TRAFFIC MONITORING WITH SERIAL IMAGES FROM AIRBORNE CAMERAS

P. Reinartz[a*], T. Krauss[a], M. Pötzsch[a], H. Runge[a], S. Zuev[b]

German Aerospace Center (DLR), [a] Remote Sensing Technology Institute, PO Box 1116, D-82230 Weßling, Germany
[b] Institute of Transport Research, Rutherfordstr. 2, D-12489 Berlin
peter.reinartz@dlr.de

Commission I, WG I/4

KEY WORDS: Aerial images, image series, traffic monitoring, image matching, velocity estimation, GIS

ABSTRACT

The classical ways to monitor traffic density and velocity depend on local measurements from induction loops and other on site instruments. This information does not give the whole picture. In order to obtain precise knowledge about the traffic flow of a large area, only airborne cameras or cameras positioned at very high locations (towers, balloons, etc.) can give an up-to-date image of all streets covered. To be able to determine precise velocities and other parameters from an image time series, exact georeferencing is one of the first requirements for the acquired image data. The method presented here for determining several traffic parameters for single vehicles and vehicle groups involves recording and evaluating a number of digital or analog aerial images from high altitude and with a large total field of view. Investigations involving visual and automatic interpretation of the images are described. The recording frequency of the individual images should be at least 1/3 Hz, but is preferably 2 Hz or more, especially for automatic tracking. Since the moving vehicles are at different locations in the subsequent images, it is possible to derive their velocity, vehicle type as well as acceleration (given at least three images). This allows capture of the immediate traffic dynamics for the recording area in considerable detail over a large region. The accuracy and possibilities of the method are analyzed and presented, as well as the integration of the results into a GIS concept for traffic monitoring, using a street data base. The traffic data obtained from induction loops or other local instruments can be checked with the help of this procedure and local traffic anomalies can be precisely identified and studied. The paper shows how the image data are acquired processed and evaluated and how this methodology can be used for traffic monitoring.

1. INTRODUCTION

Traffic research and planning requires a vast quantity of detailed information about traffic behavior. Empirical studies, measurements, modeling attempts and simulation programs are accordingly numerous and diverse. The challenge is to develop robust methods for predicting, visualizing and modeling complex traffic events (Brockfeld et al. 2002). Only on the basis of reliable data can the simulation, control and planning of traffic systems be optimized and thereby contribute to increase the capacity of the relevant infrastructure, reducing emissions, and increasing safety.

Merely equipping the streets with conventional stationary measurement systems such as induction loops, radar sensors or cameras will not provide an adequate supply of suitable data. The challenge is to develop innovative solutions which augment existing individual measurement sites, thereby closing the gaps in the traffic picture. Current traffic researchers are counting on so-called large area data collection to achieve a considerably improved data basis. In addition to new approaches currently under discussion for recording data by means of mobile measurements units which flow with the traffic (floating car data, Schaefer et al 2004), remote sensing measurement techniques have a high potential for large area traffic data collection.

Especially in areas beyond city and town limits it is difficult to obtain a complete picture of the traffic over a wide area. On German freeways and national roads numerous accident-prone sites have been identified for which the precise cause of accidents is not well known or adequately documented. It is also important to know how traffic jams develop and how vehicle velocity reduces at the tail end of a jam or response to speed limits. Stationary camera systems relatively near the ground are not useful for rigorous observation, since detailed analysis of the traffic situation is only possible within a small area of about 100 meters. For larger areas, relevant parameters such as vehicle type, distance between vehicles and velocities cannot be recorded because of significant shadowing effects.

The advantage of satellites or aircraft is that they can record a large area at once. However, because of their orbits, satellites can only provide individual images of a particular area. A return flight over the same area is only possible after the lapse of several days, the actual interval depending on the particular orbit. Continuous observation of traffic in a particular location from a geostationary orbit at 36,000 km altitude is not possible because the spatial resolution which can be achieved from that altitude is too low. If traffic is to be analyzed with the help of satellites it is necessary to record the situation on typical days and times and independent of the weather. With the upcoming German radar satellite TerraSAR-X it will be possible for the first time to record traffic situations independently of weather and daylight (Runge et al. 2004).

The usefulness of aircraft optical data both from the visible (Stilla et al. 2004, Toth et al. 2004) and thermal infrared (Ernst et al. 2003, Hinz 2004) ranges for vehicle detection has been studied using many different approaches. However, there are not many investigations of optical time series recorded by aircraft or from very high positions and with a large field of view (Mirchandani et al. 2002, Pötzsch 2005). Large area successive imaging at intervals of seconds is, however, possible with airborne sensors. Parameters relevant for traffic such as individual vehicle and

* Corresponding author

vehicle group velocities, vehicle type, distance and traffic density can be derived from such data.

With a time series acquired over a long time period it is possible to record the entire traffic history for a given area and to analyze, for example, overtaking maneuvers, merge and exit behavior, as well as traffic jams dynamically with actual data. Such results are highly relevant input data for traffic modeling programs, for testing the efficacy of traffic control measures and for the input into GIS systems for traffic monitoring (Ernst et al. 2005). The paper shows how these data are acquired and evaluated.

2. AERIAL IMAGE DATA AND GROUND REFERENCE

The analog aerial images were recorded on October 5, 2004 by a ZEISS RMK A30/23 mounted on a DLR aircraft. The regions covered are along the A99 and A9 freeways north of Munich (see fig. 1) and another part of the A9 freeway south of Nürnberg. The time difference between two consecutive images is about 2.9 sec. The A99 area is covered by 37 images and the A9 area north of Munich by 43. The images were acquired from an altitude of 2,500 m and scanned to a pixel resolution of 0.16 m.

The reference data sets for these image data are as follows:

Figure 1. Test sites north of Munich

- Data from overhead radar sensors recording velocities of single cars at four locations with a nominal accuracy of 3 km/h
- Two cars with DGPS tracking, reaching an accuracy of about 3 km/h standard deviation
- Orthoimages from the Bavarian Land Surveying Office (point location accuracy given: 1-2 m)
- Navteq data set containing information on streets and their properties like location, velocity limitation, etc.

A second data set was generated by a conventional digital frame camera, which acquired images from the Munich "Olympic Tower" at an altitude of about 200 meters above a main city freeway. In this case the image repetition rate was on the order of 0.5 sec but was not determined very accurately. This data set was used without reference, to test new methods of automatic car detection and tracking.

3. PROCESSING OF THE AERIAL IMAGE DATA

The geometric resolution was first reduced to a pixel size of 50 cm x 50 cm, since this resolution is high enough to measure single cars and the smaller digital images are much easier to handle.

3.1 Image matching and orientation

Since the inertial navigation system (INS) for the analog camera was not working nominally, the exterior orientation of the images had to be achieved using ground control points (GCP). To minimize manual interaction, a digital image matching of several consecutive images (~12) was performed first (Lehner et al. 1992). This method allows automatic matching and mosaicking with subpixel accuracy for images which contain large overlapping areas (in this case the overlap is 88% in flight direction). All images treated with this method exhibit very similar geometries. Due to the change of parallaxes this method is not exact but quite accurate for streets with low slope. The absolute orientation is calculated for this image set using one set of GCP from a high resolution orthoimage, and the single images are registered accordingly. The accuracy achieved by this method depends first on the orthoimage accuracy (GCP), which is given to 1-2 m. The standard deviations of the residual distortions at the GCP after transformation with a second order polynomial are shown in table 1, which is in the same order.

Freeway A99	σ_0 in X-Direction [m]	σ_0 in Y-Direction [m]
Part_1, 12 images	1.15	1.63
Part_2, 12 images	1.30	1.25
Part_3, 12 images	1.05	0.72

Table 1. Standard deviation for the residual deviations for every 12-image combination

Both DGPS tracking vehicles appear in five of the images showing the Nürnberg area. These images were also georeferenced using the method described above. Then the positions of the vehicles were manually measured. A comparison of these positions with the position calculated via DGPS showed a bias of 0.9 m in the images and a standard deviation of 1.2 m. This demonstrates that the orientation accuracy of the aerial images is of the same order of magnitude as that of the orthoimages: 1-2 m.

The digital frame camera data were similarly matched and show a very high level of overlay accuracy, as will be shown in section 4.

3.2 Velocity derivation and comparison

After the rectification process, the position of the vehicles can be measured in each image. Due to the relatively long time between two images (~2.9 sec) automatic tracking is difficult since a car with a velocity of about 100 km/h moves about 80 m in this time period. Since a dense traffic flow implies car distances of about 1 – 2 sec distance, automatic procedures like those mentioned in (Ernst et al. 2003) are not applicable. Therefore, manual measurements of vehicle positions were performed, which led to 100% detection accuracy. Figure 2 shows an example of the visualizer tool developed for this purpose, which allows a manual/visual easy and fast tracking of single vehicles.

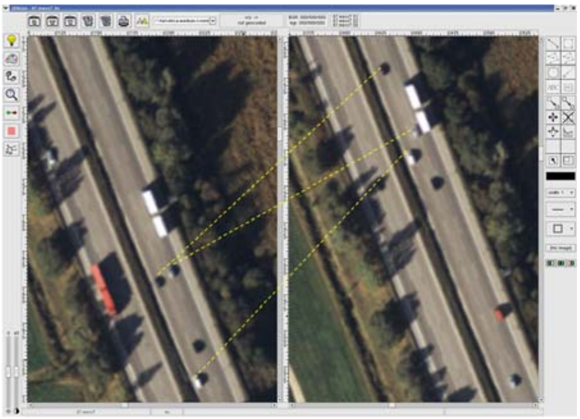

Figure 2. XDibias visualizer to measure vehicles in subsequent images

A very sensitive parameter in the velocity measuring process is the exact time between the two data acquisitions. An error of 0.1 sec implies a distance error of 3 m when traveling at a speed of 30 m/sec. The image acquisition time was estimated therefore to an accuracy of about 0.02 sec (Pötzsch 2005). The velocity estimation is therefore mainly dependent on the local accuracy of the rectified images, which is on the order of 1-2 m (and implies a velocity accuracy of about 3 km/h) and on the acceleration of the car during the 2.9 sec time span, which can be significant in cases of large changes in velocity. Table 2 compares the velocities derived by this method with the results of the DGPS measurement.

	Velocity from DGPS (km/h)	Velocity from image (km/h)	Difference (km/h)
Vehicle 1			
Images 2/3	88.3	92.3	4.0
Images 3/4	87.7	89.0	1.3
Images 4/5	86.2	82.3	-3.9
Images 5/6	85.1	87.7	2.6
Images 2-6 (mean velocity)	86.9	87.6	0.7
Vehicle 2			
Images 2/3	86.0	90.0	4.0
Images 3/4	87.6	87.9	0.3
Images 4/5	88.6	85.0	-3.6
Images 5/6	87.8	90.4	2.6
Images 2-6 (mean velocity)	87.3	88.1	0.8

Table 2: Comparison of velocities derived from DGPS and aerial images

The maximal difference in velocity is 4 km/h, which is within the law of error for the two independent measurements. If one considers the velocity over a longer period, namely over the stretch that can be covered in about 15 seconds, then the discrepancy is below 1 km/h. It should be noted that the discrepancy for both vehicles in the respective images is quite similar. This is because the image overlay was not exact (see table 1), and can lead to the same systematic effect on the velocities of all vehicles.

The second comparison concerns the overhead radar detectors at four different places (signboard bridges) along the A9 and A99 freeways. On every signboard bridge there are several detectors for each track. The accuracy of the radar sensors, which were calibrated in the laboratory, was specified to ~3 km/h and a similar accuracy is expected with the image tracking procedure. In order to compare the two data sets an accurate visual interpretation is necessary since the velocity data from the radar sensors is only given with full seconds and the measurement takes place when the vehicle is at a distance of about 20 m from the bridge. Figure 3

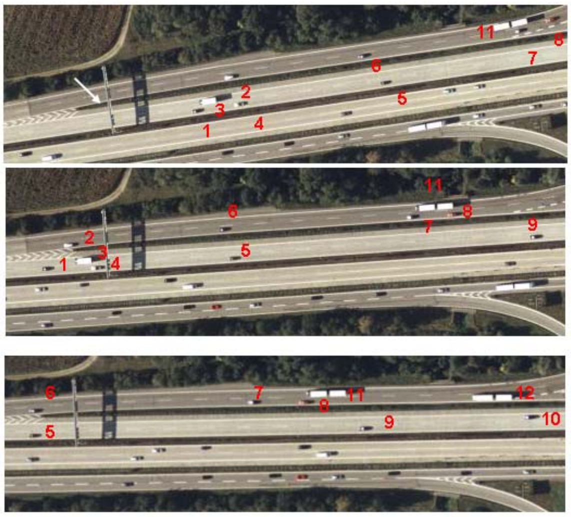

Figure 3. Three subsequent images with a signboard bridge (arrow) and several overtake maneuvers

shows three subsequent images with individual cars numbered. Several overtaking maneuvers are visible and often more than one car passes within one second. The results of the comparison vary for the different measuring locations. Table 3a shows an expected result with relatively low differences, while table 3b shows a result with a large difference. In all cases the standard deviations are higher than expected.

Sensor	Mean deviation [km/h]	σ_0 [km/h]
Track 1	4.4	7.6
Track 2	5.5	7.8
Track 3	3.4	5.8
Mean (3 tracks)	3.9	6.8

Table 3a: Comparison of velocities derived from radar sensors and aerial images (Bridge: AQ 92/610)

Sensor	Mean deviation [km/h]	σ_0 [km/h]
Track 1	2.5	4.7
Track 2	11.7	8.2
Track 3	12.7	8.3
Mean (3 tracks)	9.4	7.3

Table 3b: Comparison of velocities derived from radar sensor and aerial images (Bridge: AQ 9/370)

It is not easy to explain this behavior. After discussion with the freeway authority, it is thought to be caused by poorly adjusted sensor orientation of the radar equipment. The calibration was only performed in the laboratory but not “on location”. Even small changes in height or viewing angle will result in a large change of the velocity values. The results in figure 4, which imply systematic behavior, always measuring lower velocities with the radar sensors, could also be explained by the above mentioned reasoning. A linear regression between the two values leads to a slope of 0.8 and a linear regression coefficient of 0.93.

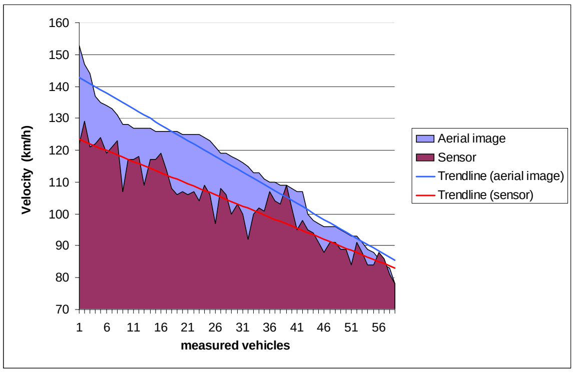

Figure 4. Velocity comparison for single vehicles measured with radar sensors and aerial images (linear regression for 60 cars at two signboard bridges)

The measured accelerations between each set of three images show typical values of increasing and decreasing speed on a freeway. With normal traffic flow, as observed during these measurements, accelerations are typically between ±0.5 m/s²; only in a few cases like starting overtakes or entering or exiting the freeway, does acceleration go up to 3 m/s² or even higher.

3.3 Navteq integration

In order to use the derived traffic parameters in a larger context, e.g., for simulations or traffic control, the data have to be stored in a traffic database. For this reason there is a need for automatic integration of the values to a common base of street lines. One of these street databases has been produced by the NAVTEQ Company (Navteq 2004). The streets are given by polygons which consist of piecewise linear “edges,” grouped as “lines” if the attributes of connected edges are the same. With this data set it is possible to calculate mean traffic density or velocity per edge/line for a certain time span. The accuracy of the Navteq database was first tested to see whether automatic assignment algorithms would be suitable. In the case of freeways the accuracy is quite high (only few m to the road/track center), while for some of the smaller streets, the absolute deviation can be on the order of 20 meters or more. There is no general statement possible about which streets are how accurate, but since the more interesting traffic performs on more frequently used streets, the NAVTEQ data set is very suitable for the mentioned application. With a simple algorithm which uses the perpendicular distance of the vehicle coordinates to the single edges, more than 95% of the cars could be assigned correctly. Only in the case of street crossings (including bridges) and freeway exits some errors appear due to wrong street assignment, which can be eliminated using algorithms containing the propagation direction of the vehicle. Table 4 shows typical distances in the automatic assigning algorithm. By this assignment a database could be established which can be used for traffic simulation studies or traffic flow investigations.

Vehicle Nr.	Distance to Navteq data (m)	Velocity from image (km/h)
001	3.4	85.5
002	1.2	88.8
003	0.5	81.8
004	3.6	135
005	3.8	89.0
006	0.5	87.1
007	2.5	121
008	2.3	85.2
009	5.0	150
010	6.9	83.8
011	2.0	118
012	2.2	128
013	0.7	71.9

Table 4. Distance from image data to Navteq data and velocities for one “Navteq-edge”

The traffic flow can be also visualized by interpolating the vehicle positions and color coding the single cars according to their velocity. With the background of the NAVTEQ data set, the assignment can be seen directly. Figure 5 shows a snapshot of the animation of real traffic flow on the Munich-North freeway interchange.

Figure 5. Clipping from a film animation showing moving vehicles on a freeway interchange north of Munich, blue=slow vehicles, orange=fast vehicles

4. AUTOMATIC DETECTION OF CARS

Manually determining vehicle positions is not satisfactory in the long run, therefore, some experiments were carried out to achieve automatic detection of moving objects. First results are given in this section. The described algorithm is based on a series of images taken over an interval of about 0.5 second by a normal digital camera placed on Munich’s Olympic Tower and aimed at the city freeway, as can be seen in figure 6. The primary purpose of the research was automatic detection of vehicles and extraction of their velocities. Surprisingly, many individual pedestrians on the images were also detected with their velocities.

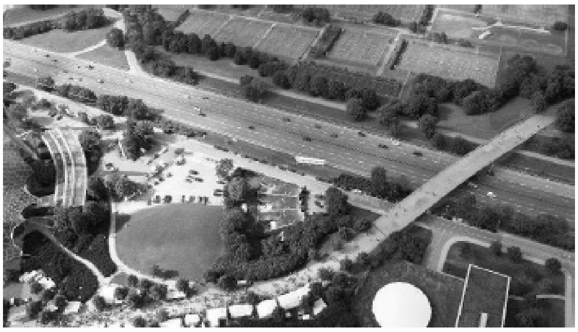

Figure 6. The first image of the series used –as a grey level image

In the first step all images were exactly mapped on the first image by the same matching procedure mentioned in section 3. Out of all these images a median image was created which does not contain the moving objects, as shown in figure 7. Next a difference image between each mapped image and the median image was calculated. These difference images were averaged over a radius of four pixels and the lower 20 % of grey values were cut off to eliminate noise.

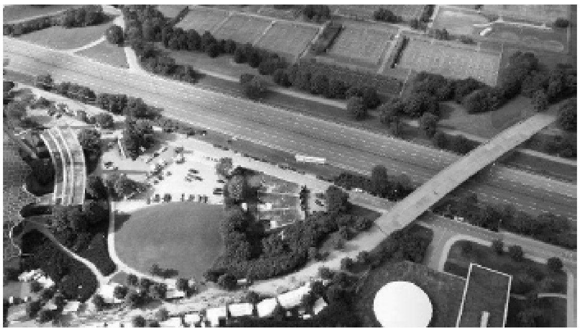

Figure 7. Median image created from all images of the series – showing no traffic on the city freeway at midday!

Applying a modified Laplace operator which averages slopes over a region of ±10 pixels detects the borders of objects. Filling these borders creates objects with specific areas. Taking into account only objects within a specific interval of areas (in the examples from 50 to 800 pixels) yields detected objects, as can be seen in figure 8. In the last step the areas of the objects detected in every single picture were correlated pixel for pixel between every two consecutive original images. Therefore, an object detected in one image is correlated with all objects detected in the following image. Correlations within a given quality number are taken as hits. The correlation is taken out within a small radius around the objects in the second image. These hits are drawn together with their movement and velocities as a vector layer to the first image (figure 9).

Figure 8. Difference to the median image, leaving only moving objects as visible traces

Figure 9. Object correlations found between the first two images and subsequently calculated velocities

As can easily be seen in figure 10 the objects are well correlated, and pedestrians were also detected if they were within the area defined for vehicles. Problems arise when too similar vehicles are in proximity. Adjusting the range of object areas and the quality number results in too few objects detected or too many mismatches found, which means that the receiver-operator curve (ROC) has to be observed carefully. The automatic method leads to detected car fraction of only 27% without generating false alarms. It rises up to 80% but generating in this case more than 20% false alarms. This has to be reduced in future works.

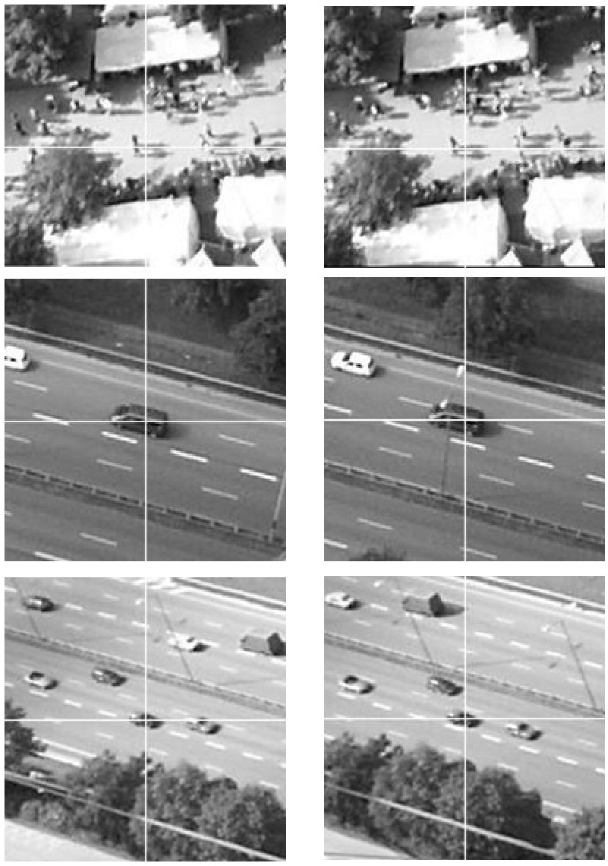

Figure 10. Three well correlated objects in subsequent images

5. DISCUSSION

The investigations show that it is possible to derive high quality traffic data from image series recorded by airborne cameras. With an easy to use visual interpretation tool 100% of the vehicles can be detected with a velocity accuracy of about 3 km/h. These data can be used for verification of standard instrumentation, for analysis of the formation of traffic jams, and for input into traffic models and simulations.

The automatic vehicle detection method leads to reduced car detection accuracies. This has to be improved by introducing constraints through the knowledge of the direction of the car movement. The interpretation works much faster but needs an image repetition frequency of at least 2 Hz. In any case the geometric identity of the overlapping images is a necessary prerequisite for high quality analysis. Future work will include the integration of the street database into the automatic algorithm and segmentation of the image into street and other classes.

Acknowledgement

We like to thank M. Ruhé from the DLR Institute of Transport Research for many valuable discussions and Dr. T. Linder, F. Frischeisen and C. Mayr from the "Autobahndirektion Südbayern" for providing data from on-site instrumentation and important input from the user side.

References

Brockfeld, E.; Kühne, R. D.; Wagner, P. 2004. Calibration and Validation of Microscopic Traffic Flow Models, In: Transportation Research Board [Eds.]: Transportation Research Record, Volume 1876, 2004, P. 62-70, TRB Annual Meeting, January 2004, Washington, DC (USA).

Ernst, I.; Sujew, S.; Thiessenhusen, K.-U.; Hetscher, M.; Raßmann, S.; Ruhé, M., 2003. LUMOS - Airborne Traffic Monitoring System, In: IEEE 6th International Conference On Intelligent Transportation Systems, 12-15 October 2003, Shanghai, China

Ernst, I.; Hetscher, M.; Lehmann, S.; Lippok, A.; Ruhé, M., 2005. Use of GIS Methodology for Online Urban Traffic Monitoring, Proceedings of 3rd International Symposium Remote Sensing and Data Fusion over Urban Areas (URBAN 2005), Tempe Arizona, USA, March 2005, on CD-ROM

Hinz, S., 2004. Car Detection in Aerial Thermal Imagery by Local and Global Evidence Accumulation, Technical Report, PF-2004-01, Remote Sensing Technology, Technische Universität München

Lehner, M., Gill, R.S., 1992. "Semi-automatic derivation of digital elevation models from stereoscopic 3-line scanner data," IAPRS, Vol. 29, Part B4 Washington, USA pp. 68-75, 1992

Mirchandani, P., Hickman, M., Angel, A., Chandnani, D., 2002. Application of aerial Videofor traffic flow monitoring and management. Percora15, ISPRS Commission I, Conference Proceedings, Denver, USA

Navteq website, 2005. http://www.navteq.com/data/whatis.html

Pötzsch, M., 2005. Messung von Fahrzeugpositionen und – geschwindigkeiten in Luftbildsequenzen und Zuordnung zu Straßendatenbanken, Diploma Thesis, Technical University of Dresden and DLR Remote Sensing Technology Institute, 97p.

Runge, H., Ruhé, M., Reinartz, P., 2004. Der Fernerkundungs-satellit TerraSAR-X hilft Unfallschwerpunkte zu analysieren. Verkehrstechnischer Tag des DLR, Braunschweig, Germany, 10p. CD-ROM and http://www.dlr.de/dlr/News/Tagungsband.pdf

Schäfer, R.-P.; Thiessenhusen, K.-U.; Wagner, P. 2002. A traffic information system by means of real-time floating-car data, ITS World Congress 11. - 14. October 2002, Chicage, USA.

Stilla, U., Michaelsen, E. Soergel, U. Hinz, S. Ender, J., 2004. Airborne monitoring of vehicle activity in urban areas. ISPRS Congress, Commission III, WG III/4, Istanbul, on CD-ROM

Toth, C.K., Grejner-Brzezinska, D., 2004. Traffic Management with state-of-the-art airborne imaging sensors. ISPRS Congress, Commission II, Istanbul, Turkey, on CD-ROM