

LAND USE/COVER REGISTRATION USING REMOTE SENSING DATA

B. Ruzgiene, J. Stadalninkaite

Dept. of Geodesy and Cadastre, Vilnius Gediminas Technical University, Sauletekio al. 11, LT-2040
Vilnius, Lithuania – Birute.Ruzgiene@ap.vtu.lt

Commission IV, WG IV/7

KEYWORDS: Photogrammetry, Remote Sensing, Land Cover, Interpretation, Updating, Database, Imagery, Accuracy

ABSTRACT

Remote sensing has provided an efficient method and perfect alternative of data acquisition for land use administration and land cover maps updating. Establishing Land information system in Lithuania there is taken into consideration not only Lithuania needs, but also requirements of European Union in land use classification. For such purpose it is investigated land cover/land use classification methods using satellite imagery and data interpretation defining imperfection of current situation, developing perspective recommendations. The strategy and approaches of geo-objects interpretation using *LANDSAT 7* imagery is outlined. The determination of most reliable land cover data have been carried out for maps at a scale of 1:50 000. Developed methodology includes determination of suitability criteria for certain group of geo-objects, applying mathematical statistical analysis and determining the confidence interval. Such methodology helps to emphasize the possibility of maps interpretation using remote sensing data. The research and accuracy investigation is based on a sample of satellite image, orthophoto map and digital vector database of Lithuanian Space Imagery Map. The accuracy analysis is referred to identifiable topographic objects that satisfy the defined suitability criteria calculating confidence intervals for each feature types. The results of research present mapping potentials for land use/land cover maps update purpose using remote sensing data.

1. INTRODUCTION

The creation of the topographic maps at a scale of 1:50 000 in Lithuania with the usage of satellite images started in 1993 within the agreement between National Geodesy Service and *SSC Satellitbild* (Swedish Space Corporation). Lithuanian Space Imagery Map at a scale of 1 50 000 (digital vector database *LTDBK 50 000-V*) was compiled using *SPOT* panchromatic and geometrically corrected multispectral images. The database is periodically updated on the base of orthophoto images at a scale of 1:10 000. Large interest is directed to application of satellite images for land parcel identification.

Satellite imagery from *LANDSAT* is used for land use/cover registration and maps updating at a medium scale, so the objective of research was to analyze on a sample of *LANDSAT 7* satellite images suitability for constructing and updating maps at a scale of 1 50 000. Such task was introduced determining identifiable land cover objects by defining the suitability criterion of objects, extracted from satellite imagery and investigating their planimetric accuracy. The vector database *LTDBK* was taken as a reference for evaluation of recognizable and interpretable objects quality from satellite imagery (Ohlhof, T., 2000) and their suitability to update this database. The different types of data for investigation were used: *LANDSAT 7* imagery (panchromatic image with 15 m ground resolution and multispectral image with 30 m ground resolution); orthophoto images; digital vector database *LTDBK 50 000* (first edition according to *SPOT* imagery) (Sjovall, A., 2001) as well as updated vector database using orthophoto images. The *ERDAS Imagine* software package have been use for pre-processing of satellite images (Ebadi, H., Azizmohammadi, M., 2001). *LANDSAT 7* imagery was converted to vector format that to determine the accuracy of land cover points position by digitizing in different data sources.

2. A SAMPLE OF REMOTE SENSING, PHOTOGRAMMETRY AND GIS DATA

The combination and separate usage of satellite images allow to evaluate the area to be investigated, perform visual and computer-aided land interpretation. Satellite data serves as an input data for further investigation. The digital orthophoto images supplies useful information about land use and provides visible ground representation. It is important information source for satellite image georectification and interpretation purpose, rather good background information for vector data edition. The digital database supplies vector information about current study area and refers to a GIS data with possibility to recognize and analyze the spatial relationships among mapped phenomena.

The investigation area (50x50 km) covered by *LANDSAT 7* satellite imagery (multispectral and panchromatic with 15 m ground resolution) is central part of Lithuania (Figure 1). Various ground objects gives a possibility to detect and evaluate land cover elements and to perform visual interpretation and comparison within different data sources. The satellite images supplied by *Eurimage* have a 1G System Corrected level (radiometrically and geometrically corrected, including output map projection, image orientation (UTM) and resampling algorithm (Cubic Convolution), is free from distortions related to the sensor, satellite (attitude deviations from nominal) and Earth (rotation, curvature). No atmospheric corrections are applied for images and do not employ ground control or relief models to obtain absolute geodetic accuracy. The purpose for usage of satellite image is based on information extraction. On process which includes the determination of spatial object position and land use identification. Considering the goal of investigation is more related to object position and geometric location.



Figure 1. *LANDSAT 7* satellite multispectral (with 7 bands) image over the central part of Lithuania

Other data sources used for research are: orthophoto maps and vector database of the Lithuania Space Imagery Map at a scale of 1:50 000. About 80% of Lithuania territory is covered by compiled orthophoto maps at a scale of 1:10 000. The object of investigation consists of 12 orthophoto maps which are constructed at National Coordinate System *LKS-94* with resolution of 0,5 m and supplied in *TIFF* format.

LTDBK 50 000 vector database was created for the whole Lithuania territory (65 000,3 sq. km) with the help of existing maps information, digital databases together with panchromatic orthophoto material from *SPOT* satellite, with geometrically rectified multispectral space image and consist of 135 map sheets. Resolution of digital information of the panchromatic image is 10 meters, for multispectral images – 20 meters. The database is created on the background of reference ellipsoid GRS80, Transverse Mercator Projection with central meridian $L_0 = 24^0$. The database *LTDBK 50 000* is accepted as reference investigating vector information obtained after classification of satellite image and conversion to vector format.

3. *LANDSAT 7* IMAGERY POTENTIALS IN LAND COVER REGISTRATION

3.1 Land Cover Classification

During visual interpretation (human factor) of *LANDSAT 7* satellite imagery was decided to segment existing feature into main topographic feature classes, because human visual acuity does not allow to identify all spectral differences in imagery (Koarai, M., Kadowaki, T., Watanabe, N., Matsuo, K. 2001). After visual image investigation, land cover was classified into five types: 1) agricultural areas; 2) urban areas; 3) forestry (vegetation); 4) hydrographic objects; 5) roads (as linear objects). Indication of main deciphering characteristics of each topographic object has been made in order to simplify object interpretation. Individual areas are recognizable by its spectral response in color bands.

The digital interpretation of satellite imagery have been made using *ERDAS Imagine* software package. The major problem in land cover supervised classification – mixed pixels there was defined. During revision of classified image was detected following problematic areas:

- many agricultural areas occurred as garden land cover class;
- spectral signature of sand objects appears in agriculture land cover class.

Mixed pixel values were detected in urban and agricultural areas. In order to avoid occurrence of such confusion approaches are suggested: usage of multitemporal images to

individualize classes (especially for agricultural land cover class); usage textural information to improve classification accuracy; usage GIS procedures based on auxiliary data (Parseliunas, E., 2001).

In unsupervised classification spectral classes were grouped first, based on the numerical information in the data – 24 classes were defined. Unsupervised classification is not completely without human intervention. However, it does not start with a pre-determined set of classes as in a supervised classification. However, the objective of research also requires linear objects such as roads, railways, highways, etc. Extraction of road axial line is less or even not affected by any environmental or seasonal factors. Considering the time of capture of satellite images the road cover or railway track is quite easily distinguishable. The extraction of road objects was performed by manual (interactive) digitizing on the view. Digitization process performed on merged image, which supplies ground resolution of 15 m and keeps predefined multispectral bands. The inaccurate in operator digitizing, also image georectification accuracy were taken into account.

3.2 Determination of Suitability Criteria for Database Updating

Using satellite imagery for land cover registration, it is important to clear how large objects can be interpreted and drawn by satellite imagery. Here was researched the possibility of interpretation for updating topographic map at a scale of 1:50 000 using *LANDSAT 7* imagery. All recognizable and identifiable objects in satellite imagery was categorized or segmented referring to the structure of geodata grouping in digital vector data base *LTDBK 50 000*. It is defined the thematic information presented in different data sources has large differences in object interpretation. For this reason it is necessary to determine and calculate the mathematical areas for the whole territory and for the separate object classes in order to evaluate objects variability. The distribution of interpretable object classes' area in satellite imagery is similar to *LTDBK*. It means that extracted information from *LANDSAT 7* represents the correct general disposition of geographic data. There are no large clear deflections as well. But in some object classes there are substantial changes of area, particularly open pits and gardens. In research area, open pits and gardens are mostly affected by nature and human factors. The comparison of calculated mathematical areas have been made which gives general overview about variability of object classes collected from different data sources (Figure 2).

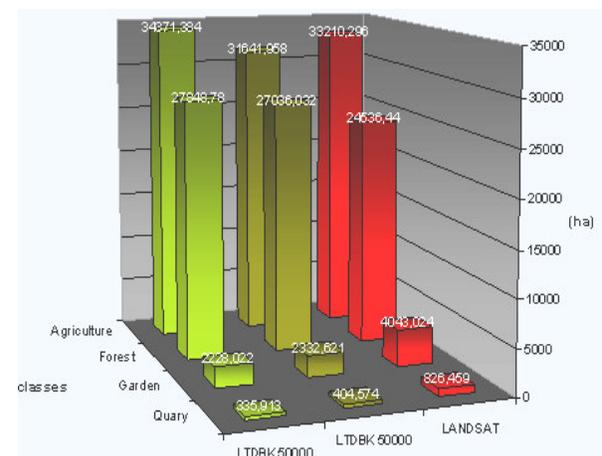


Figure 2. Areas of identified object groups

In order to determine *LANDSAT 7* vector data suitability for updating of *LTDBK 50 000*, there was defined the criterion for data selection calculating area differences and determining an average value. The area differences are directly proportional to their percent expression of inter-equivalence. The mean percent expression is obtained 76,9%. This value is accepted as suitability criteria for data selection. All data below this limit were rejected as data which quality is not efficient for updating. Value of criterion depends on quality of reference data as well. There was defined that the updating of database using *LANDSAT 7* vector data is possible just for four object classes excluding gardens and open pits object classes (Figure 3).

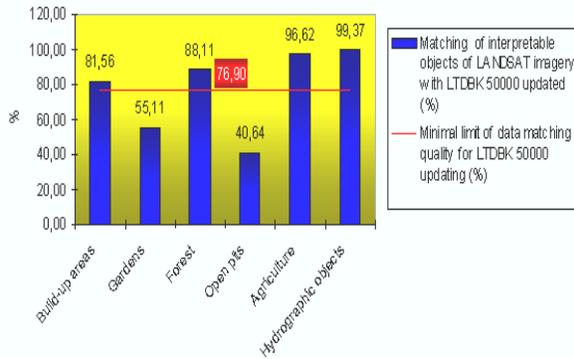


Figure 3. Suitability of *LANDSAT 7* vector data for updating of database

3.2 Accuracy Investigation and Statistical Evaluation

After determination the suitable topographic objects for updating the vector database there was investigated the planimetric accuracy of land cover features (Vainauskas, V., 1997). The topographic features that correspond to the defined criteria are: Water (Hydrographic objects), Build-up areas, Forest and Agricultural land. The investigation of planimetric accuracy is performed for such identified objects. There was used *ARC/Info* software package for measurement coordinates within created model of 115 points for each topographic object group. The picked out points were digitized and coded, depending on which topographic object group and data source they are attributed (e.g., WaLa22 Wa – Water object; WaLt22 Or – point collected from *LTDBK* database, etc.). On a basis of measured coordinate values of topographic objects, there was calculated coordinate discrepancies between vector data obtained from *LANDSAT 7* imagery and *LTDBK* and RMSE was calculated (Table 1).

Topographic objects group / RMSE (m)				
	Water	Urban	Forest	Agriculture
m_x	3,13	17,92	42,45	33,96
m_y	7,52	30,76	27,50	24,57

Table 1. Evaluation of land covers accuracy from *LADSAT 7* imagery

In order to determine, does the accuracy correspond to requirement, the analysis has been accomplished. The results of distinctive point identification in *LANDSAT 7* imagery are influenced by: georectification error; inaccurate image classification and point interpretation in raster data. According to standards, RMSE should not be higher then 10,0 m.

The largest coordinate discrepancies were obtained in a point positions where shape of object are significantly changed due to the natural (hydrograph and forest) or human (urban and agriculture areas) intervention.

The mathematical statistical evaluation consists of confidence interval calculation and determination for each of topographic object group. The interval of confidence indicates that with a certain probability all new picked up data set will be in the range of defined confidence interval. (Tamutis, Z., Zalnierukas, A., Kazakevicius, S., Petroškevičius, P., 1996). Determination of this interval, allows forecast the distribution of all model points. The probability of confidence interval was assumed 99%. Such probability was selected to find out maximal interval for maximum quantity of new data. The interval of confidence is calculated on the base of linear regressive analysis. In order to determine the suitability of analyzed data for linear regressive analysis. The test of Smirnov-Kolmogorov was applied. This test determines whether collected data has abnormal or normal distribution. Properties of collected data are: coefficient W defined by theory of Smirnov-Kolmogorov test and t_{max} , which indicates the largest differences between the same point in theoretical and practical curves (distance between theoretical and practical curves). The decision that the distribution is normal or abnormal is based on this distance and. the conditions are: hypothesis H_0 – set of data is of normal distribution; H_1 – set of data has abnormal distribution. With a probability of 99%, the correct hypothesis will be accepted or wrong hypothesis will be rejected. The conditions for correct results are: if $t_{max} > W$, then this set of data has abnormal distribution; if $t_{max} < W$, then this set of data is of normal distribution, it means that data has the properties of normal distribution. It was defined, that all collected data is normal distributed (e.g., see Figure 4).

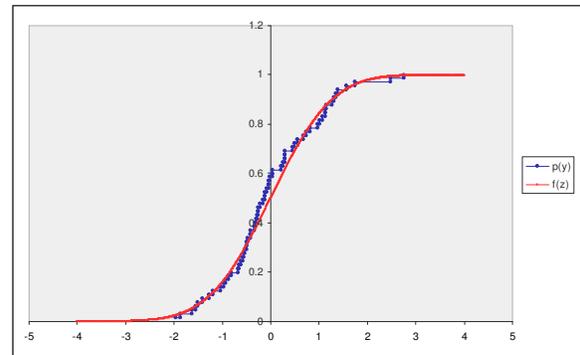


Figure 4. Distribution of abscissa discrepancies of Water object regarding theoretical curve

In figure 5 there are presented the upper and lower limit tangents of coordinate discrepancies (dx , dy), also known as regressive tangents these lines shows, that coordinate discrepancies can not exceed this limit, with particular probability. This is the tangent, optimally diverged from every collected point from data set. If all collected points would coincide with regressive tangent, discrepancies between coordinates of the same points would be equal to zero. Increasing coordinate discrepancies, regressive tangent will bend; intervals extend respectively upwards and downwards. The central part of interval presents most reliable data, but receding to the edges, reliability of the data decreases.

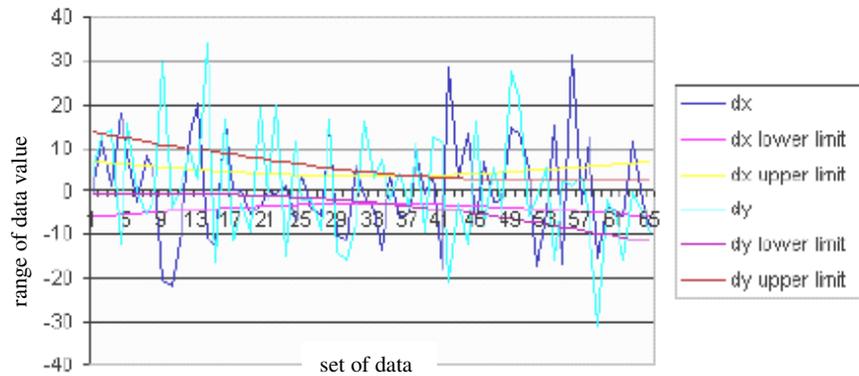


Figure 5. Confidence interval for coordinate discrepancies of Water area object

4. CONCLUSIONS

The Maximum Likelihood Classifier is most commonly used than any parametric classifiers and is well suited for accurate classification. It assumes that the input data are of normal distribution and independent. The biggest attention should be paid to the semi-automatic land cover interpretation. The results of supervised classification show that some confusion has been detected within classified image. In order to avoid occurrence of such confusions and improve the spatial classification, following approaches are suggested:

- usage of multitemporal images to individualize information classes that were confused in a single-data image (a biggest influence for a agricultural land cover);
- usage textural information to improve results of classification;
- usage GIS procedures based on auxiliary data.

The determination of vector data suitability for topographic maps updating consists of mathematical area calculation of topographical objects from reference data as well from satellite imagery data. The suitability criteria 76.9% indicates, that more than half of identified topographic features from satellite imagery could be used for map update. However, defined suitability criterion depends on accuracy of reference database.

By the investigation of planimetric accuracy there was determined the coordinate accuracy of identifiable topographic objects from satellite imagery. Regarding to sample data mathematical statistical approach with a probability of 99% and calculation of confidence intervals for each of topographic feature types should be applied. All distinctive points have to be collected in a random manner. Otherwise the application of suggested methodology would be inappropriate. Data amount compiling equal sets of points in each group of topographic objects is necessary for determination of confidence interval for vector data obtained from satellite imagery.

According to the research of usage of *Landsat 7* satellite data for land cover registration with integration to reference database (depending on the required accuracy), only Water (Hydrographic) object class satisfies the accuracy requirement.

REFERENCES

- ERDAS *Imagine*, 1997. Tour Guides V8.3, ERDAS, Inc. Atlanta Georgia, 451 p.
- Digital vector data base of Lithuanian Space Imagery Map at Scale 1 50 000 (LTDBK 50 000), 2001. Specification, Version 1.6, 23 p. (in Lithuanian).
- Ebadi, H., Azizmohammadi, M., 2001. Merging SPOT panchromatic data with multispectral data (XS) using wavelet and comparing with IHS method. The 20th International Cartographic Conference, Vol. 4, ICC2001 Beijing, China, pp. 669-675.
- Ohlhof, T., 2000. Visibility of VMAP2 data in aerial and satellite imagery. International Archives of Photogrammetry and Remote Sensing, Vol XXXIII, Part B4, Amsterdam, pp. 764-767.
- Koarai, M., Kadowaki, T., Watanabe, N., Matsuo, K., 2001. Interpretation characteristics of IKONOS imagery for mapping and coordinate accuracy of IKONOS orthoimagery. The 20th International Cartographic Conference, Vol 4, ICC2001 Beijing China, pp.712-718.
- Sjovall, A., 2001. Research of updating digital cartographic databases by IKONOS digital satellite images. Geodesy and Cartography, No 2, Vilnius: Technika, pp. 61-67 (in Lithuanian).
- Vainauskas, V., 1997. Quality control of digital mapping. Geodesy and Cartography, No 1 (25). Vilnius: Technika, pp. 71-77 (in Lithuanian).
- Parseliunas, E., 2001. Geoinformational systems: Technology. Vilnius: Technique, 229 p (in Lithuanian).
- Tamutis, Z., Zalnierukas, A., Kazakevicius, S., Petroskevicius, P., 1996. Geodesy. Vilnius: Science and Encyclopedia, p. 384 (in Lithuanian).
- Eurimage*. Multi-mission satellite data. Products. <http://www.eurimage.com/Products/landsat.shtml> (accessed December,2002).