Photogrammetrie • Fernerkundung • Geoinformation 5/2009, S. 393–406

Automatic Estimation of Vehicle Activity from Airborne Thermal Infrared Video of Urban Areas by Trajectory Classification

Wei Yao, München, Stefan Hinz, Karlsruhe & Uwe Stilla, München

Keywords: Airborne thermal IR video, detection, video co-registration, trajectory grouping, movement estimation.

Summary: Analysis of traffic data plays an important role in urban and spatial planning. Thermal Infrared (TIR) video cameras have capabilities to operate at day and night and to acquire the scene sampled with video frame rate. In this paper a strategy for the estimation of vehicle motion and the assessment of traffic activity from airborne TIR video is presented. In contrast to other approaches we handle detecting and tracking vehicles in the video separately, because moving as well as stationary vehicles are intended to be detected. Firstly, vehicles are detected in single frames of the video. Additionally, tie points are detected for co-registration and compensation the sensor movement. Afterwards, a stepwise grouping of image points considering temporal consistence and geometric relation is carried out to determine the vehicle trajectories and classify them into stationary, moving and uncertain dynamical categories. The vehicles are then integrated into the classes "moving," "stationary" and "uncertain" categories. Additionally, in consideration of matching vehicle-related image patches for moving vehicles, the topology of the trajectories are investigated and optimized in order to eliminate disturbances and estimate velocities. The algorithms were tested with video sequence of urban areas in nadir-view and oblique-view. The correctness of the results is achieved higher than 75% for both views.

Zusammenfassung: *Automatische Schätzung der Vehikel Aktivität aus luftgetragenen thermalen infraroten Videos urbaner Bereiche durch Trajektorienklassifikation.* Die Analyse von Verkehrsdaten spielt eine wichtige Rolle bei der Stadt- und Raumplanung. Thermal Infrarot (TIR)-Kameras bieten die Möglichkeit Szenen bei Tag und Nacht im Videotakt zu erfassen. In diesem Beitrag werden eine automatische Strategie zur Schätzungen von Fahrzeugbewegung und die Bewertung der Verkehrsaktivität aus flugzeuggetragenen IR-Videodaten präsentiert. Bei diesem Ansatz werden im Gegensatz zu herkömmlichen Verfahren Detektion und Verfolgung von Fahrzeug in den Videodaten getrennt behandelt, da auch stehende Fahrzeuge gefunden werden sollen. Zunächst werden Fahrzeuge in den Einzelbildern der Videosequenz detektiert. Zusätzlich werden Verknüpfungspunkte detektiert, um eine Koregistrierung der Bilder durchzuführen und die Sensorbewegung zu kompensieren. Anschließend wird unter Berücksichtigung zeitlicher Konsistenz und geometrischer Zusammenhänge eine schrittweise Gruppierung von Bildpunkten durchgeführt, um den dynamische Status des Fahrzeuges zu detektieren. Die Fahrzeuge werden dann in die Klassen „bewegt", „unbewegt" und „unsicher" eingeteilt. Zusätzlich auch mit Berücksichtigung von Matching der fahrzeugbezogenen Bildmatrizen werden für die bewegten Fahrzeuge die Topologie der Trajektorien untersucht und optimiert, um Störungen zu beseitigen und Geschwindigkeiten zu schätzen. Die Algorithmen wurden mit Bildsequenzen in Nadirsicht und in Schrägsicht von innerstädtischen Gebieten getestet. Die Korrektheit der Ergebnisse erreicht für beide Sichten Werte höher 75%.

DOI: 10.1127/1432-8364/2009/0028

1432-8364/09/0028 $ 3.50

1 Introduction

Traffic-monitoring systems rely on sensors to acquire traffic information. In the last decade many ground-based sensors, e. g., loop detectors, bridge sensors and stationary cameras have been widely used and extensively studied (Hu et al. 2004). Airborne-video data acquisition for traffic-parameter estimation has been explored as an alternative to conventional data-collection methods (Ernst et al. 2005; Rosenbaum et al. 2008), because it may enable covering a relatively broad area and potentially derive additional parameters such as travel time, relative velocity, vehicle trajectory, etc (Shastry & Schowengerdt 2005; Cohen & Medioni 1998; Reinartz et al. 2006). Thermal IR cameras provide the night vision capability and are able to capture the traffic situation at day and night. Generally, temperature measurement gives an important cue for the activity of cars, even though vehicles are in use but do not move due to a traffic jam or waiting in front of a red traffic light. Stationary vehicles often appear as cold spots (dark) on the warmer road surface (bright). But warm parts (e. g., exhaust) of active vehicles appear as hot spots (bright) (Stilla et al. 2004; Ernst et al. 2005).

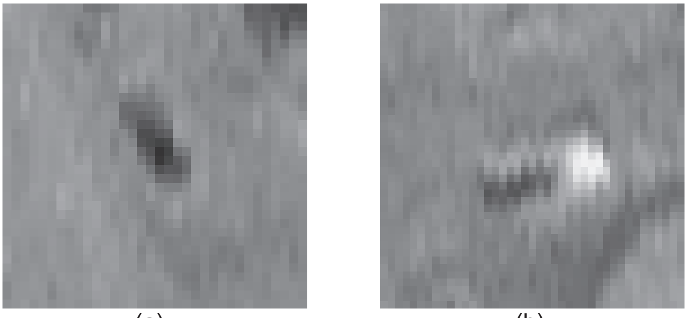
Since the videos are taken from a moving platform, the simple optical flow estimation cannot directly be used to detect object motion, we have to distinguish the sensor movement from true object movement in the scene in order to characterize traffic activity. A number of approaches have recently been proposed. However, they are either only constrained to automatically detect vehicles or vehicle queues in the IR images of dense city areas or only focused on tracking moving vehicles and estimating their movement from airborne IR video. For example, Stilla & Michaelsen (2002) have developed a method of detecting single vehicle in the airborne IR images of urban areas based on spot-filtering. In (Hinz & Stilla 2006) a detector for extracting single vehicles and vehicle queues combing global and local context is introduced, we can hardly get information about time from the single image. Concerning TIR video, Kirchhof & Stilla (2006) have applied planar homograph as geometric tool to co-register the video data and attempted to detect and track moving objects by analyzing the motion channel. Both Michaelsen & Stilla (2004) and Yilmaz et al. (2003) have analyzed and accessed different methods for pose estimation from oblique airborne videos in order to optimize processing chain for specific scene reconstruction and tracking moving objects. An up-to-date trial on multiple moving object detection is presented in (Yao et al. 2008a, 2008b and Leitloff et al. 2007), but most of them are limited to dealing with datasets aquaired in nadir-view mode.

All previous works mentioned above can be regarded as the foundations or components of traffic monitoring system from airborne TIR platforms, which only addressed individual parts of the issue. In this work we will give an integrated contemplation on the vehicle detection and the determination of their dynamical status, an off-line strategy of accessing traffic scene by airborne TIR video is proposed to locate all vehicles at first and then derive their dynamical statuses based on the trajectory construction and classification. The processing chain is implemented in the context of treating detection and tracking of vehicles separately, which will be presented and discussed in following sections.

2 Detection of Vehicles in Every Frame

In order to detect vehicles in single images an automatic approach (Stilla & Michaelsen 2002) is used. Because of low contrast and noisy characteristic of TIR imagery, several parking cars along two margins of the road have failed to be detected. The detection rate was improved by fusing the map data which is usually available for urban areas. The algorithm of single vehicle detection used here is based on following two assumptions: (i) the image is searched for cold-spots (black) which represent the vehicles themselves; (ii) for a moving vehicle, a hot-spot presenting the warmed engine bonnet must exist around a cold-spot (cf. Fig. 1).

Since we want to estimate traffic parameters in this work, the detection of moving vehicles is brought into focus. While applying this algorithm for vehicle detection, some imagin-



(a) (b)

Fig. 1: Cutouts of thermal IR-image related to vehicle (34 × 34 pixels), a) stationary vehicle b) moving vehicle.

able problems emerge instantly, e. g., a car had started just recently, the engine bonnet has not warmed up yet; the hot spot of vehicle is missing. Therefore it is at first not possible to distinguish stationary vehicles from moving vehicles. It is only about the detection of vehicles in this step. Vehicle detection results from one single TIR frame are shown in Fig. 2. All of

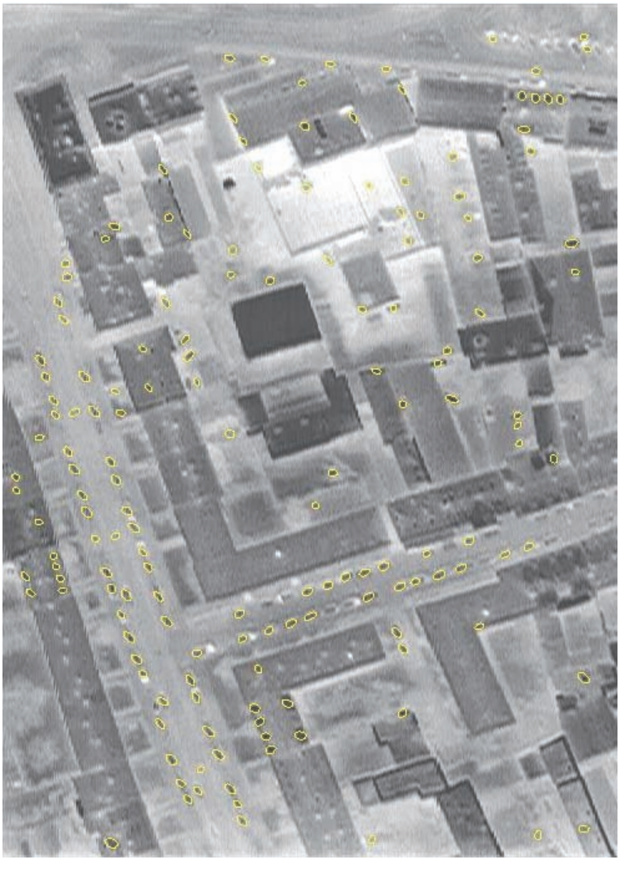


Fig. 2: Result of vehicle detector on single TIR image (494 × 686 pixels). Red-blue crosses mark the positions of detected vehicles.

test images have been processed successively to obtain the (*row, col*) coordinates of vehicles.

3 Processing of Image Sequence

The approach of processing airborne thermal infrared video consists of two major steps. The first step is video stabilization, comprising the co-registration of certain image pairs in the video. In the second step all detected vehicles in each frame of video will be transformed into a unified coordinate frame, e. g., the domain of the last frame of the video. The result is an image mosaic of the video sequence. The processing chain is depicted in Fig. 3. After it is initialized with necessary parameters, an operation loop is carried out to transform the vehicle points until the last frame of the TIR video data is reached. The details of the sub-steps are to be discussed as follows.

3.1 *Video Stabilization*

For establishing the geometric relation between every two TIR video frames, image stabilization has been chosen to perform this task. Image stabilization consists of registering the two images and computing the geometric transformation $\boldsymbol{T}$ that warps the image I_1 such that it aligns with the reference image I_0. The warping of images can be based on, e. g., (i) planar homography or (ii) affine transformation.

There are also two ways to implement the co-registration within the video frames – sequential and direct orders. We compared different results generated by various combinations of geometric tools and transformations and selected the best one – sequentially affine transformation. Being independent of geometric tools, the co-registration procedure can be divided into two main processes:

(1) Initial relative orientation
  a) Pre-processing steps namely initialize some empirical parameters for operators (e. g., offset of the search window, thresh-

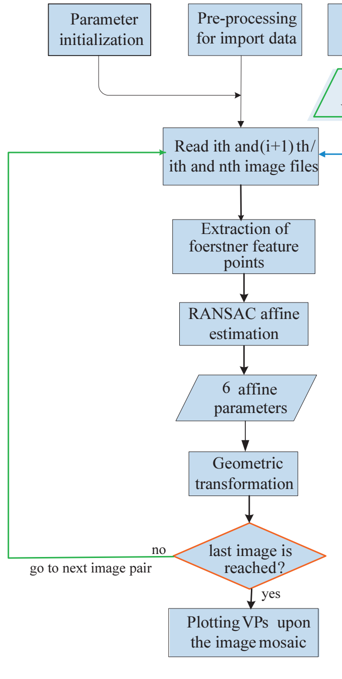


Fig. 3: Workflow of image sequences processing using affine transformation model.

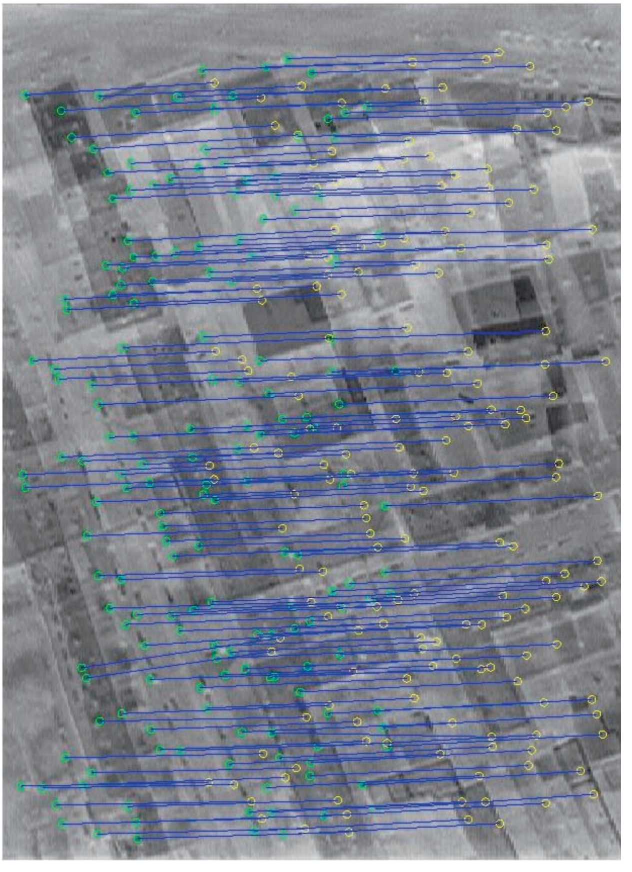


Fig. 4: Point correspondence found by RANSAC relative orientation plotted on the two overlapping frames. Green points are from 10th frame; yellow points are from 30th frame.

old for gray value matching etc.) and read the coordinates of vehicles. Since these initial parameters have a great influence over the quality of co-registration, they should be selected carefully.

b) Feature points are extracted with the operation introduced by FÖRSTNER & GÜLCH (1987).

The image domain can be reduced upon road region, so that it can better fit the prerequisite of planar homograph, but at the expense of the quantity of robust feature points. Therefore, this restriction is ignored in the case of directly transformation order due to low overlapping rate.

c) Compute an affine transformation matrix between two images based on a randomized search algorithm (RANSAC).

Feature points of both input images are read in to serve as candidates of corresponding points based on gray-value matching. Once the initial matching is complete, RANSAC is used to determine the affine transformation matrix. Matched input points filling the condition of RANSAC are output as corresponding points for affine relation (cf. Fig. 4).

(2) Refinement of the orientation

After initial relative orientation, the estimated parameters should be refined. The least squares bundle adjustment is the classical method and usually delivers the best results. Depending on the scene characteristics also some simple transformations, e. g., affine transformation, yield similar results, especially when the baseline between the images is small.

3.2 *Transformation of Vehicle Points*

After two subsequent images are co-registered, we have to sequentially project vehicle coordinates into reference image domain using affine transformation matrices derived above.

Finally, all of the vehicles points detected from different images sampled temporally have been transformed into the coordinate frame of the last image which serves as the reference frame in our experiments, and also been plotted on the mosaicked image which leads to a map. In this map, the moving vehicles are supposed to build their trajectory distributed in curvilinear form, whereas the stationary vehicles ought to accumulate nearly in the same place as compact cluster and slightly shift. Then, we can analyze and measure vehicle trajectories on the basis of this map. In this map we have plotted the position of each transformed vehicle in blue color (cf. Fig. 5).

4 Automatic Characterization of Vehicle Movement

In this step, the objective is to automate the analysis process of interpreting and inferring the vehicle movement by means of information acquired by airborne TIR camera. The stabilized map of vehicle detection is generated by last two steps; where the vehicles detected from single frames are projected into the coordinate system of the reference frame and depicted as blue cross. Multiple instances of one vehicle entity, corresponding to different discrete time tags of frame recording, tend to build and describe the temporal behavior of the vehicle trajectory. In order to characterize and analyze the traffic activity, it is required to reconstruct the trajectory of vehicles in this map, and to label them as moving or stationary ones. Our strategy to perform this task features a stepwise operational concept incorporating split-and-merge of trajectory based on temporal consistence and geometric rela-

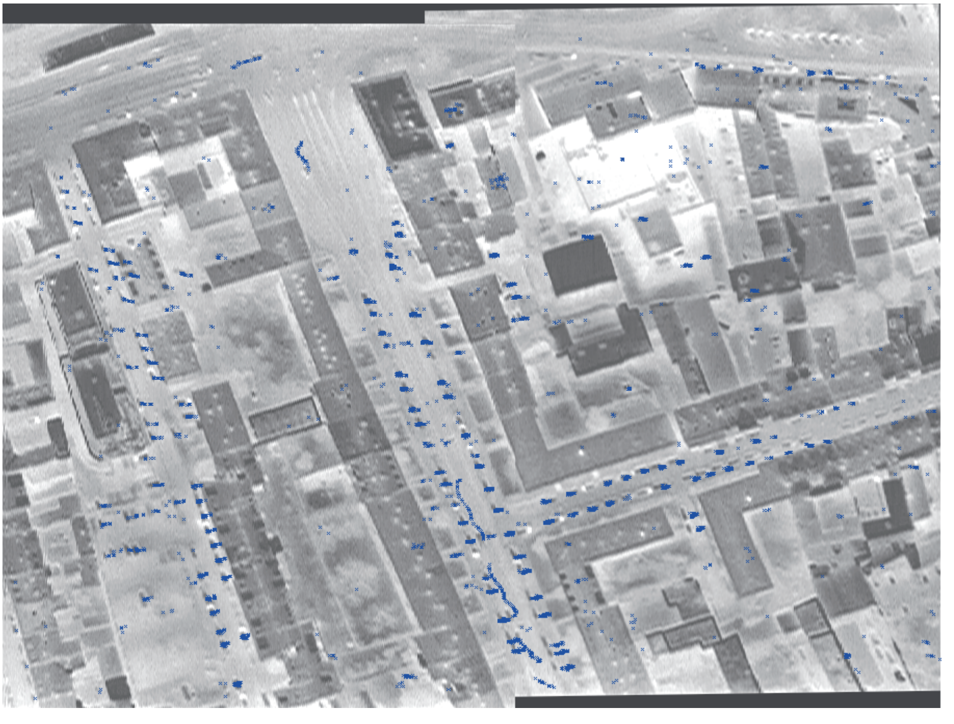


Fig. 5: Stabilized map of detected vehicles plotted overlaid upon mosaicked TIR video by sequential affine transformation.

tion. The correspondence relation between detected single vehicles of each frame is to be re-established here. We do not use image differencing and matching to characterize the moving object just as normal methods, but rather perform detection and tracking of vehicle separately.

4.1 *Preliminary Classification*

The first step of our strategy is to classify the vehicle region into four different classes by means of clustering analysis assisted by temporal consistence criterion. The vehicle region map is generated by labeling connected components in the stabilized map of vehicle detection, which can be viewed here as binary image when using an image of single intensity as background. Afterwards initial vehicle regions for trajectory delineation and classification are created; they can serve as trajectory candidates for single vehicle entity. These initial regions are to undergo the classification process according to density feature measure describing the physical compactness of vehicle clusters. This feature measure *FM* is defined as follows:

$$FM = \frac{L_{maj\_axis}}{N_{vp}} * \frac{10}{Ind_{max} - Ind_{min}} \tag{1}$$

where L_{maj_axis} is the length of major axis of region; N_{vp} is the number of vehicle points included in a vehicle region; Ind_{Max} and Ind_{Min} are the maximum/minimum index within a region, respectively.

If $FM >= 1$ and $Real\_area > 1$, classified as candidates for moving vehicle; If $FM < 1$ and $Real\_area > 1$, classified as candidates for stationary vehicle; If $Real\_area = 1$, classified as single vehicle class. A joint consideration with compatibility of temporal index within single vehicle regions is necessary. Because vehicle instances from two vehicles in reality may merge into one initial vehicle region (hybrid

Fig. 6: Vehicle region map after preliminary classification, green: stationary vehicle; blue: single vehicle; red: moving vehicle; black: hybrid vehicle.

class) displayed here, so it has to be delivered to the split process further. The resulting map after this step is showed in Fig. 6.

4.2 *Refinement of Classification Result*

Due to unavoidable existence of co-registration and detection errors, vehicle points belonging to the stationary category usually do not accumulate in connected cluster. In this intermediate step we merge the green category of vehicle region map generated from last step, and analyze the white points to split them into independent vehicle regions. For the analysis of the stationary category, we take these regions as seed point, and then do a search in a close surrounding area, in which red, blue and green vehicles to be analyzed concerning temporal and geometric accordance with the seed region. In order to generate hypotheses for stationary vehicles, we have to verify them via image matching; then, those regions confirmed by two operations above will be accepted as stationary vehicle and aggregated with the seed region to build new green class labeled as one region.

One usually has to restrict the amount or eccentricity of green regions, after or while merging green region with another stationary class, so as to exclude some ones being lacking of temporal completeness of trajectory or ones of inordinate trajectory elongation. The advantage of this step is that the problem domain and complexity can be rather reduced that we can focus on individual vehicle categories by stepwise operation.

4.3 *Grouping and Extracting Vehicle Trajectory*

Based on results generated by the last step, the green regions that are supposed to be the stationary vehicle class are relatively stable and then we focus on the red vehicle region, attempting to group the fragmented regions into reasonable trajectory of a moving vehicle.

The grouping algorithm is implemented by sequential searching based on the joint analysis of the geometric relation and temporal consistence. It starts from an arbitrary red vehicle region and guides the search direction towards the major axis of vehicle region. The criterion for testing the compatibility between temporal index and geometry is formulated as below:

$T_i \cap T_j = \{O\}$ *and* min $\{T_i\}$ – max $\{T_j\}$ *or* min $\{T_j\}$ – max $\{T_i\}$ must be consistent with the distance between border points of each region: $d_{ij} = |R_{i,min} - R_{j,max}|$ or $|R_{j,min} - R_{i,max}|$, i.e,

$$ThreshV - \Delta t \le d_{ij} / (T_{i,\min} - T_{j,\max}) \le ThreshV + \Delta t \quad (2)$$

where T_i and T_j are the temporal index sets of two vehicle regions i and j; $R_{i,min}$, $R_{j,max}$ are the border points of each region; d_{ij} is the distance between border points (usually max or min temporal index) of each region; *ThreshV* is the threshold related to the assumed vehicle velocity; Δt is the allowable deviation affected by the detection and co-registration accuracy, $(T_{i,min} - T_{j,max})$ can be replaced by $(T_{j,min} - T_{i,max})$.

After examining the assumed accordance of temporal consistence with geometric dis-

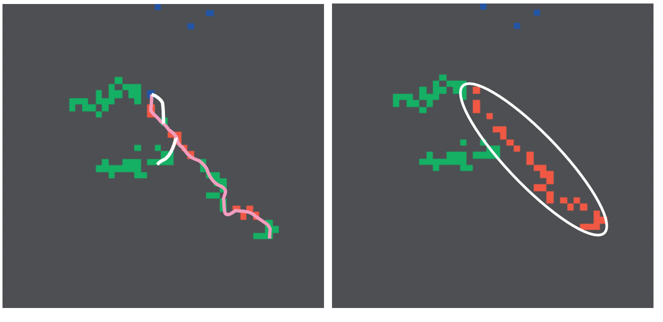


Fig. 7: Zoom into the local section of vehicle region map marked by dotted box in Fig. 6, before (a) and after (b) trajectory grouping.

tance, we extend and link the adjacent vehicle regions to create the complete trajectory of vehicle entities by taking into account the topological property of them. It is required to achieve an optimized distribution for the trajectory of each moving vehicle. The graph description (cf. Fig. 8) of vehicle regions (cf. Fig. 7) can be established to support this task. The edge of the graph is constructed to connect each two vehicle regions the instances from which guarantee a good image match. Then, the extraction of vehicle trajectory amounts to find an optimal path along each connected nodes in the graph. Defining a criterion to characterize an optimal path requires allocating a cost value to each edge. We also have to consider the relation between each node evaluated above, since the nodes describing the same vehicle entity are likely to demonstrate the best accordance of temporal consistence with geometric configurations. Therefore we assign for each edge connecting region i to j following cost:

$$W_{ij} = \frac{C_{ij}}{1+(d_{ij}/(T_{i,\min}-T_{j,\max})-ThreshV)^2+(\theta_i-\theta_{ji})^2} \tag{3}$$

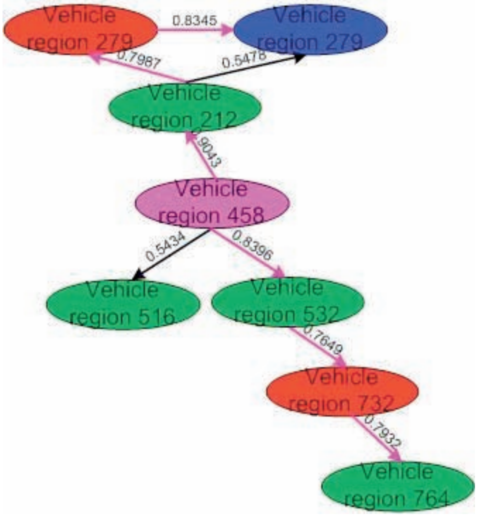


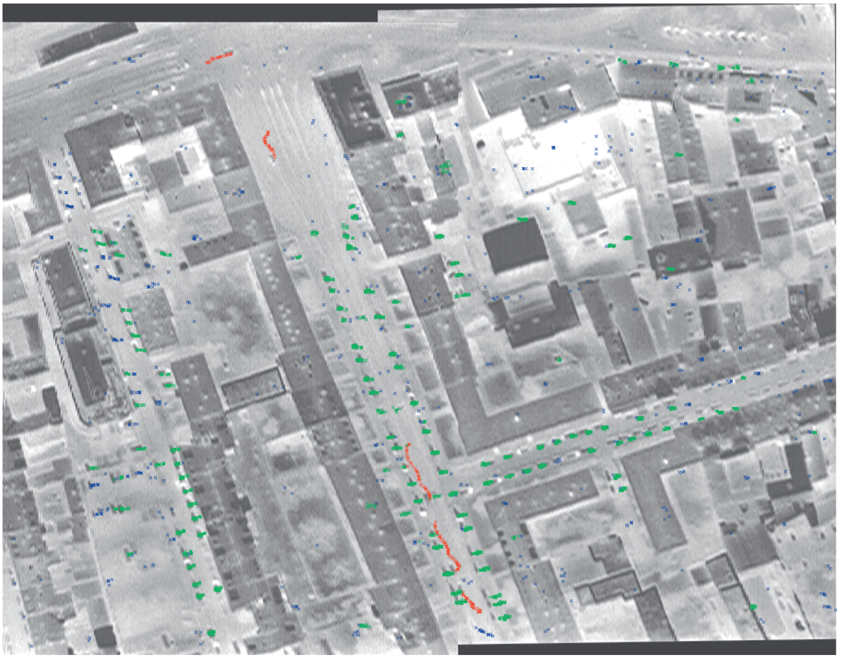
Fig. 8: Graph representation and analysis of vehicle regions in Fig. 7, pink lines denote the optimal path within the vehicle region graph starting from the seed node (pink).

where, C_{ij} is the cross-correlation between vehicle image regions i and j, and θ_i,θ_{ij} represent the orientation angles of major axis vector of region i and connection vector from region i to j.

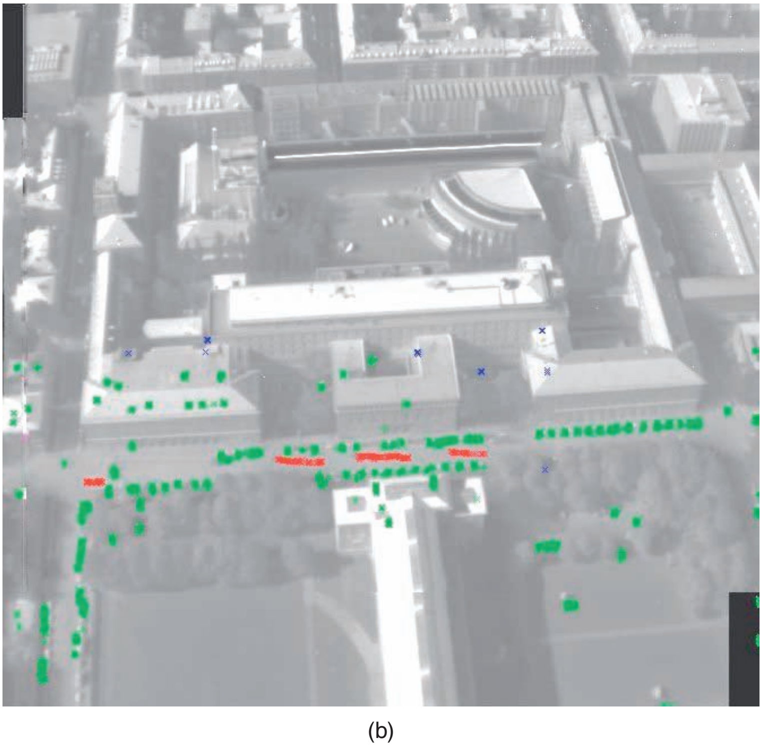
Generally, the trajectory of moving vehicles is assumed to be resolved by adequate temporal resolution, which means it should consist of enough vehicle instances detected from at least 40 % of the total frames; otherwise these trajectories have to be assigned to the vehicle object of uncertain status (behavior). Finally, all of vehicle points are grouped to vehicle entities that are classified into three categories regarding the movement status.

5 Results

We used two datasets of TIR video captured over dense build-up areas, including two main roads with moving and parking vehicles upon them, to test our algorithm proposed above. The only difference between them lies in the view angle. The first dataset was acquired under normal nadir-view, while the other dataset was recorded in forward-looking mode. Fig. 9 illustrates the vehicle classification results concerning movement of three categories: red: moving vehicle; green: stationary vehicle; blue: uncertain. For both test data, it can be seen that most of the parking vehicles along road sides located in the centre area covered by the TIR video are detected and classified into the stationary class, and a proportion of the stationary class vehicles are represented by the parking vehicles, which is reasonable for both test scenes. Due to similar appearance to vehicle, some vehicles of green class are falsely detected among the building roofs, which correspond to chimneys or small dormers in reality. For first dataset, only five moving vehicles with distinct moving trajectory have been found, while four moving ones for second test data are extracted. Although the number of extracted moving vehicles is much fewer compared to the stationary ones, it has represented essential dynamical information for regional traffic. The uncertain class contains either vehicle anomaly generated in the detection step or vehicle entities whose trajectory cannot be resolved by available temporal



(a)



(b)

Fig. 9: Vehicle classification result in respect of movement: a) nadir view dataset, b) oblique-view dataset.

resolution, e. g., vehicles located near the margin of panorama creating by mosaicking.

Afterwards, we tried to derive the velocity for moving vehicles based on their trajectories. Basic information about infrared video dataset used here can be acquired in advance:

– Pixel size (GSD) of the TIR image,
  1. Nadir-view dataset: 0.5 m,
  2. Oblique-view dataset: ca. 0.37 m at the main horizontal road area.
– Two test urban areas are covered by
  (i) Nadir-view dataset: 51 images in all, FPS = 25 frames/sec, so the duration of flight $\Delta t = 2.04$ s.
  (ii) Oblique-view dataset: 55 images in all, FPS = 25 frames/sec, so the duration of flight $\Delta t = 2.2$ s

The length of car's trajectories is obtained via sample pixel coordinates, which we have selected and read out from Fig. 10 empirically, here the all trajectory curves are approximated with 5 sample points.

It yields the following averaged velocities of vehicle:

Dataset I: V1 = 58 km/h; V2 = 51km/h; V3 = 42 km/h; V4 = 36km/h; V5 = 41 km/h,
Dataset II: V1 = 24.9 km/h; V2 = 28.6 km/h; V3 = 20.5 km/h; V4 = 12.8 km/h.

The detection and classification of stationary and moving vehicles for movement indication is evaluated in terms of completeness and correctness against reference data, respectively. Due to the lack of the simultaneously captured ground truth, the reference data used for this evaluation has been manually acquired from the same data set as used for extraction. Hence, one has to keep in mind that the above mentioned values refer to the capabilities of a human operator working with such kind of imagery.

Details of the numerical evaluation of the test video are summarized in Tab. 1 and 2. The velocity of moving vehicles derived by analyzing the trajectories is also not able to be verified strictly. However, for first TIR dataset the values lie within the velocity interval allowed in the city area, and so are plausible. For second dataset, since the oblique-view geometry of the TIR camera has further deteriorated the data quality, a lower completeness of vehicle detection particularly for moving class is shown. The detected moving vehicles travelled obviously slower than normal ones in the city, as they were approaching the road crossing when traffic light was red; second moving vehicle exhibits a bit higher velocity than first one, as the first one travelled behind a heavy truck that is not detected and had to brake early and strongly in order to avoid potential dangerous situations; the third one has a low averaged velocity due to tree occlusion even worsen by the oblique-view mode; the fourth moving vehicle became nearly stagnant, it is supposed that it was then slowing down to stop waiting for left-turn.

For the vehicle detection, certain errors could emerge: (i) Over-segmentation, which means that a mapped vehicle is detected as two or more segments. (ii) Dormer and house corners are falsely extracted due to the similar appearance and lack of exploiting context information. So we can use GIS-database information to restrict the search of vehicles only to roads and/or parking lots in order to further increase the correctness of vehicle detection. For the co-registration of video sequences, planar homography were also tested to wrap the images. However, it delivered worse results than the affine model in our cases. It can be stated that two reasons could be lead to this situation. The first one is that fewer freedom

Tab. 1: Evaluation of vehicle movement indication for nadir-view dataset.

Evaluation criteria	stationary vehicle	moving vehicle
Correct decisions	112	5
False alarms	15	0
Missing decisions	26	1
Completeness [%]	81.2%	83.3%
Correctness [%]	88.2%	100%

Tab. 2: Evaluation of vehicle movement indication for oblique-view dataset.

Evaluation criteria	stationary vehicle	moving vehicle
Correct decisions	72	4
False alarms	19	0
Missing decisions	23	3
Completeness [%]	79.1%	57.1%
Correctness [%]	75.8%	100%

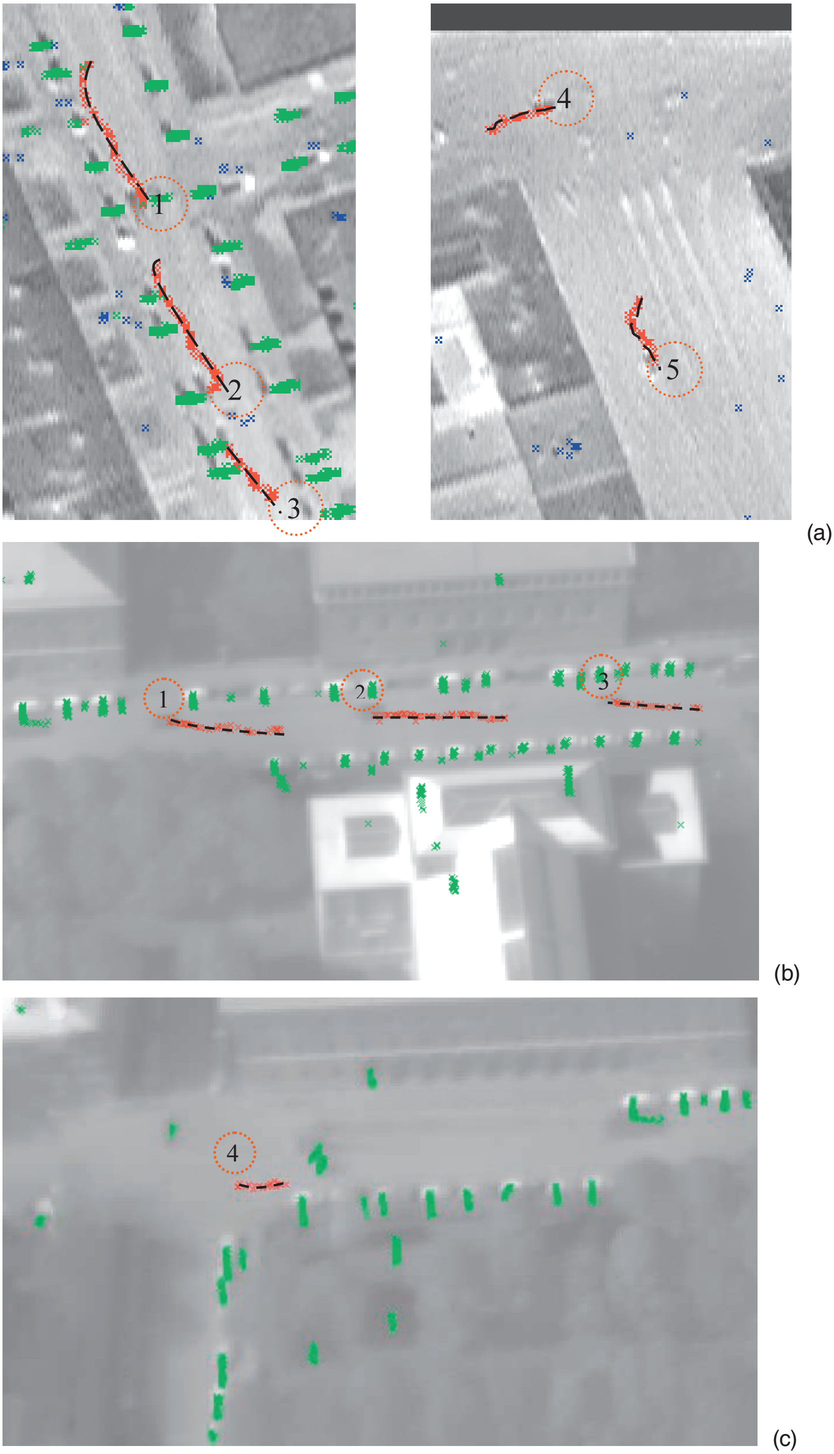


Fig. 10: Zoom into the trajectories of moving cars and their approximated curves: a) nadir-view dataset, b) and c) oblique-view dataset.

degrees allowed by affine transformation put stronger constraints on the scene structure; the second one is that the precondition of planarity for homography is not well fulfilled in the test scenes.

6 Conclusions and Future Work

In this work we have addressed issues related to the automatic analysis of airborne TIR video for vehicle movement analysis. Vehicles are automatically detected and distinguished with respect to their motion status. The vehicle velocity can be derived subsequently as an important parameter for traffic flow analysis. The proposed algorithm is driven by a stepwise operational concept. Following the compensation of sensor movement, the stabilized map of detected vehicles from different temporal tags is examined and multiple vehicle instances are grouped to construct each trajectory of vehicle entities based on temporal consistence and geometric relation. This process is realized by analyzing the distribution properties of vehicle instances under uniform temporal-spatial framework and by optimizing the trajectory topology supplemented by image matching. An implementation of the algorithm on two datasets of different view-angles has delivered us promising results, especially in terms of the detection correctness of both moving and stationary vehicles. The velocity of detected moving vehicles lies within the reasonable interval allowed in urban areas. Future works can be put on improving vehicle detector in view of the complex emissivity of objects in infrared spectrum and on exploiting physically calibrated thermal information to directly extract the object motion. Additionally, the co-registration within this strategy puts unmerited weight on the reference frame that is the last frame of the videos, it would be much better to have the reference not attached to one particular frame but to a geo-referenced system fixed to the ground.

Acknowledgements

The authors are grateful to Fraunhofer – Institut für Optronik und Mustererkennung FOM for providing the test data.

References

Angel, A., Hickman, M., Mirchandani, P. & Chandnani, D., 2002: Methods of Traffic Data Collection using Aerial Video. – 5th International Conference on Intelligent Transportation Systems: 31–36.

Cohen, I. & Medioni, G., 1998: Detection and Tracking of Objects in Airborne Video Imagery. – Workshop on Interpretation of Visual Motion, CVPR98.

Ernst, I., Sujew, S., Thiessenhusen, K-U., Hetscher, M., Rassmann, S. & Ruhé, M., 2003: LUMOS – Airborne Traffic Monitoring System. – 6th International Conference On Intelligent Transportation Systems.

Ernst, I., Hetschler, M., Lehmann, S., Lippok, A. & Ruhé, M., 2005: Use of GIS Methodology for Online Urban Traffic Monitoring. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (8/W27).

Förstner, W. & Gülch, E., 1987: A Fast Operator for Detection and Precise Location of Distinct Point, Corners and Centres of Circular Features. – ISPRS Conference on Fast Processing of Photogrammetric Data: 281–305.

Hinz, S. & Stilla, U., 2006: Car Detection in Aerial Thermal Images by Local and Global Evidence Accumulation. – Pattern Recognition Letters **27:** 308–315.

Hu, W., Tan, T., Wang, L. & Maybank, S., 2005: A Survey on Visual Surveillance of Object Motion and Behaviors. – IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews **35** (3): 435–441

Leitloff, L., Hinz, S. & Stilla, U., 2007: Inferring Traffic Activity from Satellite Images. – International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences **36 (**3/W49B).

Kirchhof, M. & Stilla, U., 2006: Detection of Moving Objects in Airborne Thermal Videos. – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3–4): 187–196.

Michaelsen, E. & Stilla, U., 2003: Estimation of Vehicle Movement in Urban Areas from Thermal Video Sequences. – 2nd GRSS/ISPRS Joint

Workshop on Remote Sensing and data fusion on urban areas: 105–109.

Michaelsen, E. & Stilla, U., 2004: Pose Estimation from Airborne Video Sequences using a Structural Approach for the Construction of Homographies and Fundamental Matrices. – Structural, syntactic, and statistical pattern recognition: S+SSPR, LNCS 3138, Springer, Berlin: 486–494.

Paska, E. & Toth, C., 2004: A Performance Analysis on Vehicle Detection from Remotely Sensed Imagery. – ASPRS Annual Conference.

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 **61** (3–4): 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.

Shastry, A.C. & Schowengerdt, R.A., 2005: Airborne Video Registration and Traffic-Flow Parameter Estimation. – IEEE Transactions on Intelligent Transportation Systems **6** (4): 391–405.

Stilla, U. & Michaelsen, E., 2002: Estimating Vehicle Activity using Thermal Image Sequences and Maps. – International Archives of Photogrammetry and Remote Sensing **34** (4).

Stilla, U., Michaelsen, E., Soergel, U., Hinz, S. & Ender, H.J., 2004: Airborne Monitoring of Vehicle Activity in Urban Areas. – International Archives of Photogrammetry and Remote Sensing **35** (B3): 973–979.

Yao, F., Sekmen, A. & Malkani, M., 2008a: A Novel Method for Real-time Multiple Moving Targets Detection from Moving IR Camera. – 19th International Conference on Pattern Recognition – ICPR: 1–4.

Yao, W., Hinz, S. & Stilla, U., 2008b: Automatic Analysis of Traffic Scenario from Airborne Thermal Infrared Video. – International Archives of Photogrammetry, Remote Sensing and Spatial Geoinformation Sciences **37** (B3A): 223–228.

Yilmaz, A., Shafique, K. & Shah, M., 2003: Target Tracking in Airborne Forward Looking Infrared Imagery. – Image and Vision Computing **21** (7): 623–635.

Anschriften der Autoren:

Dipl.-Ing. Wei Yao, Prof. Dr.-Ing. Uwe Stilla, Technische Universität München, Institut für Photogrammetrie und Kartographie, D-80333 München, Tel.: +49-89-289-22679, -22671, Fax: +49-89-280-9573, e-mail: yao@tum.de, stilla@tum.de.

Prof. Dr.-Ing. Stefan Hinz, Universität Karlsruhe, Institut für Photogrammetrie und Fernerkundung, D-76131 Karlsruhe, Tel.: +49-721 608-2314, Fax: +49-721 608-8450, e-mail: stefan.hinz@ipf.uni-karlsruhe.de.

Manuskript eingereicht: Mai 2009
Angenommen: Juli 2009