INFERRING TRAFFIC ACTIVITY FROM OPTICAL SATELLITE IMAGES

J. Leitloff [a, *], S. Hinz [b], U. Stilla [a]

[a] Photogrammetry and Remote Sensing
[b] Remote Sensing Technology
Technische Universität München, Arcisstrasse 21, 80333 München, Germany
{Jens.Leitloff | Stefan.Hinz | Uwe.Stilla}@bv.tum.de

KEY WORDS: Urban, Detection, Matching, Quickbird, Multispectral, Multiresolution

ABSTRACT:

In this paper we describe an approach to automatically estimate movements of vehicles in optical satellite imagery. The approach takes advantage of the fact that the optical axes of the panchromatic and multispectral channels of current spaceborne systems like IKONOS or Quickbird are not coinciding. The time gap that appears between the acquisition of the panchromatic and multispectral data can be used to derive velocity information. We employ a sub-pixel matching approach relying on gradient directions followed by least-squares fitting of Gaussian kernels to estimate the movement. The incorporation of the least-squares framework provides the basis to conclude about the accuracy of the movement estimates and to apply a statistical test deciding whether an object moves at all. We illustrate the matching and estimation scheme by various examples of real data.

1. INTRODUCTION

The automatic detection, characterization and monitoring of traffic using airborne and spaceborne data has become an emerging field of research. Approaches for vehicle detection and monitoring include not only video cameras but nearly the whole range of available sensors such as optical aerial and satellite sensors, infrared cameras, SAR systems, and airborne LIDAR. The broad variety of approaches can be viewed, for instance, in the compilations (Stilla et al., 2005) and (Hinz et al., 2006). Although airborne cameras are already in use and seem to be an obvious choice, satellite systems have entered the resolution regime required for vehicle detection. Sub-metric resolution is available in the optical domain and, since the successful launch of TerraSAR-X, closely followed by Spotlight SAR data. While the utilization of along-track interferometric spaceborne SAR data as delivered by TerraSAR-X is straightforward for movement estimation, optical satellite images do not provide interferometric capabilities nor they allow for acquiring image sequences with reasonable frame rate (e.g., the time gap of IKONOS single-pass stereo pairs reaches 6-12s).

However, the CCD linescanners implemented in the IKONOS and Quickbird instrument offer the potential for deriving information about moving objects. Due to constructional reasons the panchromatic (pan) and multispectral (ms) channels have non-coinciding optical axes, which lead to a small time gap of 0.2s between the acquisition of the pan and ms images (see e.g. (EURIMAGE, 2007)).

Two typical traffic scenes of a pan-sharpened Quickbird scene of an urban area are shown in Figure 1. The delay of 0.2s can be clearly seen by the typical color fringe caused by fast moving objects, which are observed at different positions in the pan and ms data.

(a) (b)

Figure 1. Spaceborne traffic scenes (Quickbird). Moving cars can be identified by their color fringe; see blue car in (a) and red cars in (b)

The appearance of color fringes at moving objects is well-known and has been mostly treated as artefact. Usually techniques are developed for removing the fringes to enhance the visual quality of images, although the potential for movement estimation has been already recognized – especially for airborne line scanner cameras like HRSC or ADS-40. In the context of spaceborne images, for instance, Etaya et al. (2004) showed the potential for moving object detection, using the delay between the acquisition of the pan and the ms channels. While this work was also done on Quickbird imagery another successful application of their approach is demonstrated for Spot data in Etaya et al. (2005). Here ocean wave movements during the Northern Sumatra Earthquake in 2004 were detected.

While these works use manual object selection in both channels, we focus on the automatic matching of hypotheses found in the high resolution channel with their conjugates in the lower resolution channels. In particular, we apply the approach to the task of automatic movement detection of vehicles.

* Corresponding author.

2. MOVEMENT ESTIMATION

First, an overview of the complete system for vehicle extraction will be given. Afterwards the procedures for movement estimation are described in detail.

2.1 System Overview

The presented work is a part of an overall framework for the estimation of traffic parameters from satellite imagery. As input the panchromatic and multispectral channels are used. These images are geo-coded and co-registered with road data taken from a geoinformation system (GIS). These give information about potential regions of interest, i.e. roads, junctions, and parking lots. The further processing is divided into the extraction of vehicles that are grouped in queues or rows and the detection of isolated standing vehicles. The vehicle queue detection has already been described in detail, see (Leitloff et al., 2006), and approaches for the extraction of single cars from sub-meter resolution images are outlined in (Hinz, 2005, Hinz et al., 2007).

This work focuses on the movement estimation of single vehicles. To thoroughly analyze the robustness of the approach, we assume the input data to be 100% complete and correct. Therefore, vehicles were selected manually from the panchromatic channel and are used as input. This step can easily be replaced by the automatic techniques mentioned above. A simplified system overview is illustrated in Fig. 2.

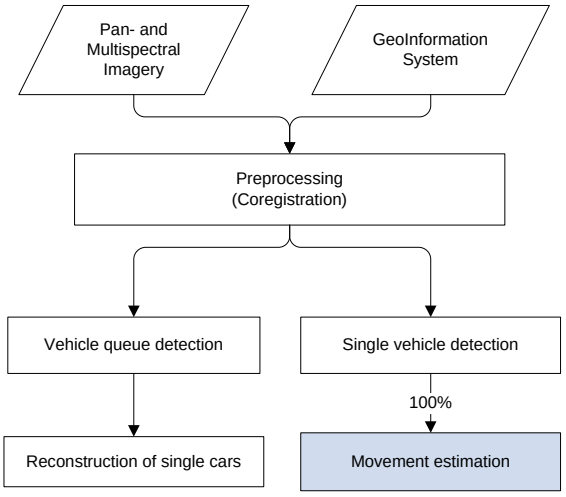


Figure 2. System overview

2.2 Initial hypotheses extraction

Starting from the manually selected vehicles' center in the pan image the position of the corresponding blob in the low resolution ms images needs to be determined. Instead of applying a search algorithm on all color channels independently, the RGB images are transformed into the Intensity-Hue-Saturation (IHS) color space. The distribution of hue and saturation values for different typical objects is illustrated in Figure 3. It depicts a horizontal slice through the IHS conic, with the intensity values projected on the slice, i.e. a polar coordinate system is used, where the length of the vector between each point and the center corresponds to the saturation and the hue is defined as angle between abscissa and the vector to each point. It comes clear that high color saturation of an object like a red or blue car on a road in at least one of the multispectral channels strongly influences the distance form the coordinate origin (Figs. 3a,b), while a gray road with less much saturation is located near the origin (Fig. 3c).

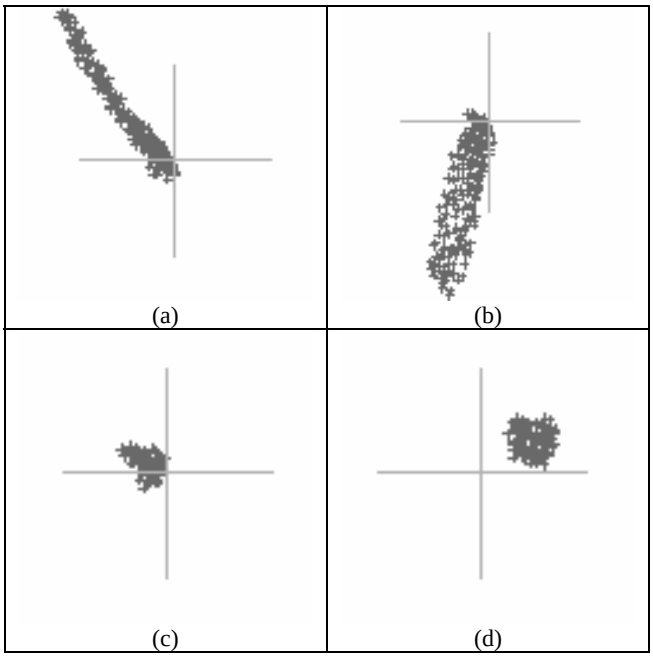
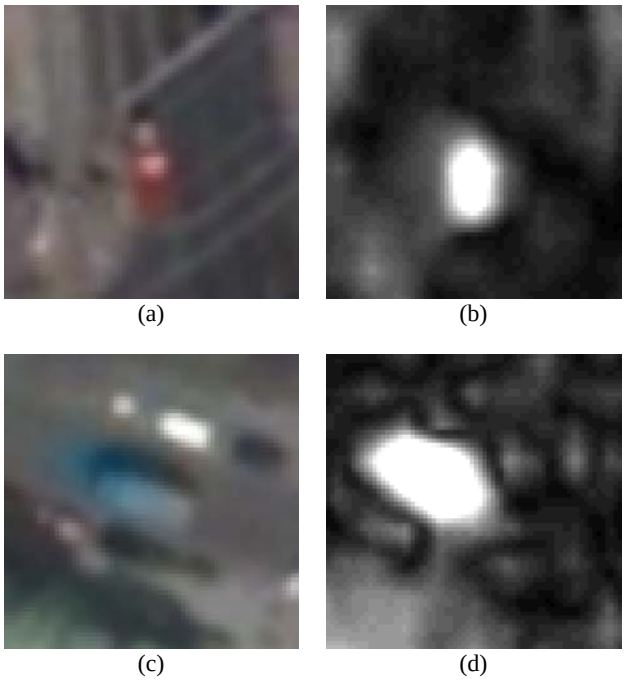


Figure 3. Hue (angle) and saturation (distance) for (a) red car and street, (b) blue car and street, (c) street, (d) grassland

Examples for the saturation channel are illustrated in Figure 4. The blob structure of vehicles in this channel becomes clearly visible and is fundamental for the following steps.

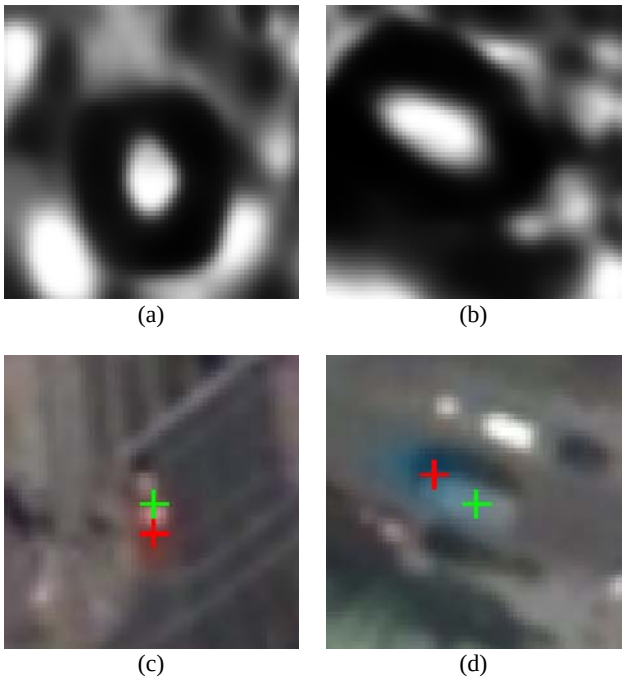
To find the translation vector between the blob of the input region and its conjugate region in the ms data, a sub-pixel matching approach is employed. The approach relies on the work of Steger (2001). The main idea of this algorithm is the use of a gradient filter to determine the gradient direction for each pixel of an image. First, a template is cropped from the pan channel around the vehicles' position. To adapt it to the scale of the coarser resolution of the ms channels the template is smoothed by a Gaussian kernel for eliminating possible substructures from the pan image. Then the gradient directions for this template and the IHS channels are calculated. An exhaustive search constrained to a reasonably limited area compares the gradient directions of the template image with the gradient directions found in the IHS images, which results in a similarity measure for each pixel (Figure 5a and 5b). Notice that the pure utilization of gradient directions makes normalization or contrast manipulations of one or both data sets virtually unnecessary. For eventually detecting the vehicle region in the ms data, the position showing the highest similarity is chosen, if a maximum distance to the position in the pan channel is not exceeded. Figure 5c and 5d show the results of the initial detection.



(a) (b)

(c) (d)

Figure 4. (a) and (c) Pansharpened images, (b) and (d) corresponding saturation images



(a) (b)

(c) (d)

Figure 5. Matching results: (a) and (b) show similarity measure, (c) and (d) corresponding detection with given position in pan image (green) and detection in ms images (red)

2.3 Accurate position estimation

The main goal of this step is to refine the matching result and – even more important – deliver statistically justified evidence about potential movement within the acquisition of the pan and the ms channels. Up to now, the precision of the positions in the higher and lower resolution is unknown. Therefore, the initial detections will be refined in terms of positional accuracy. Since this refinement is accomplished in a robust least square fitting approach, accuracies of the corrected position will be derived.

Assuming that active traffic moves along the road, a profile centered at the initial detection is spanned along the road direction (taken from the corresponding GIS axis). Values in the pan and similarity image are determined by bilinear interpolation for each profile point. Then, the parameters of the following Gaussian function are determined using Least Squares fitting:

$$y = s + \frac{a}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

with y ... ordinate value
x ... abscisse value
s ... shift in ordinate
a ... amplitude
σ ... width
μ ... shift in abscisse

The accuracies of the unknown parameters (s, a, σ, μ) are obtained from their covariance matrix after an iterative fitting process. Figure 6 shows two examples of the fitted Gaussian function (continuous line) compared to the extracted profile points (dashed line). The finally refined positions in the image and the underlying profiles can be seen in Fig. 7.

2.4 Movement estimation

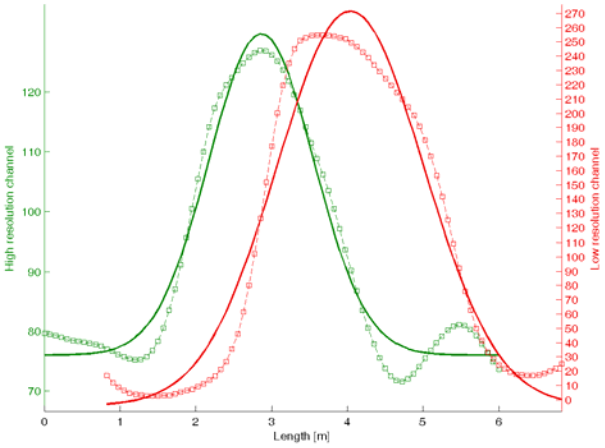
To make a decision about a significant movement of a vehicle, the distance between the position in the pan and the ms image is calculated. Using the known positional errors the accuracy of this distance can also be determined by simple error propagation:

$$\sigma^2_{dist} = \sigma^2_{\mu_{ms}} + \sigma^2_{\mu_{pan}}$$

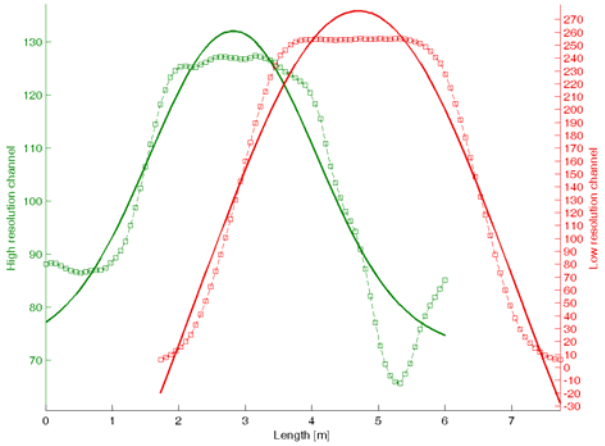
with σ^2_{dist} ... variance of distance between pan and ms
$\sigma^2_{\mu_{pan}}$ ... variance of position in pan channel
$\sigma^2_{\mu_{ms}}$ ... variance of position in ms channels

Finally, a statistical test (i.e. Student's test) is conducted with a standard value of 5% for the probability of error.

In our tests we obtained values of less then half a pixel for the tested distances to verify a movement, i.e., 1.2m when taking the coarse resolution of the ms channels as reference scale. This corresponds to a minimum velocity of 20km/h, which is well below the typical velocity in city areas.

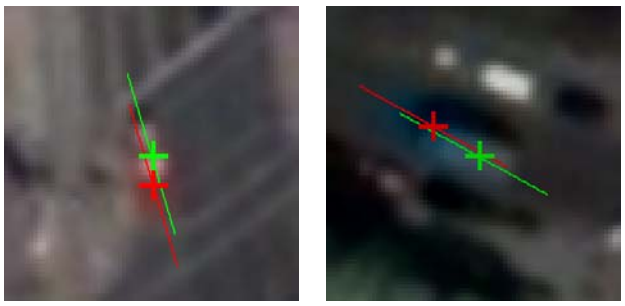


(a)



(b)

Figure 6. Refined position estimation for pan (green) and similarity (red) image. (a) and (b) correspond to the examples of initial matching in Fig. 5 (c) and (d), respectively.



(a) (b)

Figure 7. Profiles and accurate position for panchromatic (green) and similarity image (red). (a) and (b) correspond to the examples shown in Fig. 6 (a) and (b), respectively.

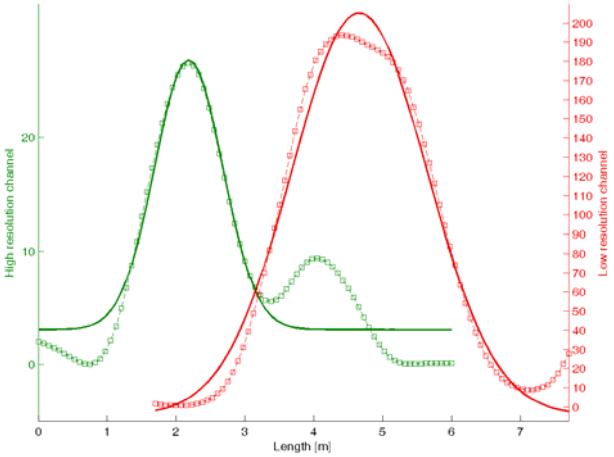
3. RESULTS

In this section more examples for the application of our approach are shown to illustrate the general performance of this approach. A numerical evaluation will be given in a later paper, once this algorithm has been added to the overall system sketched in Fig. 2.

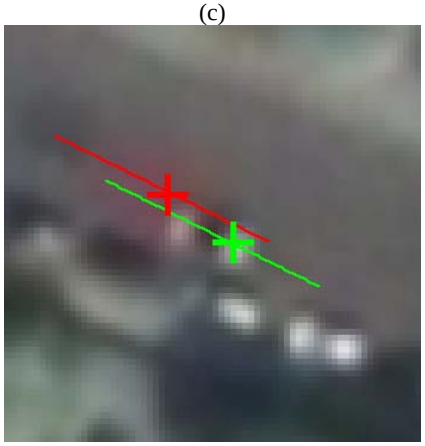
The figures of this section are organized as follows:

(a) original image with initial detection
(b) similarity image from matching
(c) results of the fitting process
(d) final results and used profiles

(a) (b)

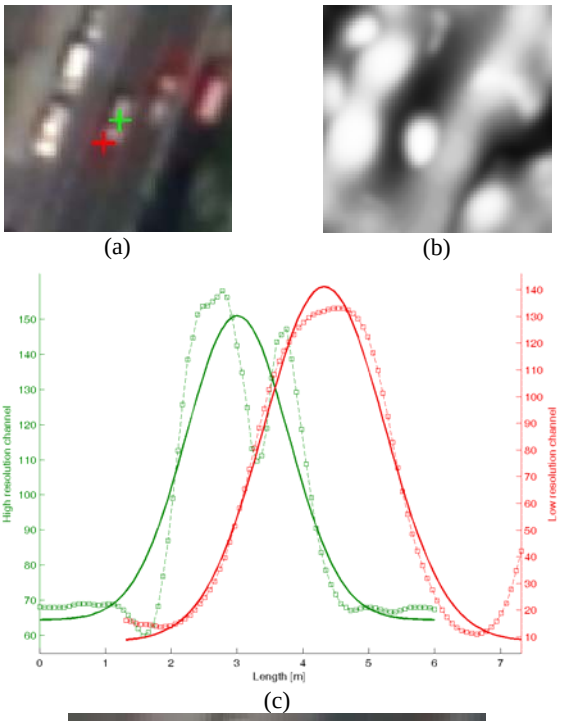


(c)



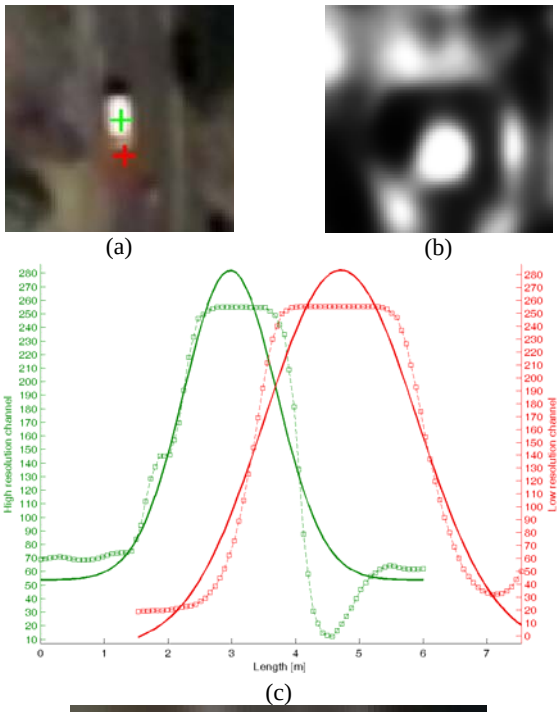
(d)

Figure 8. Example I



(a) (b) (c) (d)

Figure 9. Example II



(a) (b) (c) (d)

Figure 10. Example III

4. OUTLOOK

The examples show the potential of the presented approach. The position of all vehicles could be detected in the low resolution ms channels by the use of robust least square fitting in combination with a shape-based matching algorithm. This module will now be added to the overall detection for numerical evaluation. Furthermore, tests on gray vehicles which do not have large contrast in the ms channels will be performed. We expect that even in these cases movement detection can be performed, although it might be slightly less reliable.

REFERENCES

Etaya, M., Sakata, T., Shimoda, H., Matsumae, Y., 2004. An experiment on detecting moving objects using a single scene of QuickBird data. *Journal of The Remote Sensing of Japan* 24 (4): 357–366.

Etaya, M., Nakano, R., Shimoda, H., Sakata, T., 2005. Detection of Ocean Wave Movements after the Northern Sumatra Earthquake using SPOT Images, *Proceedings of IGARSS (2005)*, pp. 1420-1423.

EURIMAGE, 2007. www.eurimage.com, accessed Aug. 1st 2007.

Hinz, S., 2005. Fast and Subpixel Precise Blob Detection and Attribution. In *Proceedings of ICIP 05, on CD*.

Hinz, S., Bamler, R., Stilla, U., 2006 (Eds.): Theme issue "Airborne and spaceborne traffic monitoring". – ISPRS Journal of Photogrammetry and Remote Sensing **61** (3/4)

Hinz, S., Lenhart, D., Leitloff, J., 2007: Detection and Tracking of Vehicles in Low Framerate Image Sequences. Workshop on High Resolution Earth Imaging for Geospatial Information, Hanover, Germany

Leitloff, J., Hinz, S., Stilla, U., 2006. Detection of Vehicle Queues in Quickbird Images of City Areas. Photogrammetrie – Fernerkundung – Geoinformation, **4** (6): 315 - 325.

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

Stilla, U., Rottensteiner F., Hinz. S. (Eds.), 2005. Object Extraction for 3D City Models, Road Databases, and Traffic Monitoring - Concepts, Algorithms, and Evaluation (CMRT05), Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3/W24)