

MOVING TARGETS VELOCITY AND DIRECTION ESTIMATION BY USING A SINGLE OPTICAL VHR SATELLITE IMAGERY

M. Pesaresi^{a,*}, K. Gutjahr^b, E. Pagot^a

^a European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen (IPSC), via E.Fermi 1 TP 267 Ispra 21020 (VA), Italy- Martino.Pesaresi@jrc.it

^b Joanneum Research ForschungsgesmbH, Institute of Digital Image Processing, Wastiangasse 6, A - 8010 Graz, Austria

Commission VI, WG VI/4

KEY WORDS: Fast moving target detection, satellite imagery, optical very high resolution (VHR) sensors

ABSTRACT:

The present contribution demonstrates the feasibility and explores the limits of a new method for estimating the velocity and direction of moving targets using a single VHR satellite dataset. The method is based on the fact that there is a time lag between the data collection of the panchromatic (PAN) and multi-spectral (MS) sensors in the same VHR platform. Consequently, it is developed around three main steps: i) accurate image-to-image registration between MS and PAN images with a sub-pixel displacement error, ii) precise location of barycentre of targets by mathematical morphology-based image transforms, and iii) estimation of the targets ground velocity and direction using the MS-PAN spatial displacement, the known time lag, and an image-to-ground transformation taking into account the interior and exterior orientation of the sensors and a terrain height reference. An evaluation of the reliability and limits of the proposed method based on the observation of the results regarding manually-selected moving and non-moving targets is included.

1. INTRODUCTION

Most of the applications using VHR data assume that the collection of one imagery data set occurs at one point in the time line. If a multi-temporal analysis is needed, this requires a multiple collection of imagery representing the same spatial domain at different points of the time line. Then change detection techniques are usually applied in order to enhance evolving phenomena or detect moving targets. The associated intrinsic limit is to be able to observe only changes having a frequency of occurrence/change or a velocity on the ground much slower than the revisit time of the satellite platforms. These are of the order of some days in optimal conditions of cloud coverage and other tasks constraints.

The detection of moving targets with remotely-sensed data is

usually done using conventional radar or imaging radar techniques, such as SAR (Meyer et al., 2006; Pettersson, 2004) and along-track SAR interferometry approaches have been proposed for traffic monitoring using SRTM/X-SAR ATI data (Suchandt et al., 2006). In alternative, optical video cameras are used for traffic monitoring (Munno et al., 1993), but they are mostly based on ground or aircraft platforms that offer technical characteristics as image spatial resolution and frequency of image collection that unfortunately are not available in the satellite platforms at the date (Toth and Grejner-Brzezinska, 2006; Reinartz et al. 2006). For security applications aircraft platforms have also limitations related to the accessibility of remote areas especially relevant in case of conflict-prone places or areas with severe security-related concerns that are not accessible by civilian missions.

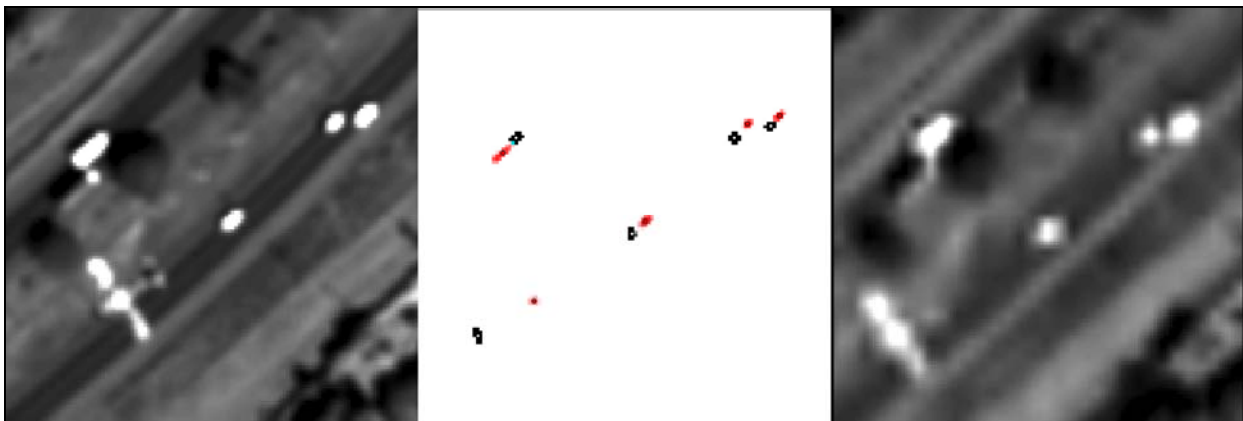


Figure 1 – sub-sample of the PAN (left) and MS (right) data used for detection of the image blobs and precise barycentres (center). In red and black the blobs coming from the PAN and MS sensors, respectively.

* Corresponding author.

The proposed methodology changes radically the perspective by exploiting the fact that in reality satellite VHR optical systems have two different sensors on board, respectively multi-spectral (MS) and panchromatic (PAN), which do not capture the information instantaneously but with a given time lag between the two data collections. By knowing the exact time lag between the two data collections, it is possible to estimate the target's ground velocity with the approximation of the pixel size. The shift of fast-moving features in the PAN and MS images produced by the time lag between the two image data collections are well known characteristics of VHR data (Dial 2003), but no systematic evaluation of the reliability of this information for estimation of the target velocity is available in the literature at the date. Usually this characteristic of VHR satellite data is considered as an artefact to be fixed for improving the image-to-image matching and digital surface extraction from stereo-pairs (Baltsavias and others 2001, Jacobsen 2005).

The detection of moving targets using remotely-sensed data has a clear connection with the support to all the activities related to traffic monitoring, estimation of traffic volume, and the modelling and planning of the traffic both in the terrestrial and maritime environment. An important application field can also be related to security and defence, where it is important to know if, and possibly where, there are moving vehicles or in general targets in the scene under analysis, with the estimation of ground velocity and direction. Moving vehicles detection using remotely sensed data could be also important in case of necessity to estimate the illegal border crossing of goods and people on remote areas difficult to control with traditional approaches.

2. THE PROPOSED METHOD

2.1 Processing flow

The concept presented here is tested using a part of VHR data coming from the Quick Bird® satellite of Digital Globe, having nominally 0.6 and 2.4 meter spatial resolution of the panchromatic and multi-spectral sensor, respectively. The processing flow consists on six steps as follow:

1. Fusion of the MS channels in order to simulate the spectral response of the panchromatic P
2. MS and P resolution matching
3. Image-to-image registration by correlation matching of the MS output of point (2) on the panchromatic channel using the latter as reference
4. Rough selection of potentially interesting target points (fuzzy markers)
5. Automatic refinement of the targets location
6. Estimation of the targets ground velocity

Steps 1 and 2 are functional to the preparation of the image-to-image registration, with the objective to use the multi-spectral data (MS) of the sensor for creating an MS-derived image as similar as possible to the spectral characteristics of the panchromatic (P) image. For Quickbird images the fusion is done by spectral averaging all the values of the MS sensor, simulating the frequency response of the P sensor. To ensure same gray value distributions a histogram matching is applied to the fused MS-image with the panchromatic image serving as reference. In order to reach the resolution of the P sensor, the MS-derived image was then over-sampled by a factor 4 through

bi-linear resampling approach. The original P image was filtered by a 4 x 4 low-band-pass convolution filter in order to create an image with similar spatial characteristics to the MS-derived image, but without losing the accuracy on the location of the barycentre of the targets.

Steps 3, 4, 5 are described in the following paragraph 4, and the final step 6 is detailed in paragraph 5.

2.2 Accuracy considerations

The time lag $\Delta t_{Pan,MS}$ between the MS and P sensor acquisition is the limiting factor for the velocity determination. Obviously a coarse estimation can be found according to

$$v = \frac{\Delta d_{Pan,MS}}{\Delta t_{Pan,MS}} \quad \text{where } \Delta d_{Pan,MS} \text{ is the length of the}$$

displacement vector between the target locations in the MS and P image. In case of Quickbird, the time lag is about 0.2 sec and the nominal spatial resolution of the panchromatic and multi-spectral sensor is 0.6 and 2.4 meters, respectively. Thus a target velocity of 50 km/h corresponds to 4.7 pixels in the P image or, the other way around, a displacement of 1 pixel corresponds to about 10 km/h.

As a consequence reliable velocity results can only be achieved if sub-pixel accuracy is ensured throughout the processing flow, most particularly for the image-to-image registration and the target localization. Since the velocity can be estimated from relative displacements, the absolute location accuracy and the quality of the reference height information are of minor importance.

3. MOVING TARGET DETECTION

As the accuracy of the estimated velocity depends on the quality of the measured displacement vectors, an image-to-image registration is required. Therefore, a dense grid of points in the modified P image is identified with the fused MS image, applying image matching techniques (Paar and Pölzleitner, 1992). These points are used to calculate an affine transformation between both images. Poor identification results are filtered by the back-matching distance and additionally can be removed during the polynomial calculation by looking at the individual point residuals.

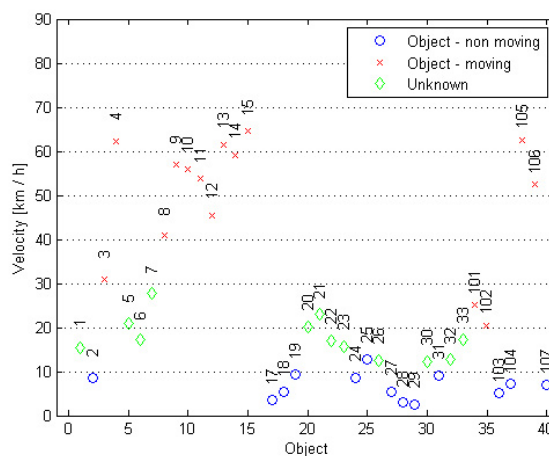


Figure 2 – estimated velocity for the selected targets (objects)

We chose a grid of 10 by 10 pixels in the area of interest (1000 by 1000 pan-pixels) yielding about 9000 tie-points (TP's) candidates. These points were first filtered using a threshold of 0.5 pixels for the back-matching distance. When fitting an affine polynomial to the remaining 1901 points again 198 tie-points were eliminated because of their poor residual statistic. The residual statistics of the polynomial fitting is summarized in Table 1. Finally, the fused MS image is warped to the P image using the found affine transformation.

1703 TP's	Column [pxl]	Line [pxl]	Length [pxl]
<i>Std.Dev.</i>	0.76	0.81	1.11
<i>Mean</i>	0.00	0.00	0.99
<i>Minimum</i>	-1.89	-2.02	0.04
<i>Maximum</i>	1.87	2.02	2.59

Table 1: Residual statistics of the polynomial transformation calculation

The rough selection of potentially interesting target points on step 4 of the processing flow has been solved by visual inspection of the images after the image-to-image automatic matching. The procedure was the following: a disk of radius r overlapping interesting targets (moving and stable objects) was manually digitised and recorded in a separate image layer. Stable targets were acquired in order to have an indirect measure of the image-to-image registration intrinsic error. The disk serves as fuzzy marker with a spatial fuzziness proportional to parameter r . In this case $r=15$ pixels (corresponding to a radius of 9 meters on the ground) was chosen in order to include any potential interesting target such as cars and/or trucks.

The precise location of the targets (step 5) was calculated using a procedure based on mathematical morphology and involving the detection of the image structures (blobs) related to the manually-fuzzy-selected targets together with the calculation of the precise barycentre of these image blobs.

In order to detect both targets brighter and darker than the

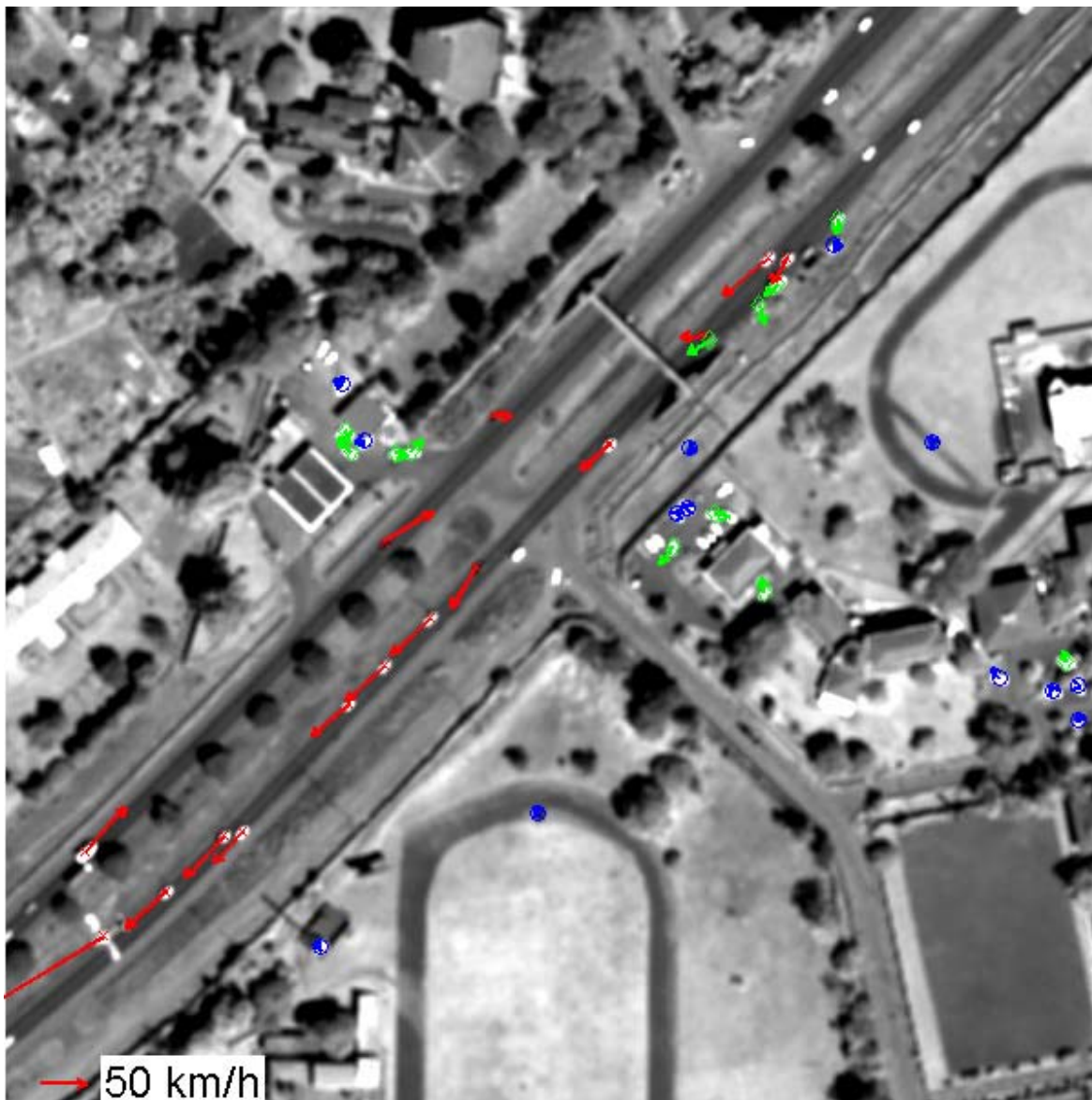


Figure 3 – estimated velocity of targets overlaid on the PAN image. Red, green and blue targets were classified as “moving”, “unknown”, and “non-moving”, respectively, by visual inspection.

surrounding image background, the procedure used a parallel approach based on the calculation of the residuals of morphological opening and morphological closing by reconstruction, respectively.

In particular, the image blobs of brighter targets of the image $T_{bright}(f)$ are calculated by the white top-hat by reconstruction (RWTH) morphological transform as follows:

$$\begin{aligned} T_{bright}(f) &= RWTH(f) > 0 \\ f - \gamma_R^{(n)}(f) &> 0 \end{aligned} \quad (1)$$

While the image blobs corresponding to the darker targets of the image $T_{dark}(f)$ are calculated by the black top-hat by reconstruction (RBTH) morphological transform as follows:

$$\begin{aligned} T_{dark}(f) &= RBTH(f) > 0 \\ \phi_R^{(n)}(f) - f &> 0 \end{aligned} \quad (2)$$

Where $\gamma_R^{(n)}(f)$ and $\phi_R^{(n)}(f)$ are the opening and closing by reconstruction, respectively, with a structuring element of size (n) of the input image (f) (Soille 2003, p.212). The target blobs were derived from the support area of the $RWTH(f)$ and $RBTH(f)$, defined as the region of pixels in the transformed image $T_{bright}(f)$ or $T_{dark}(f)$, having residuals greater than zero. The size (n) of the structuring element used here is of course linked to the size of the targets we want to detect as bright or dark structure in the image: in this case the minimal 3x3 box was sufficient.

The barycentre of the detected blobs is then calculated by simply averaging the x and y image coordinates of all pixels belonging to each specific blob. Other alternatives have been tested as the detection of the extrema (regional maxima and regional minima for bright and dark blobs, respectively) of the grey-level functions associated with the blobs, or the weighted average of the coordinates by the intensity of the blob contrast. However, by analysing the estimated displacement of stable targets, the results of these more complicated procedures were not reducing this displacement but increasing it by a factor of about 5 to 10 % with respect to what we obtained with the simple coordinate averaging method. Therefore, the last method was chosen in the following.

4. ESTIMATION OF THE TARGETS' GROUND VELOCITY

4.1 Ground-to-image and inverse transformation

To determine the velocity (step 6 of the processing flow), the barycentre pixel coordinates of the detected blobs in the pre-processed MS and PAN images had to be geo-referenced. This was based on a so-called image-to-ground transformation which requires the additional information on the interior and exterior orientation of the sensors at the acquisition time of the image and a height reference. In case of Quickbird or Ikonos sensors, the rational polynomial coefficients (RPC's) are supplied with the image data and directly describe the ground-to-image transformation by two fractions of cubic polynomials:

$$x = \frac{f_x(X, Y, Z)}{g_x(X, Y, Z)} \quad \text{and} \quad y = \frac{f_y(X, Y, Z)}{g_y(X, Y, Z)} \quad (3)$$

The original sensor polynomials depend on geographic coordinates (λ, φ, h) but we prefer the representation in an Earth-centred, Earth-fixed Cartesian system (X, Y, Z).

4.1.1 Image-to-map transformation

Calculation of East and North ground coordinate for given height. Approximations for East and North:

$$\begin{aligned} \bar{r} &= (r - r_0) / \mu_r \\ \bar{r} &= \frac{L_{1,1} + L_{1,2} \cdot \bar{\lambda} + L_{1,3} \cdot \bar{\phi} + L_{1,4} \cdot \bar{H}}{L_{2,1} + L_{2,2} \cdot \bar{\lambda} + L_{2,3} \cdot \bar{\phi} + L_{2,4} \cdot \bar{H}} \end{aligned} \quad (4)$$

$$\begin{aligned} \bar{c} &= (c - c_0) / \mu_c \\ \bar{c} &= \frac{C_{1,1} + C_{1,2} \cdot \bar{\lambda} + C_{1,3} \cdot \bar{\phi} + C_{1,4} \cdot \bar{H}}{C_{2,1} + C_{2,2} \cdot \bar{\lambda} + C_{2,3} \cdot \bar{\phi} + C_{2,4} \cdot \bar{H}} \end{aligned} \quad (5)$$

Solve for $\bar{\phi}$ and $\bar{\lambda}$:

$$\begin{aligned} (L_{1,2} - \bar{r} \cdot L_{2,2}) \cdot \bar{\lambda} + (L_{1,3} - \bar{r} \cdot L_{2,3}) \cdot \bar{\phi} &= \\ (L_{2,1} + L_{2,4} \cdot \bar{H}) \cdot \bar{r} - (L_{1,1} + L_{1,4} \cdot \bar{H}) \end{aligned} \quad (6)$$

$$\begin{aligned} (C_{1,2} - \bar{c} \cdot C_{2,2}) \cdot \bar{\lambda} + (C_{1,3} - \bar{c} \cdot C_{2,3}) \cdot \bar{\phi} &= \\ (C_{2,1} + C_{2,4} \cdot \bar{H}) \cdot \bar{c} - (C_{1,1} + C_{1,4} \cdot \bar{H}) \end{aligned} \quad (7)$$

4.1.2 Inverse normalization to get ϕ and λ .

Iterate for East (X) and North (Y):

$$\begin{aligned} r &= r(\bar{X} + d\bar{X}) \\ r &= r(\bar{X}_0) + \frac{\partial r}{\partial \lambda} \cdot d\lambda + \frac{\partial r}{\partial \phi} \cdot d\phi + \frac{\partial r}{\partial H} \cdot dH \end{aligned} \quad (8)$$

$$\begin{aligned} c &= c(\bar{X} + d\bar{X}) \\ c &= c(\bar{X}_0) + \frac{\partial c}{\partial \lambda} \cdot d\lambda + \frac{\partial c}{\partial \phi} \cdot d\phi + \frac{\partial c}{\partial H} \cdot dH \end{aligned} \quad (9)$$

$$\frac{\partial r}{\partial \lambda} \cdot d\lambda + \frac{\partial r}{\partial \phi} \cdot d\phi + \frac{\partial r}{\partial H} \cdot dH = r - r(\bar{X}_0) \quad (10)$$

$$\frac{\partial c}{\partial \lambda} \cdot d\lambda + \frac{\partial c}{\partial \phi} \cdot d\phi + \frac{\partial c}{\partial H} \cdot dH = c - c(\bar{X}_0) \quad (11)$$

4.1.3 Calculus of the solutions

The solution of the inverse transformation can be found iteratively using Taylor series expansions. It is a well known and documented problem that the meta-information has a limited and variable absolute accuracy (Jacobsen 2005, Dial and Grodecki, 2002). Generally a sensor model optimization based on ground control points (GCP's) is required. This is crucial for security related applications as the access to the area of interests is restricted in many cases. Anyhow, the accuracy of the image-to-ground transformation also depends on the used

reference height. We used the globally available SRTM C-band DEM with a nominal resolution of about 90 m and a relative height accuracy of 10 m. Obviously this level of detail and height accuracy does not meet the requirements specified above.

In our approach we try to solve, or at least release, all these problems in the image-to-image registration step. First this registration reduces the sensor geometries to be considered and optimized to one (pan-chromatic sensor). Secondly we again have to stress that we are only interested in relative displacements. Thus the absolute pointing accuracy is of minor interest if the topography around the moving objects is changing smoothly. This is a feasible condition for vehicles as the road networks in general do not show very steep height gradients.

4.1.4 Results

Figure 2 shows the velocity estimations of 40 targets that were classified by visual inspection. Three classes were assigned: "non moving objects" such as for parked vehicles or trees, "moving objects", and "unknown" when it was hard for the interpreter to decide whether or not the target was moving. It is evident that the "non-moving" targets are almost all clustered below the velocity line of 10km/h, which corresponds to roughly one panchromatic pixel of displacement with the given MS-PAN collection time lag (see point 3.2). This can be considered as the technological limit of the tested source given by the constraints on available resolution, MS and PAN image matching precision, and MS-PAN collection time lag. The targets with estimated velocity between 10 km/h and 30 km/h were almost "unknown" from the visual inspection, because the displacement was too small to be detected univocally by manual means, while the targets with estimated velocity greater than 30 Km/h were all labelled as "moving targets" by the visual inspection.

Moreover, if we check the direction and angle together with the velocity of the moving targets on the road, we can observe that the results are coherent with the expected opposite directions in the two road lanes (Figure 3).

5. CONCLUSIONS

The present contribution demonstrates the feasibility and explores the limits of a new method for estimating the velocity and direction of moving targets using a single VHR satellite dataset. The method is based on the fact that there is a time lag between the data collection of the PAN and MS sensor in the same VHR platform, and it is developed around three main steps: i) image-to-image registration between MS and PAN images, ii) precise location of barycentre of targets, and iii) estimation of the targets ground velocity and direction by image-to-ground transformation.

This preliminary study has shown that the theoretical limit of the estimated velocity accuracy given by the spatial resolution and MS-PAN time lag constraints can be reached using sub-pixel image-to-image warping and accurate target barycentre detection procedures. In particular, with the tested Quickbird sensor, we have estimated a limit of 10Km/h as minimal velocity of targets having the size of a car.

Further efforts will concentrate on testing the different limits associated to different VHR sensors available, and on the development of an automatic procedure able to select the potential moving targets without the manual intervention used in this methodology. The fully automatic target-selection

procedure will be required in case of application of this methodology for detection of unexpected (out of road) or low-velocity moving target in the entire satellite scene, that are difficult to identify by visual inspection.

ACKNOWLEDGEMENTS

The presented work has been supported by the institutional research activity of Action ISFEREA (Information Support for Effective and Rapid External Action) of the European Commission, Joint Research Centre, Institute for Protection of the Citizen (IPSC). The research concept and his implementation have been made possible because of the effective sharing of ideas and tools inside the GMOSS (Global Monitoring for Security and Stability) Network of Excellence.

REFERENCES

- Baltsavias E., Pateraki M., Zhang L., 2001. Radiometric and geometric evaluation of Ikonos geo images and their use for 3D building modeling, Joint ISPRS Workshop "High Resolution Mapping from Space 2001", Hannover, Germany, 19-21 September.
- Dial G., 2003. Measuring Boat Speed, Space Imaging technical paper.
- Dial G., Grodecki J., 2002. Block adjustment with rational polynomial camera models. Proceeding of ASCM-ASPRS Annual Conventions, Washington D.C., April 2002, pp.
- Jacobsen K., 2005. High Resolution Satellite Imaging Systems – Overview, ISPRS Hannover Workshop 2005, High-Resolution Earth Imaging for Geospatial Information, May 17-20, 2005, Hannover, Germany.
- Meyer F., Hinz S., Laika A., Wehling D., and Bamler R., 2006. Performance analysis of the TerraSAR-X Traffic monitoring concept, ISPRS Journal of Photogrammetry and Remote Sensing, 61 (3-4), pp. 225-242
- Munno C.J., Turk H., Wayman J.L., Libert J.M., Tsao T.J., 1993. Automatic video image moving target detection for wide area surveillance, Security Technology Proceedings, Institute of Electrical and Electronics Engineers 1993 International Carnahan Conference on 13-15 Oct. 1993 pp. 47 – 57.
- Paar G., Pözlleitner W., 1992. Robust disparity estimation in terrain modelling for spacecraft navigation. Proc. 11th ICPR, International Association for Pattern Recognition, 1992, pp.
- Reinartz P., Lachaise M., Schmeer E., Krauss T. and Runge H., 2006, Traffic monitoring with serial images from airborne cameras, ISPRS Journal of Photogrammetry and Remote Sensing, 61 (3-4), pp. 149-158.
- Soille P., 2003. Morphological Image Analysis, Springer.
- Suchandt S., Einedera M., Breita H., Runge H., 2006. Analysis of ground moving objects using SRTM/X-SAR data, ISPRS Journal of Photogrammetry and Remote Sensing, 61 (3-4), pp. 209-224.
- Toth C.K. and Grejner-Brzezinska D., 2006, Extracting dynamic spatial data from airborne imaging sensors to support traffic flow estimation, ISPRS Journal of Photogrammetry and Remote Sensing, 61 (3-4), pp. 137-148.