REFINING CORRECTNESS OF VEHICLE DETECTION AND TRACKING IN AERIAL IMAGE SEQUENCES BY MEANS OF VELOCITY AND TRAJECTORY EVALUATION

D. Lenhart[1], S. Hinz[2]

[1]Remote Sensing Technology, Technische Universitaet Muenchen, 80290 Muenchen, Germany
Dominik.Lenhart@bv.tu-muenchen.de
[2]Institute of Photogrammetry and Remote Sensing, University of Karlsruhe, 76128 Karlsruhe, Germany
Stefan.Hinz@ipf.uni-karlsruhe.de

KEY WORDS: Traffic Monitoring, Vehicle Trajectories, Aerial Image Sequences, Fuzzy Logic, Evaluation

ABSTRACT:

Derivation of statistical traffic data is highly dependent on the balance of detection and false alarm rates. In case false alarms have not been eliminated in the initial detection phase, they are often subsequently tracked, though, resulting in trajectories that do not match the true traffic situation. This finally leads to derivation of erroneous traffic parameters within the individual road segments. In this paper, a method is described how to eliminate false alarms by evaluating the trajectories and velocities of a tracking procedure. Basically, two types of false alarms are considered which bias the statistics of traffic data: The first type deals with redundant detections that lead to multiple trajectories biasing the statistics. The second type comprises false alarms that belong to the static background inducing zero-velocity into the statistics. We show that the presented procedure is able to increase the total correctness of detection and tracking from 65% up to 95% which allows a much more precise calculation of traffic flow parameters.

1. TRAFFIC MONITORING

The task of collecting wide area traffic parameters plays important role in today's traffic management. Aerial images offer a complement source to common measurement systems like induction loops and stationary video cameras. Besides giving a visual overview, image sequences which cover large areas can deliver a time snapshot of a spatially fully covered traffic situation of the recorded region.

In recent years, traffic monitoring using air- and space images became more and more attractive mainly due to the availability of cost-effective and flexible high-resolution systems mounted on aircrafts, i.e. the LUMOS/ANTAR system for traffic monitoring (Ernst et al., 2003; Ernst et al. 2005; Ruhé et al., 2007) or the 3K camera system (Kurz et al., 2007), or on HALE platforms and UAVs as presented in the Pegasus project (Everaerts et al., 2004). An extensive overview on recent developments is given, for instance, in (Stilla et al., 2005; Hinz et al., 2006; Lenhart et al., 2008). The following methods are especially designed for traffic monitoring with DLR's 3K camera system. This system is able to capture image sequences with a frame rate of approx. 3Hz – 7Hz depending on the imaging mode (continuous imaging or bursts) with a spatial resolution of 20cm – 50cm depending on the flight height. Concepts for deriving traffic data from these aerial image sequences have been proposed in (Rosenbaum et al. 2008) and (Lenhart et al. 2008). The traffic parameters which are calculated from image sequences are namely the mean velocity and traffic density per road segment. The resulting parameters are then integrated into traffic flow models such as the DELPHI traffic portal illustrated in (Behrisch et al.).

2. INFLUENCE OF FALSE ALARMS

Detection methods as proposed in (Rosenbaum et al. 2008) or (Lenhart et al. 2008) deliver a detection quality of about 60% completeness and 65-75% correctness. False alarms are mainly caused by structures which appear similar to vehicles, like i.e. belonging to shadows, road banks etc.

The influence of the false alarm rate on the calculation of generic traffic parameters can be studied using, e.g., Monte-Carlo simulations. In the following experiment a dense traffic scenario on a multi-lane highway was captured with an image sequence and all car trajectories were manually measured in this sequence, eventually leading to mean velocity profiles for each lane of the highway. Then, a predefined percentage of detections were selected at random positions along the road and contaminated with a specific percentage of random false alarms. Based on these data the velocity profiles were calculated for each lane again and compared to the reference data. As the estimation of the velocity profile depends strongly on the randomly selected positions of the cars, these experiments have been carried out 10000 times, in order to gain a certain statistic about the quality of the estimated profiles. The following table summarizes the RMS values and standard deviations for the estimated velocity profiles depending on the respective detection and false alarm rate.

50% detection rate 5% false alarm rate		50% detection rate 10% false alarm rate		50% detection rate 25% false alarm rate	
RMS [km/h]	σ [km/h]	RMS [km/h]	σ [km/h]	RMS [km/h]	σ [km/h]
5.22	**2.61**	**7.03**	**4.01**	**10.25**	**6.27**
30% detection rate 5% false alarm rate		**30% detection rate 10% false alarm rate**		**30% detection rate 25% false alarm rate**	
RMS [km/h]	σ [km/h]	RMS [km/h]	σ [km/h]	RMS [km/h]	σ [km/h]
5.97	**3.17**	**8.03**	**4.66**	**11.30**	**6.58**

Table 1: Monte-Carlo simulation of reconstruction of velocity profiles depending on detection and false alarm rates

As can be seen, especially the false alarm rate highly influences the quality of the estimates. For instance, it is still possible to reconstruct the velocity profile up to 6km/h ± 3km/h at a detection rate of only 30% when keeping the false alarm rate at 5%.

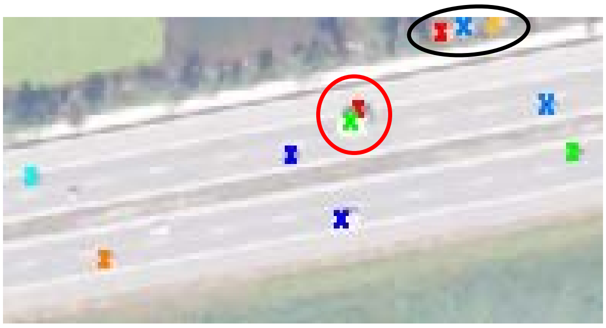

Figure 1: Detection result with false alarms. The red circle indicates redundant objects, the black mark shows objects belonging to the background.

There are mainly two ways of how false alarms are being tracked (see example in Figure 1):

- Collinear motion for redundant objects/features belonging to vehicles (trailer, car shadow or other).
- With zero velocity if objects belong to the background (road bank, shadows of trees etc.)

It is easy to see that these false alarm objects influence statistical traffic data in a manner that may lead to wrong conclusions of the traffic situation or to conflicts in model calculation.

To demonstrate such influence, two examples shall be mentioned:

- In a traffic scenario of a congestion where one lane moves slightly faster than the other (see Figure 2), false alarms belonging to (mainly larger) vehicles of the faster lane concurrently increase the derived density and raise the calculated average velocity. This obviously leads to a conflict in the traffic evaluation. If the false alarms belong to vehicles of the slower lane, the average velocity is lowered and thus implying an even higher vehicle density than there actually is.

- Let us assume a snapshot of a real situation of free flowing traffic with 30 cars moving with an average velocity of about 60 km/h (which corresponds to the speed limit). By assuming 60% completeness and 70% correctness, around 18 cars will be correctly detected and there will be 8 false alarms. In case that the false alarms belong to static background they will obtain a speed 0 km/h. This leads to a calculated average velocity of 41 km/h which implies rather dense traffic and thus feeding the traffic flow model with erroneous input data.

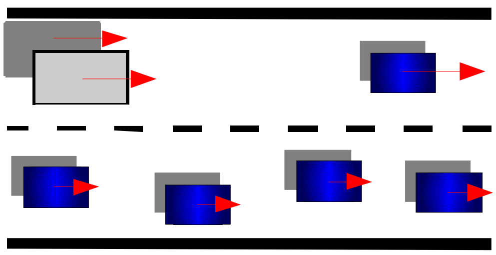

Figure 2: Congested traffic situation with different velocity in each lane.

Therefore, it is desirable to eliminate the false alarms of the initial detection to achieve a better quality of the calculated average velocity.

3. CONCEPT OF REFINEMENT

To improve the initial detection quality, we include generic knowledge about the velocity statistics and geometric layout of traffic flow in typical traffic situations (e.g. “free flowing”, “congestion”, “traffic jam”). To this end, we first track all initial detections and then eliminate the included false alarms based on an analysis of geometric layout and velocity of the trajectories.

3.1 Summary of tracking procedure

Initial vehicle candidates are extracted in the neighborhood of predefined road axes using a blob detection algorithm tuned for color images. Image triplets are then used for tracking, in order to gain a certain redundancy allowing an internal evaluation of the results. A vehicle image model is created by selecting a rectangle around a particular detection. By using the shape-based matching algorithm (Steger, 2001), car hypotheses are found in the successive images. The matching procedure delivers matches in Image 2 and in Image 3. Then, new car image models are created at all hypotheses positions in Image 2 and matched to Image 3. Of course, these matches may contain multiple match results. Finally, all results obtained in Image 3 are checked for consistency including a smoothness criterion of the trajectory to determine the correct combination of the matches. A detailed explanation of this approach can be found in (Lenhart et al., 2008).

The described tracking method is a very robust one delivering correct matches at about 99%, yet it tracks objects of any kind as long as their motion fulfills smoothness constraints similar to those of cars. Thus, trajectories of false initial detections are potentially tracked and also considered as “correct”. Based on the results of the tracking, the refinement is carried out.

3.2 Elimination of redundant objects

A first step to eliminate false detections is to remove redundant objects from the set of detections. These are the kind of objects that belong to vehicles, such as shadows or trailers.
For each pair of detections, the spatial distance is calculated. A search for very small distances delivers candidates for redundant objects. Since candidates may also include vehicles within a passing maneuver, these candidates need to be analyzed for their trajectories. The analysis includes the speed and direction of the determined trajectories and relative direction between the candidates. Identical trajectories and constant relative direction indicates redundant candidates while passing vehicles will have at least a slight difference in their speed or relative orientation.

It is now tested which of the redundant candidates is the car and which is the object to be eliminated. Therefore, a quick test of the gray or color value in the center of the objects is carried out. The darker and less colored object is assumed to be the shadow and is therefore eliminated from the set of detections. In case that both objects have a similar gray or color value, the trailing object is eliminated.

3.3 Knowledge representation of traffic situations

In order to evaluate the velocities of the vehicles we need to formulate our knowledge and expectations about the typical traffic situations such as “free flowing traffic”, “congestions” or “traffic jams”. In dependence of the state of traffic and the location with respect to intersections or traffic lights, different interactions between vehicles occur.

A well substantiated statistical concept like Bayes’ theorem would provide a sound basis for evaluation. However, determining the probability density functions is hardly feasible because extensive and sufficient samples are missing. Hence, it is advisable to avoid a concept that claims statistical integrity.

In contrast to Bayes’ theorem, fuzzy logic offers an intuitive method to represent knowledge of classes by easy parameterization (Zadeh, 1965). It is also frequently applied for modeling car following behavior (Brackstone and McDonald, 1999). Therefore, we decided to use fuzzy logic to describe our knowledge about traffic.

3.3.1 3D fuzzy membership function for active vehicles

Let us define a fuzzy set A that describes vehicles which are actively involved in traffic. Besides normally moving cars, these may be standing vehicles in traffic jams or waiting at red traffic lights or other crossroads.

Since we are only interested in the possibility of an object belonging to A, we neglect the alternative set $\bar{A}$ of inactive objects which may be false alarms of the detection, parking vehicles or erroneous tracks.

For the fuzzy set A, a membership function needs to be defined, indicating the possibility μ_A that a car belongs to A in dependence of its velocity v. However, μ_A also strongly depends on the traffic density D and the distance d from intersections. It is quite obvious, that in a free flowing situation in the middle of road segment the possibility that a car stands still is 0. In contrast to that, zero speed has a rather high possibility near intersections or in jam situations. In order to meet these different traffic situations, we have to consider the conditional possibilities $\mu_A(v|D,d)$. In the sequel, the units for the measures given shall be v [km/h], D [cars/km per lane] and d [m].

First, we should outline the ranges of D and d where μ_A may change significantly. A density of lower or equal to D = 30 corresponds to free flowing traffic while a density of D = 180 represents the maximum density of a traffic jam when there is almost no motion at all (Hall, 1999). Below 30, $\mu_A(v,D|d)$ remains constant.

The interesting range for d is approximately between 150 meters before an intersection because this describes the range where drivers start to brake and 50 meters behind the intersection where drivers accelerate until they reach their desired travel speed. Outside of the range of [-50m;150m], $\mu_A(v,d|D)$ is constant for all values of d.

Over the entire space of v, D and d, this results in a 3D membership function $\mu_A(v,d,D)$. Please note that the values also depend on the road type, speed limits intersection layout. For different road conditions, different functions have to be developed. The mentioned example function refers to a major city road with multiple lanes with a speed limit of 60 km/h and an intersection with a road of equivalent type controlled by traffic lights.

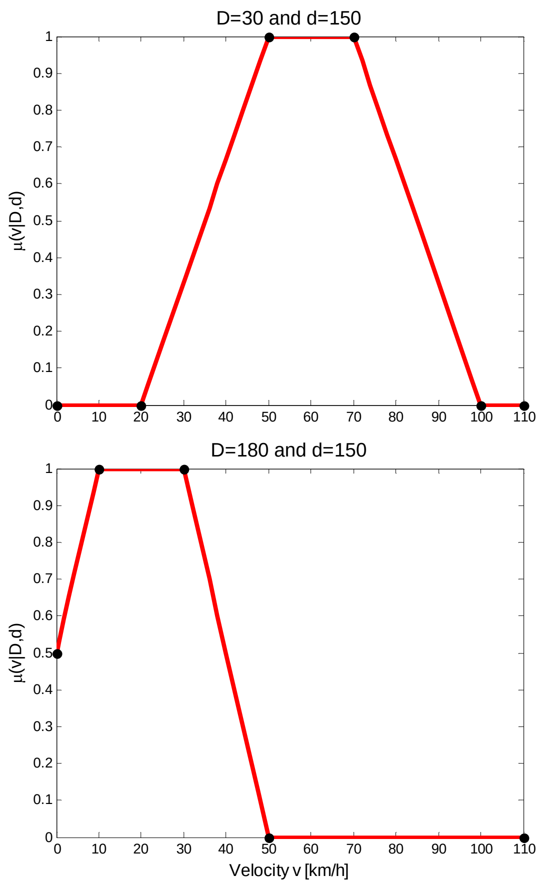

Figure 3: 1D membership functions given a traffic density D = 0 and D = 180 respectively and distance d = 150 with support points (black dots)

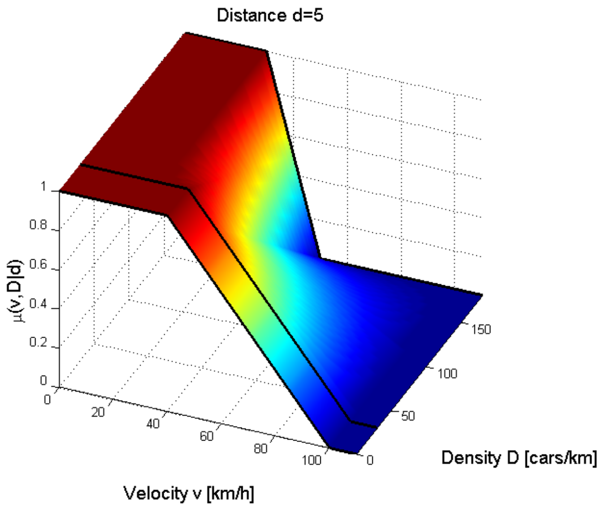

Figure 4: 2D membership function at distance d=5, with 1D support functions (black lines)

To create this 3D function, support points have to be selected. I.e., given an open traffic situation (D = 30) and long distance from an intersection (d = 150) the possibility of a vehicle moving with a speed between 0-20 km/h in shall be 0, while the possibility at the same position in the same traffic situation shall be 1 for velocities between 50-70 and becoming 0 again at v = 100.

A second function depicts a dense traffic situation at the same distance. The possibility for zero velocity shall be 0.5 (since jam cues more likely move slowly forward), while speeds of 10-30 shall be most likely and speeds larger than 50 simply impossible. By linear interpolation between the support points, this results in the functions depicted in Figure 3.

Assuming that the possibilities evolve linearly over the dimension of density, we can derive 2D functions given a certain distance. The function for a position right in front of an intersection is shown in Figure 4.

By linear interpolation along the third axis *d*, we receive a cubic membership function (Figure 5).

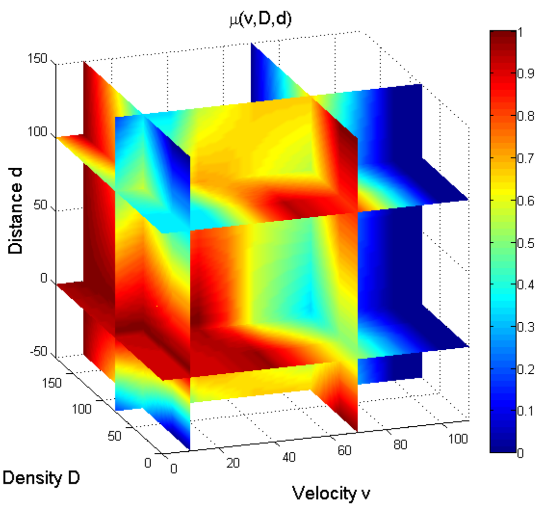

Figure 5: 3D membership function with slices at v=10, v=70, D=80, d=0 and d=100

3.3.2 Evaluation of velocity information

Before the evaluation of the speed, the road is split into sections of length 50m and, near to intersections, of only 20m (sees Figure 6). Every detected object is assigned to one section and contributes to the section density. After the determination of the section density, the possibility μ_A is derived from the above described 3D membership function for each object.

The fuzzy possibility serves a weight in the calculation of a weighted average velocity for each section:

$$\overline{v} = \frac{\sum_i v_i \cdot \mu_A(v_i, D_i, d_i)}{\sum_i \mu_A(v_i, D_i, d_i)} \tag{1}$$

and its standard deviation:

$$\sigma_v^2 = \frac{\sum_i (\overline{v} - v_i)^2 \cdot \mu_A(v_i, D_i, d_i)}{\sum_i \mu_A(v_i, D_i, d_i)} \tag{2}$$

By applying a minimum threshold on the summed up weights of a section, we meet the circumstance that there are only false alarms in a free flow section. If the sum of the weights of a section is below the threshold, all detections of this section are removed.

Finally, objects with a velocity $v_i < \overline{v} - 2 \cdot \sigma_v$ are regarded as outliers and eliminated. Then, a refined and unweighted average velocity is determined from the remaining detections. The resulting distribution is unbiased under the assumption that all false alarms have been eliminated.

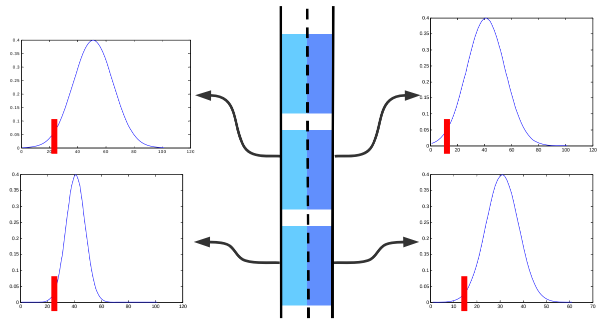

Figure 6: Velocity distribution and cutting-off criteria (red) per road section for each lane

4. RESULTS

The concept has been tested on two different sets of image data so far. The results are shown in Figure 7. One image shows a highway section with free flowing traffic. Here, the detection was carried out by a blob detection algorithm as explained in (Lenhart et al.). In this case, the refinement was able to eliminate all false alarms and redundant objects that arose from the automatic detection.

The second image shows a more complex scene with an urban highway section and an exit leading to an intersection with a traffic light. In this case, the detection has been carried out manually, however, considering a reasonable detection characteristic and quality. In this example, 12 objects have been correctly removed, leaving only one false alarm that could not be eliminated due to faulty tracking.

In both examples, the correctness of the detection could be significantly increased by approximately 30%.

Figure 7: Results of the refinement process. Green: detected and tracked vehicles; Black: redundant objects eliminated by trajectory analysis; Red: false alarms by velocity evaluation. Blue circle: false alarm which has not been eliminated.

5. DISCUSSION AND OUTLOOK

The presented concept shows a possibility to refine detection and tracking results by velocity and trajectory evaluation. This allows a more precise derivation of the average velocity which is fed into traffic models. The benefit of this concept affects exclusively the calculation of the average velocity. It is not possible to counter the problem of low completeness, since missed hits cannot be recovered. However, as has been shown by the simulation in Sect. 2, a low false alarm rate is essential for extrapolating the detection results. For low detection rates, traffic flow parameters can still be estimated with reasonable quality if the false alarm rate is small enough.
Still, many more test scenes have to be evaluated in order to give a more precise measure for the potential of this method.

REFERENCES

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

Brackstone, M., McDonald, M. (1999): Car-following: a historical review. In: *Transportation Research Part F: Traffic Psychology and Behaviour*, Volume 2, 181 - 196, 1999.

Ernst, I., Sujew, S., Thiessenhusen, K-U., Hetscher, M., Raßmann, S., Ruhé, M. (2003): LUMOS - Airborne Traffic Monitoring System. In: *Proceedings of the IEEE 6th International Conference on Intelligent Transportation Systems*, Shanghai (China), 12-15 Oct. 2003.

Ernst, I., Hetscher, M. Zuev, S., Thiessenhusen, Kai-Uwe; Ruhé, M. (2005): New approaches for real time traffic data acquisition with airborne systems. In: Stilla, U., Rottensteiner, F., Hinz, S. (eds): *CMRT05*, IAPRS, Vol. 36, Part 3/W24, 69-73.

Everaerts, J., Lewyckyj, N., Fransaer, D. (2004): PEGASUS: Design of a Stratospheric Long Endurance UAV System for Remote Sensing. In: Altan, O.(ed): *Proceedings of the 20th ISPRS Congress*, Istanbul 2004, IAPRS, Vol. 35, Part B2, 29-33.

Hinz, S., Bamler, R., Stilla, U. (2006): *ISPRS Journal Theme Issue: "Airborne and Spaceborne Traffic Monitoring"*. International Journal of Photogrammetry and Remote Sensing, 61(3/4).

Kurz, F., Charmette, B., Suri, S., Rosenbaum, D., Spangler, M., Leonhardt, A., Bachleitner, M., Stätter, R., Reinartz, P. (2007): Automatic Traffic Monitoring with Airborne Wide-Angle Digital Camera System for Estimation of Travel Times. In: Stilla, U., Mayer, H., Rottensteiner, F., Heipke, C., Hinz, S. (eds): *PIA07*, IAPRS, Vol. 36, Part 3/W49A, 83-89.

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,* 2008.

Ruhé, M., Hipp, E., Kühne, R. (2007): A model for new data - using air borne traffic flow measurement for traffic forecast. In: *TRISTAN VI* (Sixth Triennial Symposium on Transportation Analysis), Phuket, Thailand, 10-15 June 2007.

Lenhart, D., Hinz, S., Leitloff, J., Stilla, U. (2008): Automatic Traffic Monitoring based on Aerial Image Sequences. In: *Pattern Recognition and Image Analysis, Volume 18, 3*, p. 400-405, Springer, 2008.

Hall, F. (1999): Traffic Stream Characteristics. In: *Traffic Flow Theory - A State-of-the-Art Report*, updated version of Transportation Research Board Special Report 165. Chapter 2. http://www.tfhrc.gov/its/tft/chap2.pdf (accessed 27. March, 2009)

Steger, C. (2001): Similarity measures for occlusion, clutter, and illumination invariant object recognition. In: B. Radig and S. Florczyk (eds.) *Pattern Recognition,* DAGM 2001, LNCS 2191, Springer Verlag, 148–154.

Stilla, U., Rottensteiner, F., Hinz, S.: *CMRT05*, IAPRS, Vol. 36, Part 3/W24.

Zadeh, L. (1965): Fuzzy sets. In: Information and Control, Volume 8, 338 - 353, 1965.