SEMI-AUTOMATIC HYPERSPECTRAL IMAGE CLASSIFICATION OF URBAN AREAS USING LOGISTIC REGRESSION

Dirk Borghys, Michal Shimoni and Christiaan Perneel

Royal Military Academy, Signal & Image Center, Brussels, Belgium;
[Dirk.Borghys, Michal.Shimoni, Christiaan.Perneel]@rma.ac.be

ABSTRACT

Urban areas are among the most dynamic regions on earth, continuously and rapidly changing. For monitoring these changes, remote sensing has proven over the years to be a reliable source.
Current airborne hyperspectral systems with spatial resolution of a few meters, combined with very high spectral resolution, facilitate the urban scene analysis by allowing to distinguish small details in the urban environment. This paper presents part of a project aiming to classify man-made objects using hyperspectral images and to investigate the complementarity between hyperspectral and SAR data. The intention is to develop methods that are able to quickly obtain an overview of the current situation and require as little human intervention as possible. This is very important for various applications related to disasters, e.g. emergency cartography, disaster monitoring, damage assessment, mission planning, etc.

The paper describes a new method for classifying the main classes in an urban environment using hyperspectral data. The method is based on logistic regression (LR), which is a supervised multi-variate statistical tool that finds an optimal combination of the input channels for distinguishing one class from all the others. LR thus results in detection images per class that can then be combined into a classification image. The LR uses a step-wise method that implicitly performs a channel selection. The method is supervised in the sense that existing digital maps are used for learning. However, the method does not require the laboratory spectra or extensive ground truth.

The method is applied on HyMAP data of an urban area in the South of Germany. The results of the proposed approach are compared to classical methods. Furthermore, a sensitivity analysis is presented, which investigates the robustness of the detection of the different classes against various influences and in particular the influence of channel width and pre-processing level.

The classification results are better than those obtained by a classical method. The sensitivity analysis shows that the pre-processing level applied to the hyperspectral data does not influence the classification results significantly for this application. Furthermore, reducing the number of channels results in a drop of performance for some classes only when less the number of channels becomes inferior to 40.

INTRODUCTION

For many remote sensing applications e.g. emergency cartography, disaster monitoring, damage assessment, etc., it is important to be able to quickly obtain an overview of the current situation and in particular the detection and classification of man-made objects. This paper presents part of the work done in a project that aims to classify man-made objects using hyperspectral images and to investigate the complementarity between hyperspectral and SAR data. The paper describes a new method for classification of hyperspectral data for updating maps. The primary objective is to classify the mayor carthographic classes in urban areas. Six classes of interest are defined: roads, buildings, railways, paths, forests and background that mainly consists of different types of vegetation.

In classical hyperspectral classification methods, e.g. the spectral angle mapper[i] (SAM) or spectral unmixing methods[ii], the classes above would be too general. A class like ``Buildings'' would for instance have to be sub-divided according to the color of the roofs. This would require a ground truth mission or an extensive image interpretation by a human expert. As our intention was to use only information from existing maps and to limit user intervention to a minimum, methods like spectral unmixing or the SAM could not be used.

The proposed method is supervised in the sense that existing digital maps are used. The maps serve as the sole basis to construct the learning and validation databases. The idea is to avoid having to undertake an extensive ground truth mission for constructing the learning set. The classification is based on a multi-variate statistical technique called logistic regression[iii] (LR).

LR combines the different channels in a way that optimises the distinction between one class and all the others. It thus results in a detector for each class. The detection results are then combined into a classification image. The LR uses a step-wise method that only adds a channel to the used set of channels if the improvement in detection caused by this addition is statistically significant. LR thus implicitly performs a channel selection.

The results are compared to those that are obtained by a conventional hyperspectral method: the matched filtering[iv] (MF). MF is also developed to detect a class among others and is thus comparable to our method at various stages.

In conventional methods for hyperspectral image classification, the first step is a channel reduction based on Minimum Noise Fraction[v] (MNF) or Principal Component Analysis (PCA). These methods use the statistical variation in the data set but they do not take into account information about the classes of interest to lead the channel selection. In LR the channel selection is based on information about the classes of interest. A comparison of the channel selection method obtained by LR with the classical ones was already published[vi]. That paper also compares the detection performances for each class versus the detection threshold with those obtained with MF.

The current paper focusses on the final classification results. Furthermore a sensitivity analysis is presented, which investigates the robustness of the detection of the different classes against various influences and in particular the influence of channel width and pre-processing level. For each of them the impact on the results of the LR-based method is compared to those obtained by the matched filter.

DATA SET

Hyperspectral Data

For this project HyMap data were acquired over the towns of Oberpfaffenhofen and Neugilching in the South of Germany. The HyMap was operated by the German Aerospace Agency DLR and contains 126 contiguous bands ranging from the visible region to the short wave infrared (SWIR) region (0.45 - 2.48 µm). The bandwidth of each channel is 15-18 nm. The spatial resolution is 4 m at nadir and the image covers an area of 2.5 by 10 km.

The following pre-processing was done by the Flemish Institute for Technological Development (VITO): radiometric correction (level1), geocoding (level 2), atmospheric correction using ATCOR4[vii] (level 3) and the ``Empirical Flat Field Optimized Reflectance Transformation'' EFFORT[viii] (level 4). We received the data corresponding to the four different levels of pre-processing. In all processing the first and last channel were discarded. The first contains too much noise while the last is saturated.

Ground Truth Data

Cadastral digital data (from the ``Bayerischen Vermessungsverwaltung'') of Neugilching as well as digital topographic maps of the surroundings (ATKIS data from the ``Bayerishes Landesvermessungsambt'') were acquired. The topographic and cadaster maps were used to construct the learning set for the logistic regression and for the validation process. For each of the six classes of interest, i.e. Roads, Buildings, Paths, Railways, Forests and Background, approximately 140 individual points were selected to constitute the learning set. The remainder of the scene is used for validation. Figure 1 shows the dataset used in this paper. The left image shows part of the HyMap dataset in RGB color combination, the right image is the corresponding ground truth map used for validation. The railway on the image is in fact a tramway (a S-Bahn).

Figure 1: Part of the original hyperspectral data in RGB color composition (left) and ground truth map (right) Legend: red: roads, black: buildings, yellow: paths, orange: railways, green: forests, grey: background

METHODS

Classification method

The proposed classification is based on logistic regression (LR), which was developed for dichotomous problems where a target class has to be distinguished from the background. The method combines the input features (the input channels in this case) into a non-linear function, the logistic function, defined as:

$$p_{x,y}\left(\text{target} \mid \vec{\mathrm{C}}\right) = \frac{\exp\left[\beta_0 + \sum_i C_i(x,y)\beta_i\right]}{1 + \exp\left[\beta_0 + \sum_i C_i(x,y)\beta_i\right]}$$

$p_{x,y}\left(\text{target} \mid \vec{C}\right)$ is the conditional probability that a pixel (x,y) belongs to the considered class (target class) given the vector of input channels $\vec{C}$ at the given pixel. $C_i(x,y)$ is the value of pixel (x,y) in the i^{th} channel of the dataset. The LR finds a combination of the input channels that approximates optimally the $p_{x,y}\left(\text{target} \mid \vec{C}\right)$ for each class, based on learning data.

The logistic regression (i.e. the search for the β_i is carried out using Wald's forward step-wise method using the commercial statistics software SPSS. In the Wald method, at each step, the most discriminating channel is added and the significance of adding it to the model is verified. This means that only the channels that contribute significantly to the discrimination between the foreground and the background class are added to the model.

The logistic regression thus gives an optimal combination of a sub-set of input parameters for separating one class from all others, implicitly performing a channel selection.

Applying the logistic function on the complete image, a detection image for the considered class is created. The values in the detection images are proportional to the probability that the pixel belongs to the class. The detection images for the different classes are combined into a classification image by attributing to each pixel the class corresponding to the highest value in the detection image.

Sensitivity analysis

Apart from an evaluation of the classification results, this paper also investigates how the classification results for the main land-cover classes in urban areas are degraded when the data that are used are less optimal. This is the sensitivity analysis.

In particular, the influence of the number of bands (sub-sampling) in the hyperspectral image and the level of the pre-processing were investigated.

For studying the influence of the pre-processing, the classification was applied to the data resulting from the different levels of pre-processing. As maps are needed for learning and validation only the levels after geoding (level 2 –4) are considered in this study.

For the investigation of the influence of the number of channels, the data were sub-sampled. This was done using ENVI[viii] . The channels in the original HyMap data seem to be uniformly spread over three different bands (cf. top line in Figure 2). In the sub-sampling a similar structure is kept, i.e. the proportion of channels in each of the 3 bands is kept constant. In each band the centers of the available channels are uniformly spread and for the actual sub-sampling a Gaussian model with FWHM (full width half maximum) equal to the channel spacing is applied. The two classification methods were applied to each sub-sampled data set.

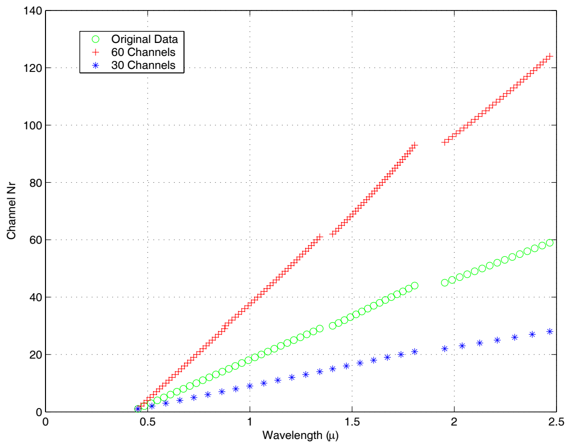

Figure 2: Principle of the sub-sampling method. Channels number vs. wavelength for original data set and the sub-sampling to 60 and 30 channels

Evaluation method

The evaluation of the results is based on the ground truth image of which a part is shown on the right side of Figure 1. Results of classification evaluated based on the confusion matrices. From the confusion matrix the user accuracy (UA) and producer accuracy (PA) are calculated. These serve to compare the classification results of the different methods as well as to study the effect of the pre-processing level.

For the study of the effect of the sub-sampling, the classification results were compared using a figure-of-merit for target detection[ix] defined as:

$$FOM = \frac{Ndt}{Ntt + Nft}$$

where Ndt is the number of correctly detected target pixels, Nft the number of false alarm pixels and Ntt the number of true target pixels actually present in the image.

RESULTS AND DISCUSSION

Figure 3 shows the results of the classification for the LR and the MF. The LR results look better than the MF ones. MF gives a lot of false alarms due to railways. It does seem to obtain a better detection of the roads. LR is better for building detection although the size of the buildings is slightly to large as compared to the ground truth image. MF classes some buildings either as roads or railways. LR has some false alarms due to paths and forests between the houses. MF presents a pepper-and-salt classification error, which could be easily removed by post-processing.

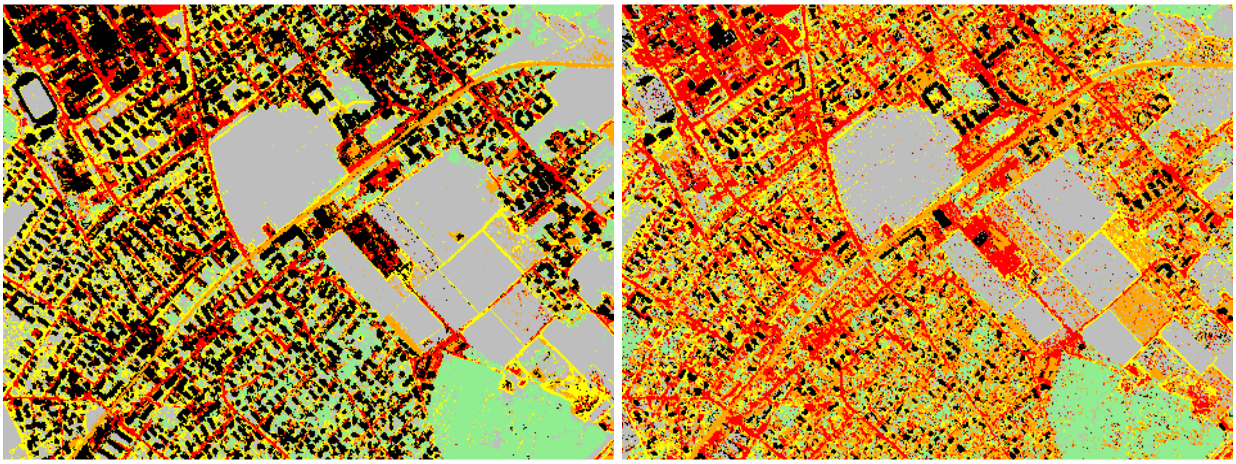

Figure 3: Classification result (left:LR, right:MF). Legend: Red:Roads, Black:Buildings, Yellow:Paths, Orange:Railways, Green:Forests, Grey:Unclassified

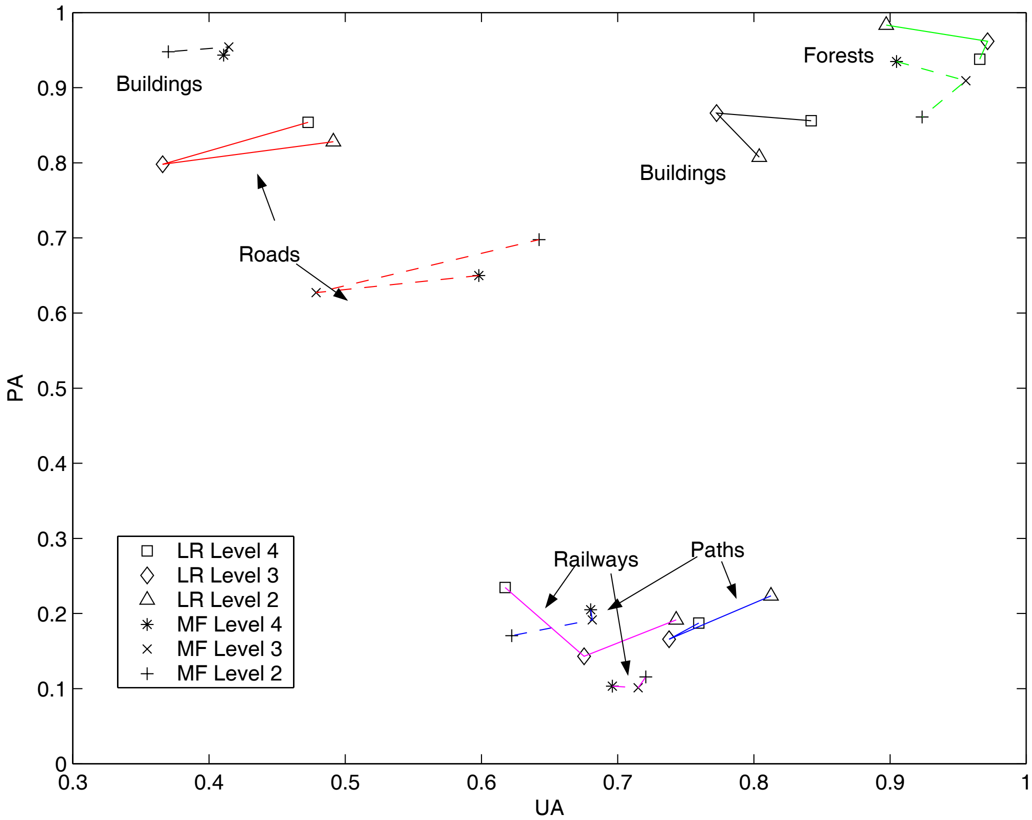

Figure 4: Classification results for LR and MF for the various classes and pre-processing levels. (Color Legend: Red: Roads, Black: Buildings, Magenta: Railways, Blue: Paths, Green: Forests)

Figure 4 shows the results of the classification as a plot of Producer Accuracy (PA) versus User Accuracy (UA). The results for different classes are represented by different colors. The solid lines connect results of the LR for a given class obtained for the three levels of pre-processing. The symbols on the lines represent the pre-processing level that was used for the classification (triangles: level2, diamonds: level 3 and squares: level 4). The dashed line represents the results of the MF in a similar way. The pre-processing levels are indicated here by following symbols: + level 2, x level 3, * level 4.

The best results are obtained for forests where PA and UA are near 90% for both methods and all pre-processing levels. For buildings UA is much higher for LR than MF at the expense of a slightly lower PA. This means that the probability of detection for buildings, given by LR is higher than by MF. The same is true for paths. For roads the UA is better for MF than for LR, but LR obtains a higher PA (less false alarms). Railways and paths have a UA of around 70% but a very high false alarm rate (low PA). As these are both linear objects, results could be improved by applying a post-processing that deletes isolated occurrences.

For the influence of the pre-processing level is concerned, it is hard to draw a general conclusion. For most classes the results for level 2 and 4 are close together in Figure 4. Level 3 is either better (forests and railways) or less good (buildings and roads). For paths level 2 gives the best results for LR and the worst results for the MF.

Figure 5 and Figure 6 show the results of the sensitivity analysis for sub-sampling. The figures show the FOM versus the number of channels that was used for the classification. Each plot presents the results of a different class and each time the LR and MF result is given. For buildings, railways and forests, the results for LR are better than those obtained by MF. For paths and roads the results of both methods are similar. For paths and railways both methods give a low FOM. Examination of the confusion matrices shows this is due to the high false alarm rate. Paths are very narrow and bordered with vegetation. They are confused with vegetation. However results for linear objects as paths and railways could be improved by post-processing. For forests both methods give very good results (FOM above 0.8), even at a high sub-sampling rate (low number of channels). The behaviour of the FOM with the number of channels is somewhat unexpected. Apparently for most classes, even dropping the number of channels to 20 does not influence the detection result significantly. Only for a few classes (roads, LR for buildings and MF for paths) there is a clear drop in performance when the number of channels gets below 30-40.

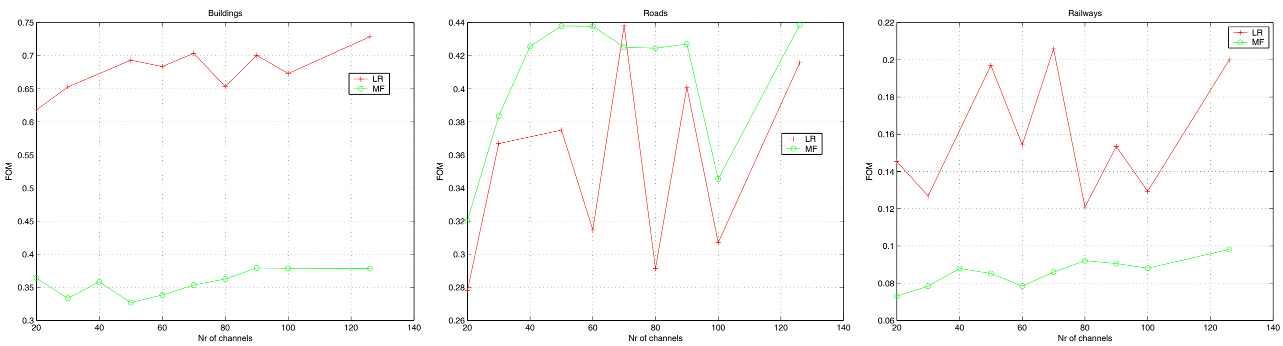

Figure 5: Sub-sampling sensitivity analysis results: FOM vs. Nr of channels (left: Buildings, middle: Roads, right: Railways)

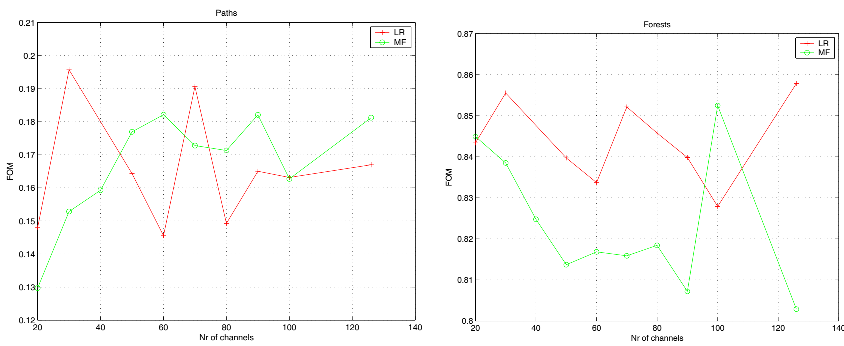

Figure 6: Sub-sampling sensitivity analysis results: FOM vs. Nr of channels (left: Paths,right: Forests)

CONCLUSIONS

A method, based on logistic regression (LR), was developed for classifying the main carthographic categories of urban areas using hyperspectral images. The method detects each class separately and combines the detection images into a classification. It was applied to a HyMap image of a village in the South of Germany. Results were compared to those obtained by a matched filter (MF). LR gives better results than MF for detecting buildings, forests and railways. For paths and roads results are comparable. Classification results are very good for forests and buildings. For roads, some gaps are present. For paths and railways the results are degraded by a large number of false alarms. As these are linear objects, results could be improved using a post-processing method that eliminates isolated occurrences of these classes. This is a topic for further work.

The paper also presents part of a sensitivity analysis. Rather unexpectedly, the pre-processing level applied to the hyperspectral data does not influence the classification results in a significant manner for this application. This is probably due to the fact that the defined classes are very rough and do not required a detailed analysis of the spectra. This could also explain the second result of the sensitivity analysis, i.e. the performance does not change dramatically when the number of channels is reduced. Only when less than 40 channels are used, the performance drops for some of the classes.

Later, we intend to investigate the robustness with geographic location, i.e. applying the logistic function determined in one location to an image acquired at another location. We also intend to investigate the influence of a miss-calibration of the image by studying the degradation of classification results when some channels are spectrally shifted.

ACKNOWLEDGEMENTS

The research presented in this paper is the result of a collaboration between three projects: HYSAR: “Man-Made object classification using fused polarimetric SAR and hyperspectral imagery data", funded by the Belgian Federal Science Policy Office, in the frame of the STEREO Program (project nr. SR/00/044); a project funded by the EC in the frame of the PASR Program: Advanced Space Technology to Support Security Operations (ASTRO+) and a study funded by the Belgian MoD on the semi-automatic use of hyperspectral and SAR images for general security problems project nr. F05/02.

REFERENCES

[i] F. Kruse, A. Lefko, J. Boardman, K. Heidebrecht, A. Shapiro, P. Barloon, and A. Goetz, 1993. The spectral image processing system (sips) - interactive visualization and analysis of imaging spectrometer data, Remote Sensing of Environment, Special issue on AVIRIS, 44:145-163

[ii] N. Keshava and J. Mustard, 2002. Spectral unmixing, IEEE SP Magazine,19:44-57

[iii] D. Hosmer and S. Lemeshow, 1989. Applied Logistic Regression, John Wiley and Sons.

[iv] J. Boardman, 1998. Leveraging the high dimensionality of AVIRIS data for improved sub-pixel target unmixing and rejection of false positives: Mixture tuned matched filtering, Summaries of the Seventh Annual JPL Airborne Geoscience Workshop. Pasadena, CA. 55.

[v] A. Green, M. Berman, P. Switzer, and M. Graig, 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal, IEEE-GRS 26(1):65-74.

[vi] D. Borghys, M. Shimoni, C. Perneel, 2005. Supervised classification of hyperspectral images using a combination of spectral and spatial information, In Proc SPIE Image and Signal Proc. For Remote Sensing XI, ed. L. Bruzzone, SPIE Vol 5982.

[vii] R. Richter, 2005, Atcor-4 user guide, v4.0, DLR.

[viii] ENVI User's Manual (v4.0), Research Systems Inc., Boulder, 1993.

[ix] D. Borghys, L. Sévigny, T. Sams, R. Gabler, B. Hoeltzener, P. Schwering, J. Haddon, and J. Knecht, 1999. Multi-Sensor Image Fusion for the Detection of Targets in the Battlefield of the Future , Final report of a Cooperative Project in Image Processing conducted by RTB(SET-TG01)., NATO, Brussels.