TRAFFIC MONITORING WITHOUT SINGLE CAR DETECTION FROM OPTICAL AIRBORNE IMAGES

K. Zeller*, S. Hinz, D. Rosenbaum, J. Leitloff, P. Reinartz

German Aerospace Center (DLR), Remote Sensing Technology Institute, PO Box 1116, D-82230 Weßling, Germany
klaus.zeller@dlr.de

Commission IV, WG 3

ABSTRACT:

This article describes several methods for traffic monitoring from airborne optical remote sensing data. These methods classify the traffic into free flowing traffic, traffic congestion and traffic jam. Furthermore a method is explained, which provides information about the average speed of dense traffic on a defined part of the road. All methods gather the information directly from image features, without the use of single vehicle detection.
The classification of the traffic is done by stacking at least three overlapping images on top of each other and calculating the standard deviation of the gray values of each overlying pixel. In addition to that a texture analysis is implemented to differentiate the traffic. The average speed of dense traffic is calculated employing disparities of two following images of the same scene.
All methods, which were presented in this article were tested on various data sets and compared with interactive measured reference data.
These methods will be applicable in combination with methods using single car detection in the ARGOS-Project (AiRborne wide area hiGh altitude mOnitoring System).

1. INTRODUCTION

One important goal of the ARGOS-Project is the derivation of traffic data such as vehicle speed and vehicle density by airborne remote sensing in near real-time. After the completion of this system, flights will be mainly conducted in case of disasters or major events. After the automatically extraction of the traffic data, these data will be directly distributed to security authorities and organizations as well as action forces via the traffic portal DELPHI (Behrisch et al., 2008).
For this aim the actuality of the traffic data seems most important apart from the completeness and the reliability of the traffic data. The exact information about the position and speed of every single vehicle is not necessary. Hence the demand is to determine the traffic density and mean vehicle speed of all vehicles on a defined road part.
In this work only optical airborne images and no radar data, which are used in ARGOS too, are applied.
In combination with methods of airborne image processing in which every vehicle is extracted and tracked separately, also methods shall be inspected, which get traffic information directly out of image features.
A detailed description of the complete traffic extraction and characterization strategy is given in Hinz et al. (2008). In this article only algorithms, which get traffic information directly out of image features are explained. These information derived are a traffic classification, distinguishing between free flowing traffic, congestion, and holdup, and the computation of the speed of dense traffic.
Single vehicle extraction and tracking in low framerate aerial image sequences is also explained in Hinz et al. (2007). Another approach for traffic monitoring, which uses a change detection algorithm is described in Palubinskas et al. (2008).
Images are recorded with the 3K-Camera system, which was developed within the ARGOS-Project. This camera system consists three non-metric off-the-shelf cameras (Canon EOS 1Ds Mark II, 16MPix), which are adjusted in different directions.
The images are taken in a burst mode, which means that at least three images are taken at a high frame rate up to 3 fps, followed by a break of a few seconds until the next images are recorded in quick succession.
The georeferencing of the images takes place by use of GPS and IMU without any ground control points. Orthorectification is done with the use of a DSM. The ground sampling distance amounts from 0.14m at a flight altitude of 1000m to 0.43m at a flight altitude of 3000m.
The orthorectificated images as well as the traffic data are computed in near real time on board of the airplane. Afterwards, the data is transmitted via datalink to a mobile ground station. A detailed description of the workflow and the system components in the ARGOS-Project are given in Rosenbaum et al. (2008) and Kurz et al. (2009).

2. TRAFFIC MONITORING

The workflow of the traffic monitoring in the ARGOS-Project starts with the distinction of the traffic into three classes. Afterwards the final traffic parameters are extracted. The three traffic classes are free flowing traffic, traffic congestion and traffic jam. For each of this traffic classes a special extraction strategy is chosen, because no extraction strategy works well in all traffic situations.
In an operational processing chain for online traffic data extraction from aerial images, that will be accomplished in the end of 2009, it is planned to classify the traffic situation first and then to extract traffic data by different methods depending on traffic state.

After classification single vehicle detection is used to determine the traffic density in case of free flowing traffic. In that case, the average speed of defined road segments is calculated by tracking the vehicles from a single vehicle detection. A detailed description of the methods used for this kind of traffic data extraction based on single vehicle detection are not concern of this paper. It can be found in Rosenbaum et al. (2008).

In dense traffic the average speed is determined using disparities from following images. In the following the classification of the traffic and the calculation of the average speed of dense traffic are explained.
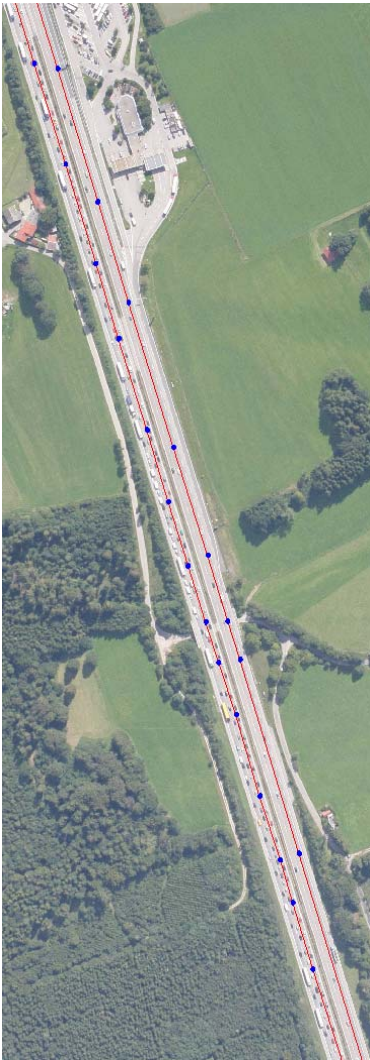
2.1 Data basis

Figure 1. Image section with NAVTEQ-nodes

The initial data for the traffic monitoring are three channel georeferenced and orthorectificated images with a resolution on ground from 0.14m to 0.43m. The images were recorded in a burst mode, so that the images of a burst feature a height overlap. Employing the GPS measurements the recording time of the images can be determined, which is necessary to calculate the speed of the vehicles.

The position of the roads in the images is gathered from the NAVTEQ database. This GIS database contains discrete points of the street axis in irregular distance (figure 1). Furthermore this GIS data also contains information about the street-type, the number of the lanes of the road and the direction of travel.

Thus Regions-of-Interest (RoI) can be defined. For each RoI an average speed and density shall be calculated as result of all extraction strategies.

It is advantageous to straighten the roads by resampling the image parts along the GIS axis (figure 2) before the calculation of the average speed, in order to approximate epipolar geometry. To simplify the workflow of the traffic data extraction, these resampled images are used for all extraction methods.

Figure 2. Resampled image and road segments

2.2 Traffic Classification

The traffic classification is realized by determining a rough measure of the traffic density inside the RoI.

One method to calculate a rough measure of the density is to stack all overlapping images on top of each other and calculate the gray value standard deviation of the overlying pixels. This is possible due to the used burst mode, which leads to large overlapping areas of at least three consecutive images.

The calculation of the standard deviations of the stacked pixels provides an image in which standard deviations are illustrated as gray values.

If a pixel has approximately the same gray value in all images of a burst, the conclusion can be made, that no vehicle has moved through this pixel on ground. In this case the calculation of the standard deviation returns zero and the considered pixel is illustrated black in the standard deviation image. The standard deviation of a pixel is not zero, if a vehicle moves through this pixel on the ground during a burst. The standard deviation image of the scene which is illustrated in figure 3.a is shown in figure 3.b.

The colour differences between the vehicles and the lanes have influences on the extend of the temporal standard deviation. To eliminate this effect a binary image is created adapting a threshold to the standard deviation image. Figure 3.c illustrates a binary image from free flowing traffic and figure 4.b from dense traffic.

In the following, mean gray values inside of the RoI of the binary image are calculated, to get a relation between pixels where changes took place and pixels with no changes. Hence road segments with a high traffic volume have a higher mean gray value then less frequented roads.

Figure 3a. First image of a scene with free flowing traffic

Figure 3b. Standard deviation image

Figure 3c. Binary image of free flowing traffic

Figure 4a. Scene with static and slow moving vehicles

Figure 4b. Binary image of dense traffic

Figure 4c. Binary image with shifted georeferencing

An accurate quantification of the traffic volume is not possible, because the variable and unknown size of the vehicles reacts directly on the result. It is also not possible to deduce the number of vehicles from the number of regions in the binary image, due to the existence of overlapping vehicles in the time series especially in dense traffic.

It should also be noted, that parts of roads with only stationary vehicles (traffic jam) return the same result as roads with no vehicles, namely a black standard deviation image. Therefore two strategies to distinguish road segments with only static vehicles and empty roads are developed.

One strategy is to shift the georeferencing of one image of the burst (figure 4c). This has the effect that stationary vehicles are recognizable in the image of the temporal standard deviation. Whereas empty parts of the roads still result in dark parts of the standard deviation image in case of a homogeneous road surfaces.

The other strategy to differentiate between traffic jam and empty roads is a texture analysis.

This texture analysis computes the local gray value standard deviation of the RoI. Furthermore the contrast is calculated out of the co-occurrence matrix. Regions with a high local gray value standard deviation and a high contrast correspond to traffic jam, whereas regions with a minor standard deviation and a minor contrast correspond to empty roads.

The biggest problem during the texture analysis as well as using shifted georeferencing is shadows from the surroundings which causes false traffic information.

That is the reason why these methods are only used to differentiate between empty roads and road segments with static vehicles, whereas they deliver good results in images with good illumination conditions.
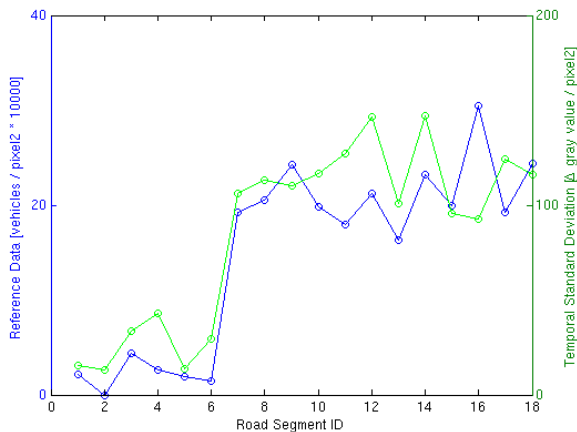

Figure 5. Temporal gray value standard deviation
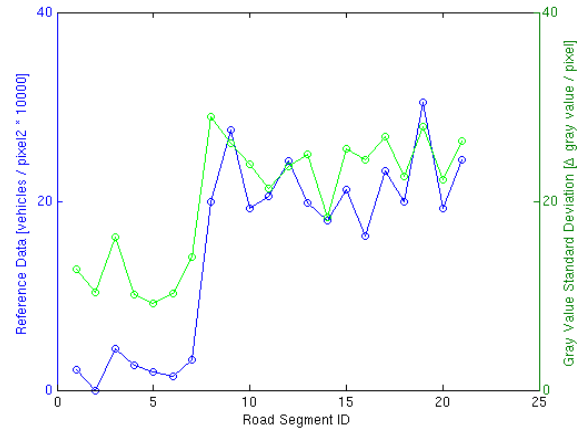

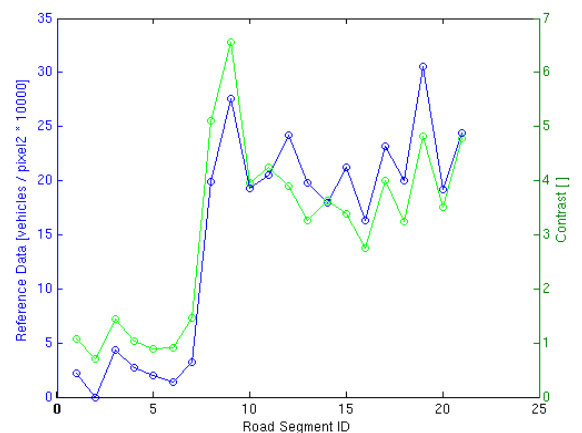Figure 6. Local gray value standard deviation

Figure 7. Contrast

The results of these methods were tested on several data sets. Figure 5 to figure 7 shows the calculated and the interactive measured reference data from the scene, which is shown in figure 1.

Figure 5 shows the temporal gray value standard deviation of the road segments, which are illustrated in all images of the burst. Road segment ID 1 to 6 are part of a free flowing traffic lane, whereas in road segment ID 7 to 18 the traffic data is calculated for a traffic lane with dense traffic. Because of the different size of the road segments the numbers of manual measured vehicles are normalized to the segment area. The chart shows, that it is possible to determine a roughly measure of the traffic density of the road segments out of the temporal standard deviation.

Figure 6 and 7 illustrates the results from the texture analysis. The texture analysis is done for all road segments of the first image of the burst. In the results of the local standard deviation and the contrast, the behaviour of the traffic density is observable in the test data set. In cases with worse illumination conditions as in the test data set (figure 1), a worse result can be expected.

The results show, that the combination of the methods to determine a rough measure of the density is adequate to classify the traffic into three classes. However these methods are not applicable to determine the accurate density of the RoI.

For this purpose methods which use single vehicle detection are advisable (Hinz et al., 2007). Another algorithm which provides good results for the traffic density is explained in Palubinskas et al. (2008). The algorithm is based on a change detection method

to extract the vehicles. The speed of the vehicles is calculated by applying the traffic density to a traffic model.

A non model based method, which gets the speed information out of measurements, is the tracking of extracted vehicles. The tracking of vehicles provides good results in free flowing traffic but is problematically in dense traffic, because of ambiguity in vehicle allocation. Therefore it is useful to apply the following approach to compute the speed of dense traffic.

2.3 Computing the speed of dense traffic

In this part the average speed of road segments with dense traffic is computed, without the extraction and tracking of single vehicles.

Therefore it is useful to straighten the illustrated roads by resampling the image along the GIS axis. Then the geometry of the resampled images is nearly the same as in epipolar geometries and two consecutive images of a burst can be used as image pair. The correspondences of the two images are determined by use of the normalized cross correlation.

In dense traffic the correspondences are only found on the vehicles, because the vehicles cause the most marked texture in the road scenes. Hence the disparities of the correspondences can be considered as shift vectors between the vehicles of two images. With the knowledge of the pixel resolution on ground and the time difference between two images, the average speed of the road segments is calculated from the disparities.

It should be noted, that this method only works in dense traffic and not in free flowing traffic, because the disparities of the correspondences would also be calculated for the background (roads) and not only for vehicles. Figure 8 illustrates the comparison between manual measured reference data and the calculated data. Here it is shown, that the method to calculate the speed doesn't work in free flowing traffic but provides good results in dense traffic.
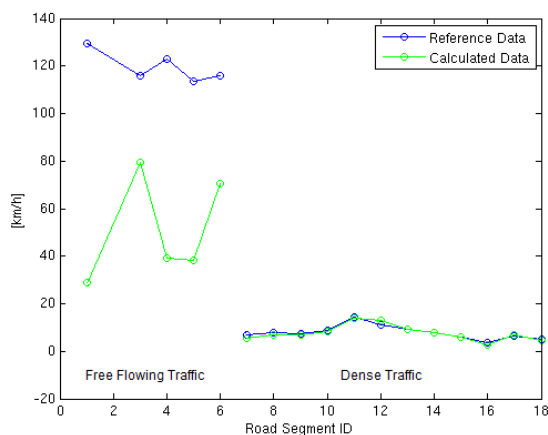

Figure 8. Average speed

3. CONCLUSION

Comparisons with reference data have shown that it is sufficient in certain cases to calculate traffic information directly out of image features without extracting single vehicles. It should be noticed that the measure of the traffic density calculated from the temporal standard deviation or the texture is not applicable to derive absolute vehicle densities directly, as it only reflects the number of pixels where changes took place. Thus a direct vehicle count for deriving vehicle number densities is not possible. However, the used measure of the traffic density is appropriate to distinguish traffic into three classes. Then, in case

of dense traffic or complete holdup the derivation of velocities without single vehicle detection is possible with the disparity algorithm. It could be shown that the velocities obtained from disparities provide good results. Since the proposed method does not work in free flowing traffic situations, tracking algorithms based on single vehicle detections should be preferred in those cases.

REFERENCES

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

Hinz, S., Lenhart, D., Leitloff, J., 2007. Detection and tracking of vehicles in low framerate aerial image sequences. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences,* 36(1/W51). 6 pages (on CD-ROM). http://www.commission1.isprs.org/hannover07/paper/Hinz_lenhart_leitloff.pdf (accessed April 8, 2009).

Hinz, S., Lenhart, D., Leitloff, J., 2008. Traffic extraction and characterisation from optical remote sensing data. *The Photogrammetric Record 23(124).*

Kurz, F., Rosenbaum, D., Thomas, U., Leitloff, J., Palubinskas, G., Zeller, K., Reinartz, P., 2009. Near real time airborne monitoring system for disaster and traffic applications. In: *ISPRS Hannover Workshop 2009, High Resolution Earth Imaging for Geospatial Information,* 2-5, June, 2009, Hannover, Germany, ISPRS. In print.

NAVTEQ, 2009. http://www.navteq.com/about/data.html (accessed April 8, 2009).

Palubinskas, G., Kurz, F., Reinartz, P., 2008. Detection of traffic congestion in optical remote sensing imagery. *Proc. of IEEE International Geoscience and Remote Sensing Symposium (IGARSS'08), 6-11 July, 2008, Boston, USA, IEEE.*

Reinartz, P., Lachaise, M., Schmeer, E., Krauß, T., Runge, H., 2006. Traffic Monitoring with Serial Images from Airborne Cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 61, pp. 149-158.

Rosenbaum, D., Kurz, F., Thomas, U., Suri, S., Reinartz, P., 2008. Towards automatic near real-time traffic monitoring with an airborne wide angle camera system. *European Transport Research Review, 1,* Springer, ISSN 1867-0717.