vehicle list, a closing of the shapes of the potential vehicles is performed. We can see the effect of the closing algorithm in Fig 3.

Potential car pixels detected before closing — Potential car pixels after closing

Fig 3. Closing the shapes of potential car pixels

Using the closed shape, the properties of vehicle shapes are described by its direction, its area, the length and width following the direction, and its position on the road. Based on these parameters, the vehicles are assumed to have rectangular shapes with a specific length and width oriented in the road direction. Their area should be about the length multiplied by the width and vehicles must be located on the roads. We set the values for the vehicle length to *5.7 m* and for the width to *2.6 m (for standard cars)*. In case of 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.

By applying vehicle detections to an image sequence taken within a small time interval, which is the case for the 3K image series, we are able to match a vehicle in this sequence, in order to measure its velocity.

Within the tracking algorithm (see Fig 4), a search area for each vehicle detected in the first image is defined in the second image based on the predicted position. During short time steps between two images, vehicles do not change their direction significantly, so it can be assumed that the matching vehicle from the other image lies within the search area. Thus, we can find potential corresponding pairs of detected vehicles. However, in case of a high vehicle density on the images, as it is the situation in traffic congestions, many vehicles in the second image are located in the search area of one vehicle in the first image. At the same time each car in the second image has lots of possible origin cars in the first image. In order to choose the right pair, it is checked that the direction of both vehicles in a possible vehicle pair is the same. As criteria, the road direction from the road database is considered for the selection correct pairs.

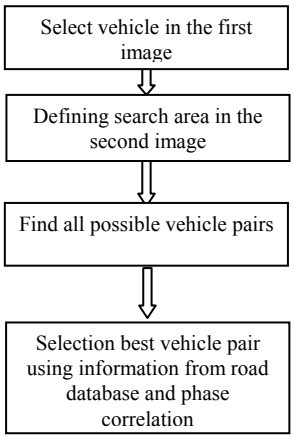

Fig 4 Car tracking algorithm

Then, a phase correlation algorithm to find best corresponding vehicles is applied for each pair by using small image patches around the car. With these tests a quality coefficient is determined for each pair of matching candidates. At the end, when a car in the first image has several correspondents in the second image, we keep only the pair with the best quality value.

4. SENSORS AND DATABASE

4.1 3K camera system

The 3K camera system (3K = "3Kopf") consists of three non-metric off-the-shelf cameras (Canon EOS 1Ds Mark II, 16 MPix). The cameras are arranged in a mount with one camera looking in nadir direction and two in oblique sideward direction (Fig 5), which leads to an increased FOV of max *110°/31°* in across track/flight direction.

Fig 5. DLR 3K-camera system consisting of three Canon EOS 1Ds Mark II, integrated in a ZEISS aerial camera mount

The camera system is coupled to a GPS/IMU navigation system, which enables the direct georeferencing of the 3K optical images. Fig 6 illustrated 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 coverage, and pixel size is shown, e.g. the pixel size at a flight height of *1000 m* above ground is *15 cm* and the image array covers up *2.8km* in width.

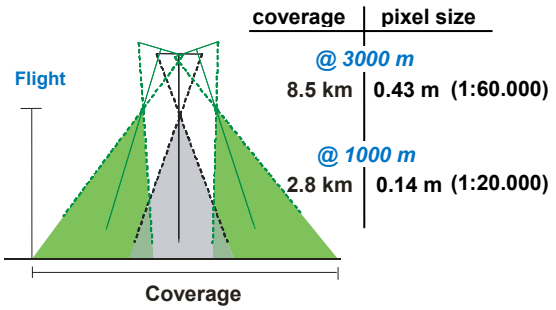

Fig 6. Illustration of the image acquisition geometry. The tilt angle of the sideward looking cameras is approx. 35°.

4.2 Test-Site

The motorway A8 south of Munich is one of the busiest parts of the German motorway network with an average load of around 100.000 vehicles per day. Fig 7 shows the *16 km* motorway section between motorway junctions "Hofolding" and "Weyarn", which was selected as test site on 2. Sep. 2006. At this time, heavy traffic was expected at this section caused by homebound travellers.

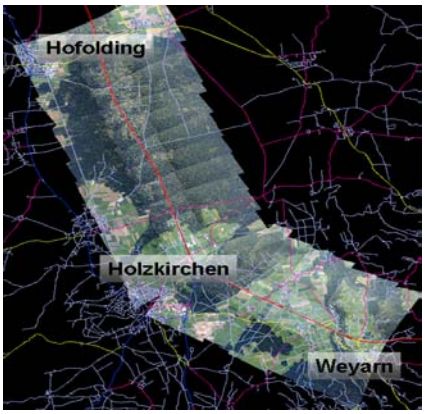

Fig 7. 16km motorway strip (A8) south of Munich as imaged by 3K camera system

4.3 3K imagery

Three 3K data takes were acquired on *2. Sep. 2006* between *14:01* and *15:11* from *2000m* above ground. Table 1 lists the exact acquisition times of each data take. During each overflight, 22 image bursts were acquired each containing four consecutive images. The time difference within these bursts was *0.7 s*, so that each car was monitored at least *2.1 s*.

ID	Date	Pixel size	Images	H.a.G.
3K-Ia	02-Sep-2006 14:01-14:12	30cm	22x4x3	2000m
3K-Ib	02-Sep-2006 14:30-14:40	30cm	22x4x3	2000m
3K-Ic	02-Sep-2006 15:01-15:11	30cm	22x4x3	2000m

Table 1 3K camera data takes at A8 south of Munich

For further analysis, 3K images were geocoded using onboard GPS/IMU measurements with an absolute position error of 3m in nadir images and less than one pixel relative. The last error has great influence on the derived vehicle velocities.

4.4 Road database and ancillary data

Data from a road database are 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 (NAVTEQ 2006). 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 5m accuracies for motorways.
Data from other sources and sensors were collected to make an overall comparison of derived travel times. Although traffic messages from the traffic channel do not contain travel times but congestion lengths, this information represents the state of the art and was therefore collected. Additionally, data from local detectors are used to calculate instantaneous travel times:

$$TT(t) = \sum_{i=1}^{n} \frac{l_i}{v_i(t)} \qquad (2)$$

With *TT(t)* being the travel time for the whole stretch at time instant t, v_i being the speed reported at detector station *i* and l_i being the length of the segment assigned to detector station *i*.

ID	Time	Traffic message channel	Travel times from detector data
3K-Ia	-14:06	Halting traffic 14km	34 min
	-14:11	Congestion 7km	
3K-Ib	-14:40	Congestion 12km	28 min
3K-Ic	-15:03	Halting traffic 18km	
	-15:07	Congestion 7km	26 min
	-15:11	Congestion 12km	

Table 2 Traffic message channel information and estimated travel times from detector data for A8 south of Munich

Table 2 lists data from other sources and sensors, which are linked to the acquired 3K data takes. Obviously, the traffic messages vary strongly within short time periods between congested and halting traffic.

4.5 Reference vehicle

As ground truth for travel times, two runs with a GPS equipped vehicle (ADAC) in northbound direction and one run in southbound direction was conducted. Table 3 lists the links between reference vehicle runs to 3K data takes. It should be mentioned, that a direct comparison between travel times from reference vehicle and 3K data takes contain systematic errors, as the northbound runs take around half an hour and the 3K data set represent a time span of ten minutes.

ID	Date	Strip	Travel-time	Data take	
R1	02-Sep-2006 14:02-14:36	16km	35 min	3K-Ia	North-bound
R2	02-Sep-2006 14:38-14:47	16km	9 min	3K-Ib	South-bound
R3	02-Sep-2006 14:48-15:18	16km	31 min	3K-Ic	North-bound

Table 3 Data from reference vehicle at A8 south of Munich

4.6 Manual measurements

For manual measurements, the 3K images have been mosaiked in two orthophotos according to Fig 7 (taking the 2. and 4. image of the bursts). This mosaic represents the traffic situation with a relative time difference of *1.4s*. Measurement of vehicle positions and corresponding vehicles resulted in the velocity of vehicles. These measurements were basis for calculating the travel times which are illustrated in Table 4.

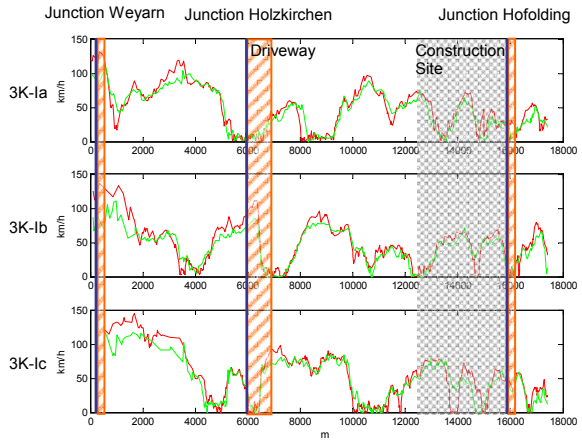

Fig 8. Manual measured northbound velocity profiles on the left lane (red) and right lane (green) for all 3K datatakes

Fig 8 shows velocity profiles of all vehicles going northbound during the three 3K data takes. The vehicle velocities vary strongly between *150 km/h* and *0 km/h* with small differences between left and right lane.

5. RESULTS

5.1 Comparison of travel times

Travel times are calculated using measured vehicle positions and velocities based on equation (1). For this, the minimum vehicle velocity was set to *7.2 km/h* according to the METANET settings, and average vehicle velocities $\overline{v}_i$ were calculated based on segments *L* of *1 km* length. Table 4 lists the travel times derived from the 3K data sets separated in left and right lane, the travel times from the detector data, and the travel times of the reference vehicle.

In general, left lane travel times are slightly shorter than right lane travel times. In comparison with the reference travel times, the 3K derived times are higher and the detector derived times are shorter. Besides, the 3K times decrease with higher minimum vehicle velocity or longer road segments. Thus, with a minimum vehicle velocity of *15 km/h* or a road segment length of around *10 km*, the derived travel times correspond with the reference times. These two parameters are free and must be adjusted to the road type and traffic type, but in this experiment they were fixed to the METANET settings.

ID	Ref	Travel time 3K		Detector data	
		Left lane	Right lane		
Ia	35’	39’48’’	39’57’’	34’	S-N
Ib	X	38’36’’	40’42’’	28’	S-N
Ib	9’	08’04’’	X	X	N-S
Ic	31’	37’06’’	37’27’’	26’	S-N

Table 4 Comparison of travel times derived from 3K images with reference travel times

Instantaneous travel times derived from detector data do not completely comply with the reference measurements. That is because only speeds at certain stations are used that are more or less a set of random observations of the complete freeway stretch. Also, instantaneous speeds do not reflect the dynamics of traffic, as they do not consider the time a vehicle actually needs to pass the different stations.

5.2 Automatic detection of vehicles

Fig 9. Zoom into the two images used for tests on the automatic vehicle detection. Rectangles mark automatic detections, triangles point into the travel direction.

We tested our programs for automatic road and vehicle detection on several image sequences of traffic on motorways taken by the 3K camera system. Fig 9 shows a zoom into two example images which were taken from a sequence obtained at

the motorway A8 near "Holzkirchen". The lower image was taken *1.4s* later than the upper image. Using these images the programs for automatic road and vehicle detection were tested (see also section 3), and the results of the automatic detection were compared to manual measured reference.
Table 5 shows the results of the automatic and manual detection of cars on these two images as well as the completeness and correctness of the automatic car detection with respect to the reference. It can be seen that the correctness is quite high, whereas the completeness is lower. We used these results of the automatic vehicle detection as input for the vehicle tracking program. However, the results in correctness and completeness of corresponding pairs of vehicles in these two images obtained from the vehicle tracking program are quite low due to the low completeness of the automatic vehicle detection.

ID	Total automatic detections	Correct automatic detections	Manual detections	Correctness of automatic detection	Completeness of automatic detection
Ia	117	95	193	81%	49%
Ib	56	47	170	84%	28%

Table 5 Comparison of automatic and manual car detection for two images

6. CONCLUSIONS

The investigations show the high potential to use airborne image time series for the estimation of travel times. Firstly it is shown that the automatic road detection works sufficiently good for restricting the search area of the vehicle detection and for giving information about road directions. Also the automatic vehicle detection has already reached a high level of accuracy concerning the correctness, although the completeness is still too low for car tracking applications and has to be improved. A further challenge of the methodology is a better estimation of the minimum velocity for modelling the vehicle speed in a traffic congestion. Models have to be established to estimate this parameter out of vehicle density and/or other measurable values. It is expected that overall accuracy could be enhanced if the system would be combined with a dedicated traffic flow model in order to utilize the strength of both approaches.
The results of such measurements can be used for obtaining travel times and other relevant traffic parameters in cases of catastrophes or other special events (e.g. mass events), where the costs for an airborne data acquisition is justified.
It is planned to test the airborne based camera travel time measurement system against a real reference database with high sampling rate, e.g. vehicle re-identification via automatic licence plate recognition.

7. REFERENCES

Angel, A., Hickman, M., Mirchandani, P., Chandnani, D. (2003). *Methods of Analyzing Traffic Imagery Collected From Aerial Platforms.* IEEE Transactions on Intelligent Transportation Systems, Vol 4 No 2 pp 99 – 107

Busch, F.; Glas, F.; Bermann, E. (2004). *Dispositionssysteme als FCD-Quellen für eine verbesserte Verkehrs-lagerekonstruktion in Städten - eine Überblick.* Straßen-verkehrstechnik 09/04

Daganzo, F. C. (1997). *Fundamentals of Transportation and Traffic Operations.* Elsevier Science Inc., New York.

Deriche, R. (1987). *Using Canny's criteria to derive an optimal edge detector recursively implemented.* The International Journal on Computer Vision, Vol. 1, No. 2, pp 167-187

Ernst, I., Hetscher, M., Zuev, S., Thiessenhusen, K.U., Ruhé, M. (2005). *New approaches for real time traffic data acquisition with airborne systems.* In: TRB Transportation Research Board, TRB 2005, Washington.

Gipps, P.G. (1981). *A behavioural car-following model for computer simulation.* Transportation Research Part B, 20, 403-414.

Gipps, P.G. (1986). A model for the structure of lane-changing decisions. Transportation Research Part B, 15, 105-111.

Hinz, S., Kurz, F., Weihing, D., Meyer, F., Bamler, R. (2007). *Evaluation of Traffic Monitoring based on Spatio-Temporal Co-Registration of SAR Data and Optical Image Sequences.* PFG – Photogrammetry – Fernerkundung – Geoinformation, in print.

Kotsialos, A., Papgeorgiou M., Diakaki, C., Pavlis, Y., Moddelham, F. (2002). *Traffic Flow Modeling of Large-Scale Motorway Networks Using the Macroscopic Modeling Tool METANET.* IEEE Transactions on Intelligent Transportation Systems. Vol 3 No 4, pp 282–292

Krauß, S., Wagner, P., Gawron, C. (1997). *Metastable states in a microscopic model of traffic flow.* Physical Review E, 5, pp 5597 - 5602

Kurz, F., Müller, R., Stephani, M., Reinartz, P., Schroeder, M. (2007). *Calibration of a wide-angle digital camera system for near real time scenarios.* In: Heipke, C.; Jacobsen, K.; Gerke, M. [Hrsg.]: ISPRS Hannover Workshop 2007, High Resolution Earth Imaging for Geospatial Information, Hannover, 2007-05-29 - 2007-06-01, ISSN 1682-1777

Nagel, K.; Schreckenberg, M. (1992). *A cellular automaton model for freeway traffic.* J. Phys. I France 2 (1992) pp 2221-2229.

Paillau, P. (1997). *Detecting Step Edges in Noisy SAR Images: A New Linear Operator.* IEEE Transactions on Geoscience and Remote Sensing, Vol. 35, No.1, pp 191-196

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.

Shen, J., and Castan, S. (1992). *An optimal linear operator for step edge detection.* CVGIP, Graphics Models and Image Processing, Vol. 54, No. 2, pp 112-133