# DETECTION OF MOVING VEHICLES WITH WORLDVIEW-2 SATELLITE DATA

Andrea MARCHESI[a], Marco RUSMINI[a], Gabriele CANDIANI[b], Giorgio DALLA VIA[a], Federico FRASSY[a], Pieralberto MAIANTI[a], Francesco ROTA NODARI[a], Marco GIANINETTO[a]

[a] Laboratory of Remote Sensing (L@RS), Politecnicodi Milano - Building Environment Sciences and Technology (BEST) Department.Via Ponzio 31, 20133 Milano, ITALY;Tel:+39-0223996204
[b] Optical Remote Sensing Group, CNR-IREA, Via Bassini 15, 20133 Milano, ITALY;Tel:+39-0223699550

E-mail: andrea.marchesi@polimi.it; marco.gianinetto@polimi.it

**KEY WORDS**:WorldView-2, Moving objects, Traffic monitoring, OBIA, Feature extraction.

**Abstract:**Traffic monitoring in urban areas is a complex issue and recent Remote Sensing technologies can play an important role in planning and monitoring the urban environment.In this studya semi-automatic object-oriented workflow was designedto detect moving vehicles and their speed from single pass WorldView-2 multispectral data.The time lag in data recording between each spectral band causes a small image displacement of moving objects andthis discrepancy is used to detect moving vehicles, their speed and direction of travel.The method proposed was applied to a very complex study area in the historical core of city of Multan, in the Pakistanisouthern province of Punjab, where very small and extremely dense built-up old style houses are mixed together with narrow roads and bazaar streets.First results show interesting applications of this new technology, with achieved accuracies of about 67% evaluated comparing automatic detection vs. manual interpretation.

## 1. INTRODUCTION

The rapid growth of Asian economies in the last decades caused a dramatic increase of traffic congestion.In metropolitan areas, trafficcreates a number of negative effects including pollution, carbon dioxide emissions increase, and, on average, a decrease of pedestrians, drivers and passengers'safety, causing enormous casualties and economic losses. Like most of developing countries, Pakistan experienced a massive traffic growth.

There are major factors which influenced this trend. During the last five years, banks' activities focused on leasing/financing cars purchasing. The country overcame a population of 160 million with an annual increase of 1.96%. There is a consequent lack of public transportation, and the local route busses are in deplorable conditions. Furthermore cultural changes in the society and lifestyle forced people to abandon the joint family system for a life more individual, where vehicles are shared among less and less people (Ahmed, 2011). In this scenario, image-based traffic monitoring systems play a crucial role in gathering traffic information. Nowadays, vehicle tracking and traffic control monitoring systems include the use of closed circuit television, traffic counters, variable message signs and other means of monitoring traffic by authorities to manage traffic flows and providing advice concerning traffic congestion. This methods, of course, provides very high resolution in terms of space and time but the areal extent of each observation is extremely narrow.

With the increasing availability of high spatial resolution satellite imagery, Remote Sensing asserts itself as a promising tool for road traffic monitoring (Larsen*et al.*, 2009).Nevertheless, up to now only few studies have successfully extracted vehicle information from satellite data by exploiting the time lag in data acquisition between the panchromatic (PAN) and multispectral (MS) data recording (Liu*et al.*, 2011; Xiong and Zhang, 2008).Moreover, high-resolution satellites usually have the possibility to acquire along-track stereo imagery as well. In past this characteristic has been used for different applications related to terrain modeling (Gianinetto, 2008a; Gianinetto, 2008b; Gianinetto, 2009) but the delay between the two stereo images collection can be also used to detect movements on the ground (Kääb, 2011) as well. More recently, the unique configuration of the WorldView-2 (WV-2) focal plane allowed to exploit the very small inter-band time lag between each imaging spectrometer for movement detection, as demonstrated by Kääb (2011) and Tao and Yu (2011). However, the majority of previous studies mainly focused on traffic detection on highways or multi-lane main roads (Mishra and Zhang, 2012; Salehi*et al.*, 2012), being a simpler task than traffic flow monitoring in urban networks.

This paper describes a semi-automatic object-oriented workflow to detect moving vehicles, their speed and direction of travel from single-pass WorldView-2 multispectral dataalong narrow city roads. Obviously, the method could be also applied to fast-flowing roads.

## 2. STUDY AREA AND DATASET

The study area is the cityof Multan (Figure 1), located in Punjab province at almost the geographic center of Pakistan. The analysis has been focused on the historic core of the Multan Walled City, an extremely dense urban area characterized by old style small and low houses built up close to each other with narrow roads and bazaar streets.

WorldView-2archive imagery (ID 1030010005C48C00 and ID 1030010005AB6300)collected on 6 June 2010 (06:13:06 UTC) were selected for this study. The images covered an area of about 100 $km^2$and were taken with an average off-nadir angle between 1° and 7°.Data have been delivered as an ortho-rectified bundle (0.5 meters PAN and 2.0 meters MS) corrected with a fine digital elevation model, resulting in a product with an approximate 10.2 m CE 90 accuracy (5.0 m RMSE) meeting 1:12,000 NMAS requirements.

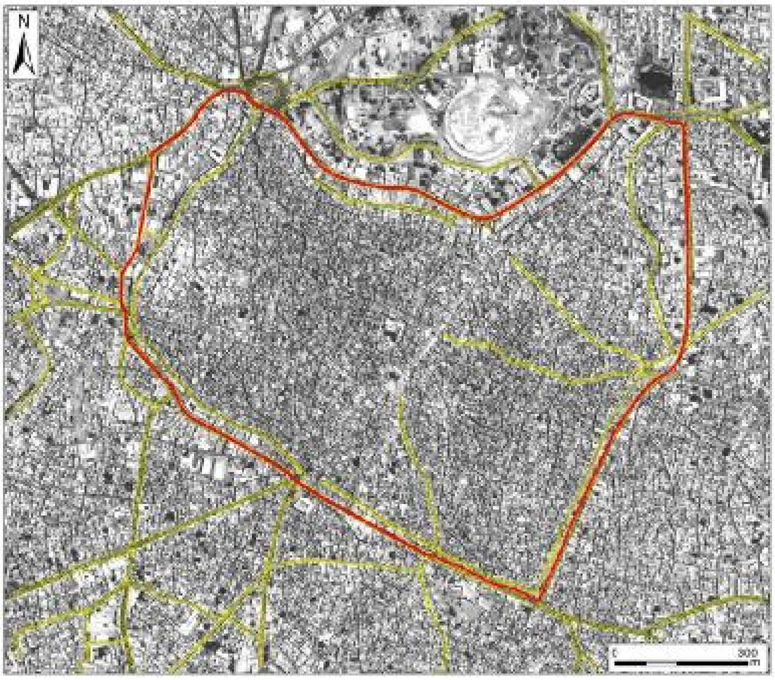


**Figure 1:**Study area.WorldView-2 panchromatic of Multan Walled City (red) withthe main road network vectorized from the satellite data (yellow)

## 3. METHODS

### Backgrounds

WorldView-2's focal plane consists of one PAN and two MS imaging sensors (MS1: Blue, Green, Red, Near-IR1 and MS2: Coastal Blue, Yellow, Red Edge, Near-IR2), recording asynchronously. Consequently, a very small time lag in data recording between each spectral band, few hundreds of milliseconds, causes imperceptible displacements of moving objects in the image space.The band recording order and the inter-band time lag is shown in Table 1 (Kääb, 2011).

In this study we focused on the use of Coastal Blue (time $t_0$) and Blue (time $t_1$)spectral bands for traffic detection due to the high spectral contrast between roads and vehicles in this specific wavelengths.The delay in data recording between Coastal Blue and Bluehas been estimated in 316ms (Kääb, 2011), thus, an object with a speed of 30 km/h (18.6 mph) will show a ground displacement of 2.6 m between the image sensed at time $t_0$(Coastal Blue) and time $t_1=t_0+316$ ms (Blue). In terms of image space, the ground displacement corresponds to about 5 image pixels in the PAN and less than 2 image pixel in the MS.Figure2 shows an example where moving vehicles can be easily visualized by using a false color composition. Here, the moving objects' position at time $t_0$is represented in red and in cyan at time $t_1$, while in white we have the common part of the objects in both the two spectral bands (same time).

**Table 1:** WorldView-2's recording properties

| Band recording data order | Sensor name | Wavelengths (nm) | Inter-band Time lag (s) | Time lag (s) from start | Sensors Time lag (s) from start |
|---|---|---|---|---|---|
| Near-IR2 | MS2 | 860-1040 | Recording start | Recording start | Recording start |
| Coastal Blue | MS2 | 400-450 | 0.008 | 0.008 | ↓ |
| Yellow | MS2 | 585-625 | 0.008 | 0.016 | |
| Red-Edge | MS2 | 705-745 | 0.008 | 0.024 | 0.024 |
| Panchromatic | PAN | 450-800 | | | |
| Blue | MS1 | 450-510 | 0.3 | 0.324 | 0.324 |
| Green | MS1 | 510-580 | 0.008 | 0.332 | ↓ |
| Red | MS1 | 630-690 | 0.008 | 0.34 | |
| Near-IR1 | MS1 | 770-895 | 0.008 | 0.348 | ~0.35 |

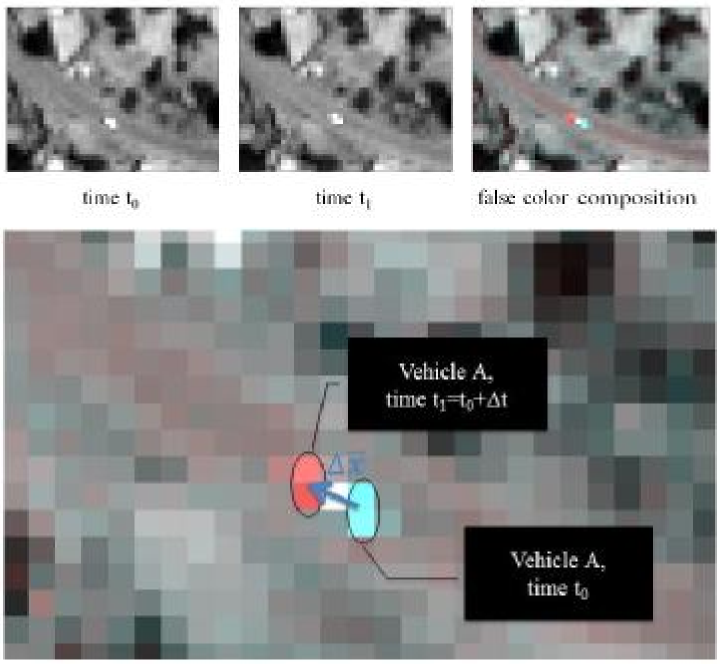


**Figure 2:** World-View-2false color composition (MS2-Coastal Blue is in red, and MS1-Blue both in green and blue). Here, the moving objects' position at time $t_0$ is represented in red and in cyan at time $t_1$, while in white we have the common part of the objects in both the two spectral bands.

The methodology outlined to derive traffic information can be splited into two main workflows (Figure 3):
a) Detection of moving objects by means of object-based feature extraction on both PAN and MS data;
b)Estimation of vehicles' speed and direction of travel by means of GIS analysis.

Finally, some products such as the city map of vehicles' density can be derived.

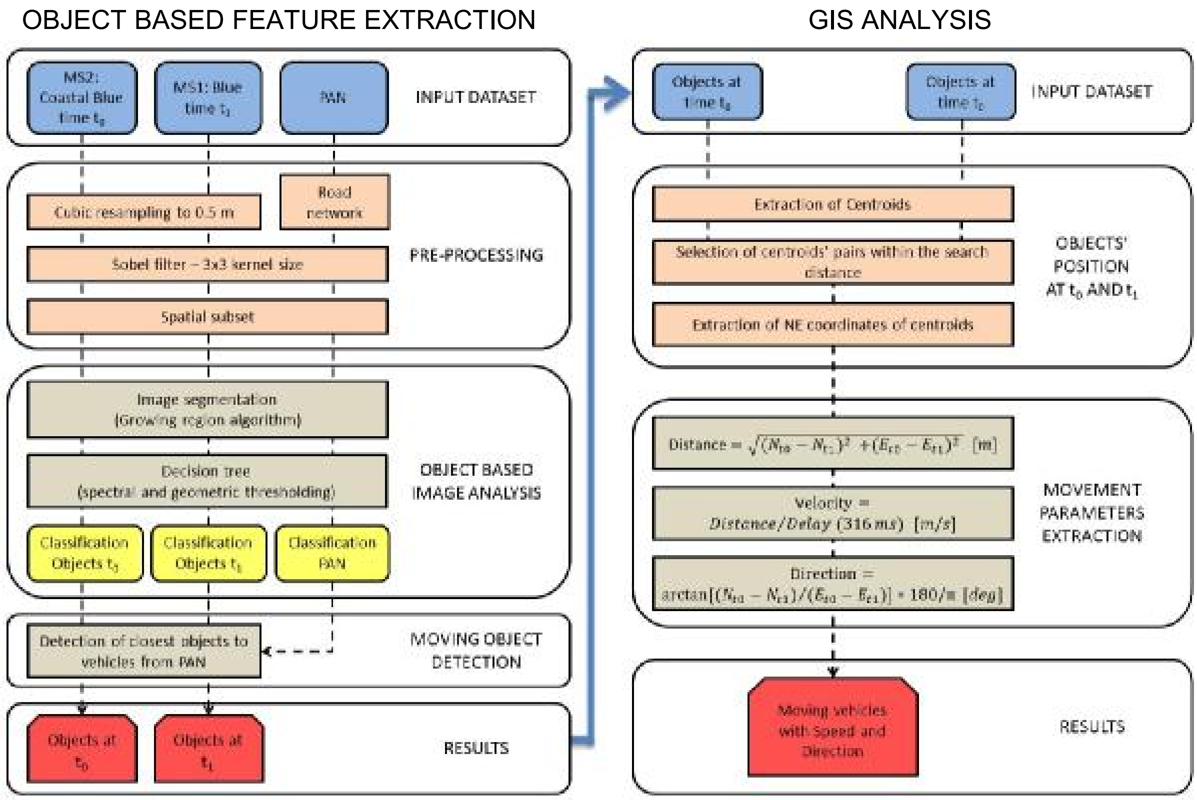


**Figure 3:** OverallFlowchart for vehicle detection and measurement

**Object-based feature extraction**

First of all, the main road network of the study area was manually extracted and vectorized from the WV-2 PAN imagery (Figure 1). This allowed to constrain the analysis to the road network only, so that biases resulting from misclassification of vehicle-like objects outside roadways could be avoided. Furthermore, the computing time for image segmentation and classification was sensibly reduced.
Since the ground sampling distance of WorldView-2's MS bands is 2.0m, shapes of small objects like cars, motorcycles, small vans or even carts are not clearly defined. Consequently, for enhancing the image contrast and making possible to fuse together both the PAN and MS data, Coastal Blue and Blue bands have been first super-sampled to the PAN geometric resolution (0.5 m) with a cubic convolution algorithm (Mishra and Zhang, 2012). When dealing with high-resolution single bands, some authors suggested to improve the classification results with the introduction of texture features (Su *et al.*, 2008; Johansen*et al.*, 2007; Murray*et al.*, 2010, Rusmini*et al.*, 2012), therefore we applied the non-linear edge detection Sobel filter with a 3x3 kernel size (Kittler, 1983) prior image classification. Results are shown in Figure 4.
Regarding the detection of moving objects, because the small size of vehicles never allowed the use of automatic image matching algorithms for homologous search (Gianinettoand Scaioni, 2008; Gianinetto, 2012),an ad-hoc spectral workflow was designed.Vehicle-like objects were distinctly detected on the three WV-2 bands through an object-based classification as follows:
1. A region growing algorithm (Haralick, 1985) was used for image segmentation. Table 2 lists all the input layers and the segmentation parameters used;
2. Vehicles appear as bright objects with respect to pavement and with a very compact shape.The classification of each spectral bandhas been carried out through the threshold-based rule sets listed in Table 3 (BaatzandSchäpe, 2000),where the parameters' values have been set using a trial anderror method. An example is shown in Figure 4.
3. As can be seen in Figure 4, vehicledetection from the panchromatic data is much more accurate than from the multispectral. Therefore, among the image objects extracted from the multispectral data only those with a small displacement from the corresponding extracted from the panchromatic band have been retained and labeled as "moving vehicles".

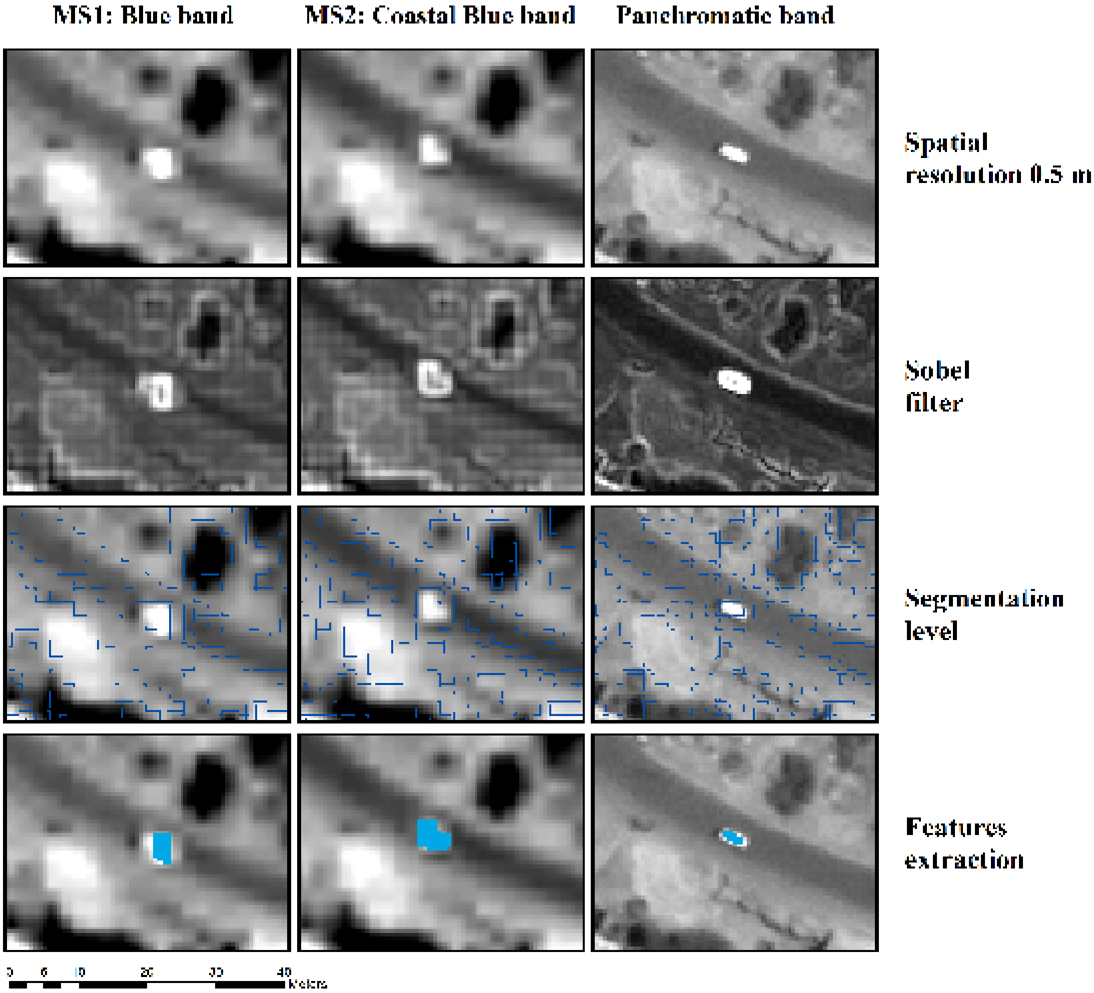


**Figure 4:** Example of object-based image analysis applied to the multispectral and panchromatic WorldView-2 data.

**Table2:** Input layers and segmentation parameters used for the image classification (PAN: panchromatic; MS: multispectral; CB: Coastal Blue; B: Blue).

| Input Layers | PAN segmentation | MS1 (Blue) segmentation | MS2 (Coastal Blue) segmentation |
|---|---|---|---|
| PAN | No | No | Yes |
| PANSobel filter | No | No | Yes |
| CB | No | Yes | No |
| CB Sobel filter | No | Yes | No |
| B | Yes | No | No |
| B Sobel filter | Yes | No | No |
| Road Network | Yes | Yes | Yes |
| Scale | 25 | 25 | 25 |
| Shape | 0.1 | 0.1 | 0.1 |
| Compactness | 0.9 | 0.9 | 0.9 |

**Table 3:**Threshold-based rule sets for the image classification (n.d. not defined).

| **Features** | **PAN** | **MS1** | **MS2** |
|---|---|---|---|
| Compactness | 0.9 – 1.65 | 0.9 – 1.50 | 0.9 – 1.50 |
| Max diff. | <= 1.40 | <=1.20 | <=1.40 |
| Band value | >=500 | n.d. | n.d. |
| Shape index | <1.51 | n.d. | n.d. |

Finally, the shape of moving objects at time $t_0$and time $t_1$ were stored in as vector files.The overall object-based feature extraction workflow is summarized in Figure 3.

**GIS Analysis**

The vector files of the moving objects resulting from the image classification were imported into a GIS environment (ArcGIS 10) as shapefiles. For the analysis of speed and direction of travel a specific ArcGIS tool has been implemented with ModelBuilder©.

The first step was the extraction of the objects' centroids at time $t_0$ and time $t_1$. From each centroid attime $t_0$, the closest centroid belonging to time $t_1$was detected and $t_0 – t_1$ object pairs were created. The search was limited to a radiusof 8.78m calculated on the basis of the bands delay (316 ms), so that objects pairs with a speed higher than 100km/h (62 mph) would not be selected. This constraint allowed to exclude errors in centroids pairs detection and seems a reasonable cut-off if we consider that the analyzed network is composed of narrow and well-travelled roads.

For each $t_0 – t_1$centroid's pair, their Easting (E) and Northing(N) coordinates were extracted and speed, distance and direction were calculated using the following equations:

The road network was then divided into segments representing single main roads, so that the traffic density could be evaluated for each section. Traffic density was calculated as the number of vehicles per unit area ($m^2$). The densities were classified as:

1.Slower (<12 moving vehicles/100$m^2$);
2.Medium (12-23 moving vehicles/100$m^2$);
3. Higher (>23 moving vehicles/100$m^2$).

Figure 6 shows the traffic density at the time of the image acquisition.

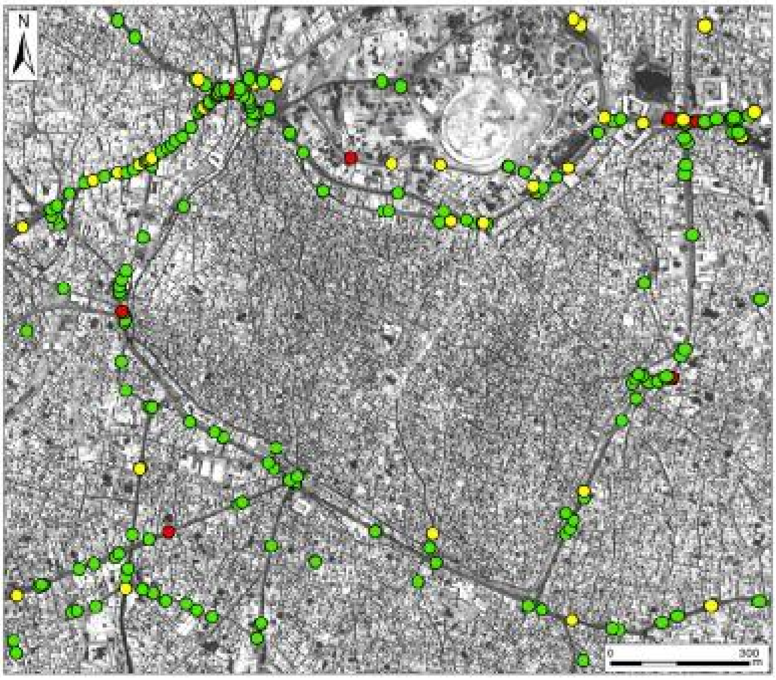


**Figure 5:** Study area. Map of moving vehicles classified by their speed range (green: slower; yellow: medium; red: faster). Image collected on Saturday June 6th 2010, 11:13:06 local time.

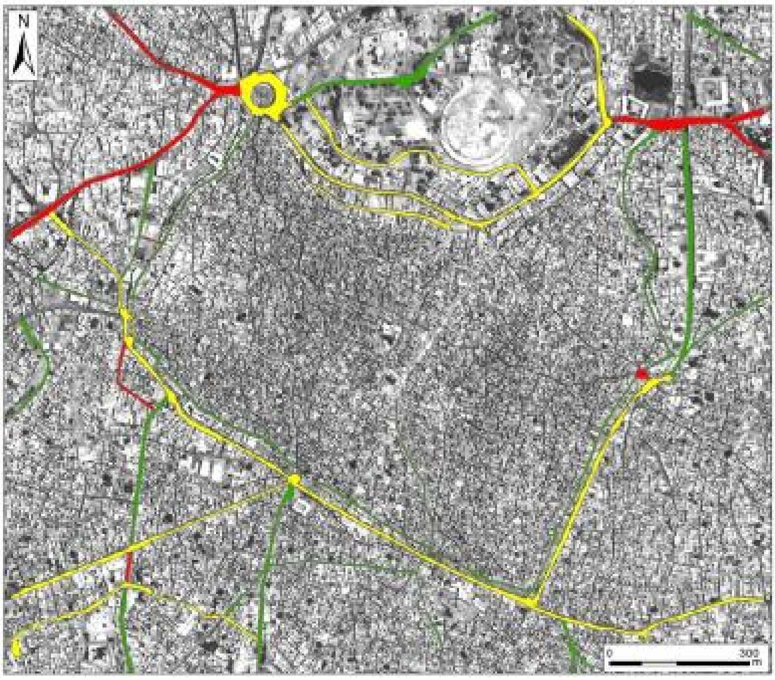


**Figure 6:** Study area. Map of vehicles density (green: low; yellow: medium; red: high). Image collected on Saturday June 6th 2010, 11:13:06 local time.

The accuracy assessment wasperformedonly for the detection of the moving vehicles. Results obtained through the automatic workflow described in Figure 3 have been compared with 219 moving objects manually detected by a trained operator. On the whole, the user's and producer's accuracies have been estimated to be 67% and 68% respectively. Hence, the algorithm was able to correctly detect about 70% of the moving vehicles on the road network without a priori knowledgeand in a very complex and dense urban area such as the historic core of Multan Walled City.

On the other hand, no data were available for evaluating the estimated speed. Nevertheless, results are in accordance with the knowledge of the traffic conditions for the Multan Walled City nearby. It is to highlight that both still and very slow vehicles, below the detection threshold of 5 km/h (3 mph), where not mapped because not belonging to the class "moving vehicles" and downtown Multan is full of stands, slow cattle trucks and other stuffs on the road network causing traffic jam.

## 5. CONCLUSIONS

In this study was evaluated the possibility to collect traffic-related data from single-pass WorldView-2 satellite imagery using a semi-automatic procedure developed for detecting, mapping and estimating speed and direction of travelof moving vehicles.

Even if the overall user's and producer's accuracies have good values, they are slightly worse than those reported in other similar studies (Mishra and Zhang, 2012; Salehi*et al.*, 2011; Zheng, 2006). However, most of the past researches focused on moving vehicle detection along fast-flowing roadssuch as highways or multi-lane roads where only fast cars or tracks occupy the road lanes, so the detection is easier. Contrariwise, this study addressed the complex task of traffic monitoring in a very dense and overcrowded urban area such as the historical core of Multan (Pakistan), characterized by narrow roads and bazaar streets, with the road network full of pedestrian, animal-drawn carts, little van and motorcycles so that the traffic speed is generally very low and the moving vehicle detection is difficult.The validation of results requires further efforts, especially in assessing the accuracy of the vehicles' speed and direction, not considered in this work,even if the speed rates estimated with the described method are reasonable considering the characteristics of the roads and vehicles in Multan.

Future researchwill focuseson the integration offully automatic object-based procedures for both road network and vehicles extraction. Overall, the present study demonstrated the capabilities to detect moving vehicles and traffic information from single-pass high-resolution satellite imagery, not only when dealing with highways or similar situations, but also in dense urban environments. This can open new frontiers in the use of optical Remote Sensing data for wide areal traffic monitoring and satellite data could also be integrated with ground based (local) monitoring stations. Furthermore, the same methodology can be applied to other civilian and military fields, like trains monitoring, ships monitoring or other unconventional vehicles tracking.

## ACKNOWLEDGEMENTS

This study was part of the project "Sustainable, Social, Economic and Environmental Revitalization in the Historic Core of Multan City", funded through the Agreement on debt for development swap between the Government of the Italian Republic and the Government of the Islamic Republic of Pakistan. Authors would like to thank FondazionePolitecnico di Milano (Daniele Bignami and Juan XabierMonjasKanpandegi) for the project management, Nexus Solutions Pakistan (Waqar Malik) and all DigitalGlobe'stechnical department for the technical support with WorldView-2 data.

## REFERENCES:

**References from Journals:**

Ahmed, K., 2011. Evaluation of Low Cost Technique "Indirect Right Turn" to Reduce Congestion at Urbanized Signalized Intersection in Developing Countries. Procedia - Social and Behavioral Sciences, 16, pp. 568-577.

Gianinetto, M., and Scaioni, M., 2008. Automated Geometric Correction of High-resolution Pushbroom Satellite Data. Photogrammetric Engineering and Remote Sensing, 74(1), pp.107-116.

Gianinetto, M., 2008a. Automatic digital terrain model generation using Cartosat-1 stereo images. Sensor Review, 28(4), pp. 299-310, ISSN 0260-2288.

Gianinetto, M., 2009. Evaluation of Cartosat-1 Multi-Scale Digital Modelling Over France. Sensors, 9, pp. 3269-3288.

Gianinetto, M., 2012. Automatic co-registration of satellite time series. The Photogrammetric Record, in press, pp.

1-9, doi: 10.1111/j.1477-9730.2012.00689.x.
Haralick, R.M., Shapiro, L.G., 1985. Image segmentation techniques. Computer Vision, Graphics, and Image Processing 29(1), pp. 100-132.
Johansen, K., Coops, N.C., Gergel, S.E., Stange, Y., 2007. Application of high spatial resolution satellite imagery for riparian and forest ecosystem classification. Remote Sensing of Environment, 110 (1), pp. 29-44.
Kittler, J., 1983. On the accuracy of the Sobel edge detector. Image and Vision Computing, 1 (1), pp. 37-42.
Larsen, S.Ø., Koren, H., Solberg, R., 2009. Traffic monitoring using very high resolution satellite imagery. Photogrammetric Engineering and Remote Sensing, 75 (7), pp. 859–869.
Liu W., Yamazaki F., Vu T.T., 2011. Automated Vehicle Extraction and Speed Determination From QuickBird Satellite Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4(1), pp.75-82.
Murray, H., Lucieer, A., Williams, R., 2010. Texture-based classification of sub-Antarctic vegetation communities on Heard Island. International Journal of Applied Earth Observation and Geoinformation 12(3), pp. 138-149.
Salehi, B., Zhang, Y., Zhong, M., 2011. Automatic Moving Vehicles Information Extraction From Single-Pass WorldView-2 Imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5 (1), pp.135-145.
Su, W., Li, J., Chen, Y., Liu, Z., Zhang, J., Low, T.M., Suppiah, I., Hashim, S.A.M., 2008. Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. International Journal of Remote Sensing, 29(11), pp. 3105-3117.
Xiong, Z. and Zhang, Y., 2008. An initial study on vehicle information extraction from single pass Quickbird satellite imagery. Photogrammetric Engineering and Remote Sensing, 74(11), pp. 1401-1411.

**References from Books:**

Baatz, M. and Schäpe, A., 2000.Multiresolution segmentation: an optimization approach for high quality multiscale image segmentation, In: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), AngewandteGeographischeInformations-Verarbeitung XII. WichmannVerlag, Karlsruhe, pp. 12–23.

**References from Other Literature:**

Kääb, A., 2011. Vehicle velocity from WorldView-2 satellite imagery. IEEE Data Fusion Contest 2011. 4p.
Gianinetto, M., 2008b. Multi-scale digital terrain model generation using Cartosat-1 stereo images for the Mausanne les Alpilles test site. The International Archive of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVII, pp. 1331-1335.
Mishra, R.K., Zhang, Y., 2012.Moving Vehicle Extraction from One-Pass WorldView-2 Satellite Imagery. Proceedings of Global Geospatial Conference 2012.
Rusmini, M., Candiani, G., Frassy, F., Maianti, P., Marchesi, A., Rota Nodari, F., Dini, L., Gianinetto, M., 2012. High-resolution SAR and high-resolution optical data integration for sub-urban land cover classification. 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2012), pp. 4986-4989.
Tao, J. W., and Yu, W. X., 2011. Preliminary Study on Imaging Time Difference among Bands of Worldview-2 and Its Potential Applications. 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 198-200.
Zheng, H., 2006. Automatic Vehicles Detection from High Resolution Satellite Imagery Using Morphological Neural Networks. Proceedings of the 10th WSEAS International Conference on COMPUTERS, Vouliagmeni, Athens, Greece, July 13-15, 2006, pp. 608-613.