A Traffic Object Detection System for Road Traffic Measurement and Management

Carsten Dalaff
Institute for Transport Research,
German Aerospace Center DLR ,Berlin, Germany
Carsten.Dalaff@dlr.de

Ralf Reulke
Institute for Photogrammetry, Stuttgart University, Germany
Ralf.Reulke@ifp.uni-stuttgart.de

Axel Kroen
ISSP Consult, Stuttgart, Germany
kroen@stgt.ssp-consult.de

Thomas Kahl
ASIS GmbH, Berlin, Germany
Thomas.Kahl@asis-it.de

Martin Ruhe, Adrian Schischmanow, Gerald Schlotzhauer, Wolfram Tuchscheerer
German Aerospace Center DLR ,Berlin, Germany
firstname.lastname@dlr.de

Abstract

OIS is a new Optical Information System for road traffic observation and management. The complete system architecture from the sensor for automatic traffic detection up to the traffic light management for a wide area is designed under the requirements of an intelligent transportation system. Particular features of this system are the vision sensors with integrated computational and real-time capabilities, real-time algorithms for image processing and a new approach for dynamic traffic light management for a single intersection as well as for a wide area. The developed real-time algorithms for image processing extract traffic data even at night and under bad weather conditions. This approach opens the opportunity to identify and specify each traffic object, its location, its speed and other important object information. Furthermore, the algorithms are able to identify accidents, and non-motorized traffic like pedestrians and bicyclists. Combining all these single information the system creates new derivate and consolidated information. This leads to a new and more complete view on the traffic situation of an intersection. Only by this a dynamic and near real-time traffic light management is possible. To optimize a wide area traffic management it is necessary to improve the modelling and forecasting of traffic flow. Therefore the information of the current Origin-Destination (OD) flow is essentially. Taking this into account OIS also includes an approach for anonymous vehicle recognition. This approach is based on single object characteristics, order of objects and forecast information, which will be obtained from intersection to intersection.

Keywords: Traffic observation, traffic control, sensor network, sensor fusion

1 Introduction

Traffic observation, control and real-time management is one of the major components within future intelligent transportation systems (ITS). One central postulation of the European Government is the increase of road safety, so the number of killed people should be halved till the year 2010. There are nearly 41.900 road casualties and more than 1.7 million seriously injured persons each year in the European Union (EU). This causes about 45 billion Euro direct and approximately 160 billion Euro external costs per year. Daily 4.000 km of traffic congestions stress only the European highways. This means 10% of the complete European highway system. The economical damage is tremendous, but till now there is no common approach to calculated the real amount. So the official numbers differ. The necessary investments for the European transportation field will reach more than 10% of the EU gross national product (GNP). The needed financial capabilities for transport infrastructure of the acceding countries (e.g. Poland, Estonia) will increase this expense enormously. To realize just the priority projects in these countries, the EU has to be spend 91 billion Euro up to 2015. Taking this into account it seems effective to use parts of this money for innovative approaches for traffic management in future intelligent transportation systems. Having such an intelligent transportation system an increased road safety can be realized. So the economical damage can be reduced. One appropriate approach could be the use of the mentioned traffic object detection system for road traffic measurement and management. This systems is different to state-of-the-art traffic measurement equipment, e.g. induction loops, which does not suffice anymore the growing demand of transport research and traffic control.

The project OIS [1] uses optical and informational enabling technologies for an automatic traffic data generation with an image processing approach. Its main purpose is to acquire and evaluate autonomously traffic image sequences from roadside cameras. Traffic parameters will be obtained from extracted and characterized objects of this image sequences. To meet these requirements, numerous image processing algorithms have been developed since more then 20 years (e.g. special issue [2]), with simple web-cameras and more complex systems (e.g. [3]).

Traffic scene information can be used to optimize traffic flow on intersections during busy periods, identify stalled vehicles and accidents, and is able to identify non-motorized traffic like pedestrians and bicyclists. Additional contributions can be obtained for the determination of the Origin-Destination (OD) matrix. The OD matrix contains the information where and when the traffic participants start and end their trip and which route they have chosen. OD matrix is one basic element for an optimized modelling and forecast of traffic flow. The estimated traffic flow is necessary for a dynamic wide area traffic control, management and travel guidance.

Furthermore, the recent advances in computational hardware can provide high computational power with fast networking facilities at an affordable price. The availability of specific solutions in the low-cost general-purpose range allows special image processing and avoids some basic bottlenecks. A couple of traffic data measurement systems already exist. Best known is the induction loop. Induction loops are embedded in the pavement. They are able to measure the present of a vehicle, its speed and rough classification. These are only local information but for a wide area traffic management a big coverage of the area is needed.

Another approach is based on the idea that moving vehicles transmit information about there position and velocity via mobile communication, e.g. GSM to a traffic management center. These data are called Floating Car Data (FCD). To get an overview of the traffic situation of a complete city at any time a huge number of vehicles has to be equipped with hardware and mobile communication units. To get reliable data every one minute the current position and velocity of the vehicles is needed. This causes enormous costs for the mobile communication. The FCD approach provides spatial traffic information, but the spatial and time resolution doesnt fit the requirement of a traffic signal control. OIS as a new and innovative traffic observation system that opens the opportunity to deliver all necessary input for a local traffic signal control as well as for a dynamic wide area traffic management system. Next challenge is the implementation of OIS in a wide area city.

2 System Requirements

A modern system for traffic control and real-time management has to meet the following requirements:

- Reliability under all illumination and weather conditions
- Working period non-stop, 24hours 7 days a week
- complete overview over the intersection from at least 20m in front to 20m behind
- working in real-time
- real-time (every half second a complete data set of the traffic situation) on an intersection

This requirements should be taken into account for the design of all parts of the system, which covers all procedures and processes from image data acquisition, image processing and traffic data retrieval up to traffic control. For realizing an operating system for 24 hours and under different weather conditions infrared cameras should be used. Algorithms had to be developed for a special camera arrangement (with real-time demands) for vehicle detection and deriving relevant parameters for traffic description and control.

3 System Overview

To get a complete overview over the intersection from at least 20m in front to 20m behind it is necessary to have more than one camera. The number of required cameras depends on the intersection geometry and installation possibilities for the cameras on house walls or lampposts. The time synchronous image data acquisition from one observation point with different cameras will be done in a so-called camera node. The camera node is part of the OIS-philosophy and consists of different sensors as a combination of VIS and IR-cameras. To fit the real-time processing requirement, time consuming image processing parts will be implemented in a special real time-unit. This approach is in implementation at the moment for two intersections in

Berlin, Germany equipped with camera-systems. The architecture of the complete system is shown in the first figure.

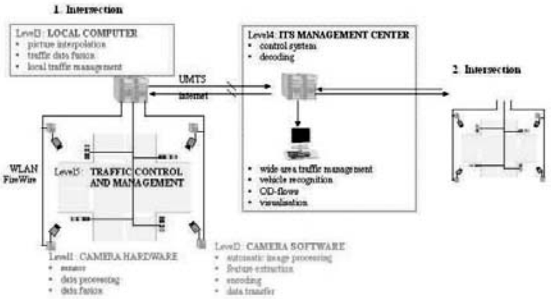

Figure 1: System design for the test bed.

As shown in fig. 1 the camera nodes are part of a hierarchy, which are linked via Internet or Wireless LAN with the computer systems on the intersection and the management center. Due to the limited data transfer rate and possible failure of a system on an intersection the camera node works independently. Image processing is decentralized and will be done in the camera node, so that only objects and object features will be transmitted. Together with error information the object data will be collected on the next level in the hierarchy. For synchronization purposes time signals can be incorporated over a network or from independently received signals (like GPS). The camera node is part of a hierarchy starting with camera nodes, junctions, subregions, etc. The next so-called junction level unifies all cameras observing the same junction. Starting with the information from the camera level, relevant data sets are fused in order to determine the same objects in different images. After that, these objects are tracked as long as they are visible. So traffic flow parameters (e.g. velocity, traffic jam, car tracks) can be retrieved. Level number 3, the so-called region level, uses the extracted traffic flow parameters out of the lower level and feeds this data into traffic models in order to control the traffic (e.g. switching the traffic lights). Additional levels can be inserted. For test applications the camera node will generate a compressed data stream. For typical working mode only object features in an image should be transmitted and collected in a computer of the hierarchy. Generally, optical sensor systems in the visible and near-infrared range of the electromagnetic spectrum have reached a very high quality standard, which meet the requirements even for high-level scientific and commercial tasks, above all concerning the radiometric and geometric resolution and data rate. Otherwise sensors working in the thermal infrared range (TIR) are still a research topic for traffic applications. Technology development for the next few years will not be focused on higher resolutions or faster read-outs, because for most applications the performance of these sensors is sufficient. The emphasis will be put on smart, intelligent sensor systems with different measurement parameters (e.g. different resolutions, different spectral sensitivities) which are connected within a network similar to the internet and being able to convert the incoming physical signals not only to digital data but to process them to user needed information. Therefore, image fusion as well as fast and reliable algorithms are needed preferably near the sensor itself. Real-time processing and programmable circuits will play an important role.

4 Hardware Concept

The hardware concept is oriented on the requirements, as described above. The system has to operate 24 hours and 7 days a week and to observe the whole junction and at least 20m of all related streets. To improve the opportunities for image acquisition more than one camera system for one observation standpoint should be in use. Such a camera node fits the first requirement with a high (spatial) resolution camera and a low resolution thermal infrared camera. Also stereo and distance measurements are possible with two identical cameras. The camera node is able to acquire data from up to four cameras in a synchronized mode. The camera node has real-time data processing capabilities and allows synchronous capture of GPS and INS (inertial navigation system) data, which should be external mounted. Due to limitations of camera observing positions and possible occlusions (e.g. from buildings, cars and other disturbing objects) for intersection observation mostly more than one standpoint for a camera node (camera-system) is necessary. Therefore communication or data transmission between camera nodes and the computer in the next higher hierarchy becomes critical. To reduce data volume in the network the real time data processing capability of the camera-node is used to speed up image processing and to transmit only object data. For the real-time unit a hardware implementation was chosen. Large free programmable logic gate arrays (FPGA) are available now. A programming language is available (VHDL) and different image processing algorithms are implemented.

5 Data Processing

The processing of data in the sensor web follows successive steps. In the first step image data are generated and pre-processed. After that the objects are extracted out of the images. In the last step all the object features from different camera nodes will be collected, unified and processed into traffic information. The data processing is optimized to the logical design of the system. There are operations for the system configuration, for the operational working for the data mining and visualization. Results are new data and information. A complex data management system regulates the access, the transmission, handling and the analysis of the data.

5.1 Image Processing

The essential processing steps are the elimination of noise and systematic errors, compression/decompression, higher level image processing and spatial (or geo-)- and time-referencing. The last point is necessary to determine space and time coordinates of observed objects as an essential feature for data fusion. Most of the work on vehicle detection or recognition was done on ground images, mainly as pre-processing before tracking for surveillance or traffic applications [4]. Different approaches, e.g. finding edges [5] or deformable model for vehicle [**?**] can be found. Generally, there are problems in image processing with car occlusion [6] and shadows [7]. Beside stationary image acquisition from ground also from moving platforms and arial images [8] are used. Image processing approaches will also be used for detecting lanes and obstacles by fusing information [9]. Image processing for OIS was described in detail in [1]. The main processing is a classical object detection and identification task. Following major problems have to be solved - Object discrimination from a spatial and time variable background (cross, street, buildings, etc.) - Removing disturbing structures between object and camera, as well as shadow regions around the object - Identification of cars in a row, which are occluded by other The first problem can be solved at least in two different ways:

1. Working with the image sequences, which are subtracted from the image before or
2. following the time changing background.

For the first approach background can be eliminated very simple, but stalled vehicles are invisible. The other approach is an on-going update of the background. This needs much more expense, but allows a detection of moving and also stalled cars. As a result of this procedure, background can be subtracted from the current image and objects can be derived as shown in fig. 2. Additional morphological operation removes clutter and close objects structures. The right image shows the grey coded objects after labelling.

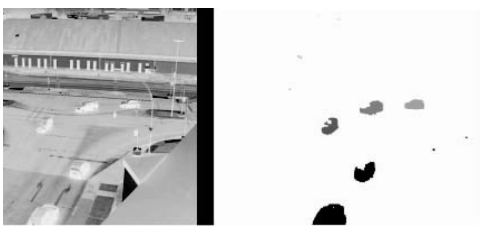

Figure 2: Object detection in a traffic scene (left: original image, right processed image.

The other major problem occurs after detecting the object. To determine size and shape, disturbing effects like shadows have to be removed. A straightforward way is the analysis of grey and color values, as well as texture within the found object boundaries.

5.2 Data and object fusion

To ensure an operating system 24 hours a day even under bad weather conditions, e.g. rain, fog, and at night the fusion of a visible and an IR sensor is a promising approach. The fusion is possible on different stages of information processing:

- data level (e.g. image data from different cameras)
- object level (objects extracted from the image data)
- information level.

Image matching and registration as one part of the data fusion is a procedure that determines the best spatial fit between two or more images acquired at the same time and depicting the same scene, by identical or different sensors. To fuse the different images and/or object data a synchronization of time is necessary. Time synchronization can be realized by internal clocks (e.g. computer) or external time information (e.g. GPS). For the application of OIS we merge images on camera node level and fuse the position information and object features on junction or region level. The object information are fed in from different camera nodes. Both procedures should be explained more detailed. Fusion on data level: To fit the requirement for 24-hour observations, typical CCD-cameras fail because of limited illumination of the objects. Car headlights and rear lamps seems not be sufficient. To overcome this problem the car self-radiation, which has a maximum in the thermal spectral region (TIR), can be used. There are a bunch of different detectors sensitive in this spectral range. Most of the sensors are expensive and needs additional cooler. Recent developments show, that bolometer arrays are a candidate for cheaper and uncooled detector arrays. Therefore, a camera development was started, which gives full access to the sensor, control, data correction and dataflow. First experiments were done with a commercial system. An example of the data and the fusion of both on data level is shown in fig. 3.

Observation was done on late afternoon. The left image is a typical CCD-image. Contrast becomes smaller, only reflections from sun glitters on car roofs are visible. The middle image was taken from a bolometer sensor. The whole intersection and the street are visible. The right image is the fusion of both. For visualization grey level image was put in the green channel and the TIR image in the red. Merging the visible image and the IR-image a affine transformation was used. To fuse data from different sources is obvious, but needs spatial and time synchronization, because of different imaging system and observation conditions. For this example, the synchronization task is based on manual procedures, like finding equivalent points in both IR and VIS images and the calculation of the necessary transformation. Especially the automatic spatial synchronization is a research topic. All these operations are done in each camera node. Another advantage of TIR images shows the following image pair, which was taken from the same observation point, but at daytime.

Figure 3: Sensor fusion of VIS an IR images (I).

Figure 4: Sensor fusion of VIS an IR images (II).

Fig. 4 shows an example of fusing RGB and TIR images at daytime. After coregistration the IR-image to the RGB-image and applying affine transform, a direct comparison is possible. In difference to figure 1 the infrared image was fit directly into the grey level image (a color separation of the RGB-image). The RGB image has a much more higher resolution than the TIR-image (720x576 pixel). Spatial and true color object data can be derived from this image. The TIR-image has a smaller resolution (320x240 pixel). The combination of both images gives a more complex result, because of the thermal features, which can be observed on the engine bonnet and as reflected radiation from under the car. These are also new features for object detection and description within the image-processing task. Object Data Fusion: Occluded regions can only be analyzed with additional views on these objects. Because camera nodes always transmit object information, a data fusion process on object level is necessary. The object list from different camera nodes has to be analyzed and unified. Object features like position, size, and shape vary from each view angle. Therefore the different object features from different perspectives have to be compared. The result of this operation is one traffic object with an exact position, size and shape, etc. The assumptions for this operation are time and spatial synchronized image data. The sequential processing of this list allows the derivation of more detailed information, e.g. track of the traffic participants, removing erroneous objects, etc.

5.3 Data Acquisition and Georeferencing

Transformation from image to world coordinates is an essential need in order to calculate metric based traffic data. Standard photogrammetric procedures for the transformation of coordinates within monocular images are used. Basic assumptions are well distributed and accurate measured ground control points (GCP) in object and image space as well as an exact camera calibration. The GCPs are calculated via DGPS within WGS84 and UTM-projection. The calculated camera calibration parameters (interior and exterior orientation) and image coordinates are input for the transformation equations. Due to monocular image accusation the object or vehicle positions have to be projected on a XY-plane in object space. The vehicle positions projected on that plane depend on the camera distance, the camera declination and the position of the point within the vehicle representing its position. Once the camera calibration is set the vehicle positions can be transformed to world coordinates within image sequences until the camera position changes. Additionally, intersection geometry for example lanes etc... can be transformed from image into object space.

5.4 Calculation of Traffic Characteristics

The following figure 5 shows the general traffic characteristics calculation process.

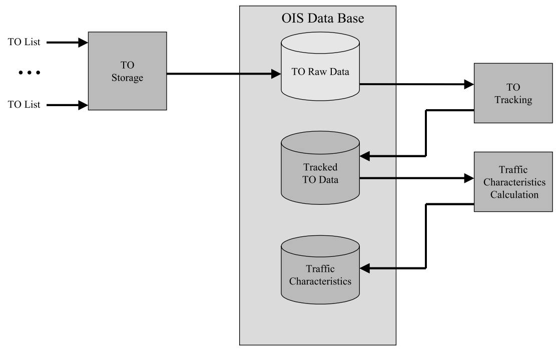

Figure 5: General traffic characteristics calculation process.

The image processing (not shown above) cyclically delivers the parameters of all identified traffic objects such as cars, trucks, cyclists and pedestrians in traffic object lists (TO Lists), containing type, size, speed, direction, geographic location etc. of the objects for a certain sample time. A storage procedure (TO storage) writes this raw traffic data into a data base for further processing. A tracking procedure (TO Tracking) marks traffic objects appearing in consecutive time samples by an unique object identification. Thus, traffic objects can be pursued throughout the observed traffic area. The tracked traffic objects and their parameters are stored in the Tracked TO Data area of the data base. Using this data, a the Tracking Characteristics Calculation module computes traffic characteristics, e.g. traffic density or flow rates. To compute the traffic density the number of motor vehicles on a certain road segment is necessary. Using their geographical coordinates and direction all motor vehicles moving along a road segment are selected form the data base. By simple counting the number of this vehicles and scaling their number to a one kilometer segment the traffic density can be obtained. Flow rates can be determined by counting the number of cars crossing a defined traverse section. Based on the traffic object parameter set over a number of sample times and using the tracking information, the number of vehicles crossing the section is counted and a traffic flow measure in vehicles per hour or so can be obtained.

6 An Adaptive and High Dynamic Network Control

The video based traffic sensor developed in this project creates possibilities for new concepts of traffic control for intersections and wide area networks. This approach also includes the development of a new dynamic and adaptive traffic control models for traffic lights. State-of-the-art in traffic observation is the induction loop. Induction loops are embedded in the pavement and register about $1 \cdot 2m$. This kind of sensor is able to measure the present of a vehicle, its speed and rough classification. In a next processing step it is possible to calculate time intervals between vehicles. This data is needed to control traffic light signals on intersections. For a real dynamic and demand based traffic signal light control, you would need the data of numerous induction loops on one single intersection. This is neither efficient nor realizable. OIS sensor web offers a new kind of traffic data/information. It is based on so called traffic-actuated signals. This means, the system is detecting information about the real-time traffic situation on an intersection automatically. For example the length of the queue for all different lanes, the traffic flow at the intersection, the density of traffic and the current velocity of each vehicle. The OIS sensor web automatically processes data-position, speed vectors of each vehicle, queue length as well as other relevant features. This leads to a complete and real-time overview at least 20 m before and behind an intersection. Based on this new quality of data, new approaches are developed to control traffic light signals on dynamic demand. At the moment most of traffic lights are controlled by two different ways: 1. control by fix-time signals 2. control by so called actuated signals Fix-time signals means: green and red time is fixed over the time and independent of the actual situation on the intersection. Actuated signals means: a number of fix-time signals are used for different demands and situations. At this time there are nearly no sensor webs or measurement systems available, that are able to measure each object on a intersection and that provides all necessary spatial information for a real dynamic traffic control system. OIS is a system that is able to acquire features like size, shape and other object features to classify and identify traffic objects. Summarizing this, OIS gives a complete data information (overview) over a whole intersection.

The project is funded by the German government (Federal Ministry of Education and Research, registration number: 03WKJ02B).

References

[1] R. Reulke, A. Bœrner, H. Hetzheim, A. Schischmanow, and H. Venus. A sensor web for road-traffic observation. In *Image and Vision Computing, New Zealand 2002*, pages 293–298, 2002.

[2] A. Broggi and E. D. Dickmanns. Applications of computer vision to intelligent vehicles. *Image and Vision Computing*, 18(5):365–366, April 2000.

[3] F. Pedersini, A. Sarti, and S. Tubaro. Multi-camera parameter tracking. *IEE Proceedings - Vision, Image and Signal Processing*, 148(1):70–77, February 2001.

[4] M. Betke, E. Haritaoglu, and L. S. Davis. Real-time multiple vehicle detection and tracking from a moving vehicle. *Machine Vision and Applications*, 12(2):69–83, 2000.

[5] J. Canny. A computational approach to edge detection. *PAMI*, 8:679–698, 1986.

[6] I. Ikeda, S. Ohnaka, and M. Mizoguchi. Traffic measurement with a roadside vision system - individual tracking of overlapped vehicles. In *Proceedings of the 13th International Conference on Pattern Recognition*, pages 859–864, 1996.

[7] G. S. K. Fung, Nelson H. C. Yung, Grantham K. H. Pang, and Andrew H. S. Lai. Effective moving cast shadow detection for monocular colr traffic image sequences. *Optical Engineering*, 41:1425–1440, June 2002.

[8] R. Chellappa. An integrated system for site model supported monitoring of transportation activities in aerial images. In *Proceedings of the 1996 ARPA Image Understanding Workshop*, pages 275–304, 1996.

[9] M. Beauvais and S. Lakshmanan. Clark: a heterogeneous sensor fusion method for finding lanes and obstacles. *Image and Vision Computing*, 18(5):397–4, April 2000.