AUTOMATIC ROAD EXTRACTION FROM LIDAR DATA BASED ON CLASSIFIER FUSION IN URBAN AREA

F. Samadzadegan[a], B. Bigdeli[a], M. Hahn[b]

[a]Dept of Geomatics Engineering, Faculty of Engineering, University of Tehran, Tehran, Iran, (samadz, bigdeli)@ut.ac.ir
[b]Stuttgart University of Applied Sciences, Stuttgart, Germany, m.hahn.fbv@fht-stuttgart.de

KEY WORDS: LIDAR, Road Extraction, Classifier Fusion, Classifier selection, Genetic Algorithm

ABSTRACT:

LIDAR is a powerful remote sensing technology for acquisition of 3D information from terrain surface. Algorithms used for LIDAR data in urban area are used mostly to deal with the 3D points cloud and to identify objects, such as buildings, trees and roads. In contrast to the well studied problem of building and tree detection from LIDAR data, the detection of roads from LIDAR is in its formative years. Road detection from remotely sensed data is a difficult problem that requires more research due to the many unsolved questions related to scene interpretation. Existing road extraction techniques are characterised by poor detection rates and the need for existing data and / or user interaction. To improve the potential of road extraction process, other information in addition to the height of cloud points is required, such as laser intensity. However, few algorithms for classification using intensity data have been deeply investigated. The laser intensity differs from material to material. The intensity of reflection for most instances of the same material is similar; while pulse intensity from different materials differs. This article deals with using as much of the recorded laser information as possible thus both height and intensity are used. To extract roads from an LIDAR point cloud, a multiple classifier system is used to classify the LIDAR points into road and other non-road objects. We experiment classifier selection and combination in classification of LIDAR data over an urban area, wherein we aim to select an optimal or sub-optimal subset of classifiers from available combination candidates through an evolutionary strategy. The performance of the selected classifiers is measured by the combination accuracy using plurality of votes. The obtained results show that optimum subset classifier selection, improves the combination performance compared to the combination of all classifiers.

1. INTRODUCTION

Automatic extraction of roads in complex environments is one of the challenging issues in photogrammetry and computer vision, since many tasks related to automatic scene interpretation are involved (Hinz, 2004). Modelling of the roads required for a variety of tasks such as urban planning, network planning for mobile communication and tourism information systems.

For a long time, the only data sources, used for road extraction, were aerial and satellite imagery. Approaches designed to process satellite or low-resolution aerial images generally describe roads as curvilinear structures (Wiedemann and Ebner, 2000; Gecen and Sarp, 2008; Heipke, 2007), while those using large scale imagery (i.e., a ground resolution less than 1 m) model roads mostly as relatively homogeneous areas satisfying certain shape and size constraints (Baumgartner et al., 1999; Hinz, 2004).

More recently, advancements in LIDAR enabled the acquisition of dense point clouds. Major benefits of this technique are its high level of automation during data capturing and its spatial resolution. With point densities of up to several points per square meter laser scanning data have become a valuable additional data source for the reconstruction of different urban objects such as building, tree and roads.

There have been several attempts to extract roads from LIDAR data. Alharthy and Bethel used both the intensity and height information were used to filter the raw LIDAR data and remove “noise” that was unrelated to the road class (Alharthy, Bethel, 2003). Clode Perform road classification in a manner similar to Alharthy and Bethel (Clode, 2004). In their work, the road extraction process will be split into two stages, classification and vectorisation. Classification is accomplished by applying a hierarchical method yielding a binary image of ground elements (hence called “pixels”) classified as belonging to a road. Vectorisation of roads is then performed by convolving this binary image with a Phase-Coded-Disk (PCD) in order to extract the road centerline and to determine the road width.

Compared to the relatively high number of research groups focusing their work on road extraction from images, only a few groups work on the automatic extraction of roads in urban environments from LIDAR data. Usually, for automatic road recognition or reconstruction from LIDAR, we need to classify roads from LIDAR data to get the extraction results. In order to efficiency classify LIDAR data, other information, such as laser intensity, in addition to the height of cloud points is required. However, few algorithms for classification using intensity data have been deeply investigated (Mao, 2008).

The ultimate goal of most traditional methods in road extraction is to achieve the best possible classification performance for recognition of different objects such as buildings, roads and trees. This objective traditionally led to the development of different classification schemes for the recognition problem to be solved. The results of an experimental assessment of the different designs would then be the basis for choosing one of the classifiers as a final solution to the problem. (Kuncheva, 2004; Biggio and Roli, 2008)

It has been observed in recognition of roads from LIDAR data, that although one of the designs would yield the best performance, the sets of patterns misclassified by the different classifiers would not necessarily overlap. This suggested that different classifier designs potentially offered complementary information about the roads which could be harnessed to improve the performance of the selected classifier. These observations motivated the relatively recent interest in combining classifiers.

The idea is not to rely on a single decision making scheme. Instead, all the designs, or their subset, are used for decision making by combining their individual opinions to derive a consensus decision. Various classifier combination schemes have been devised and it has been experimentally demonstrated that some of them consistently outperform a single best classifier. However, there is presently inadequate understanding why some combination schemes are better than others and in what circumstances. In this paper, we present a road extraction method from both intensity and height information from LIDAR data, based on optimum classifier selection and fusion of them.

2. MULTIPLE CLASSIFIER SYSTEMS (MCS)

Combining classifiers is an established research area shared between statistical pattern recognition and machine learning. It is variously known as committees of learners, mixtures of experts, classifier ensembles, multiple classifier systems, consensus theory, etc. In such systems the optimal set of classifiers is first selected and then combined by a specific fusion method. If we have many different classifiers, it is sensible to consider using them in a combination in the hope of increasing the overall accuracy.

Methods that used for combination of classifiers depending on output type of single classifier. Hard classifier is a classifier only outputs a unique class, and soft classifier is a classifier that associates for each class a confidence measurement and at the end produced a vector for every classifier and a matrix for ensemble of classifier. Hard fusion methods are methods only use hard classifiers and soft fusion methods use soft classifiers.

Majority Voting (MV), Naïve Bays (NB) are two popular hard methods that fused hard classifiers.

One of the simplest combiners operating on binary classification outputs (correct/incorrect) is the majority voting (MV). Due to its simplicity, MV can be applied to classifiers producing different types of outputs as they all can be converted to the uniform binary outputs: 1/0 (correct/incorrect). Applications of MV for pattern recognition have already been studied in detail in (Lam, Suen, 1997). Lam and Suen studied MV performance for both odd and even number of independent classifiers supported by conditions of beneficial addition of one and two classifiers to the MCS.

MV considers only the most likely class provided by each classifier and chooses the most frequent class label among this crisp output set. In order to alleviate the problem of ties, the number of classifiers used for voting is usually odd (Kuncheva, 2004). A trainable variant of majority voting is weighted majority voting, which multiplies each vote by a weight before the actual voting. The weight for each classifier can be obtained; e.g., by estimating the classifiers' accuracies on a validation set. Assume that the label outputs of the L classifiers are given as c-dimensional binary vectors $[d_{i,1},...,d_{i,c}]^T \in \{0,1\}, i=1,...,L$ where $d_{i,j}=1$ if D_i Labels x in w_j, and *0* otherwise.

$$\sum_{i=1}^{L} d_{i,k} = \max_{j=1}^{c} \sum_{i=1}^{L} d_{i,j} \tag{1}$$

Where, D_i is *i*th classifier and w_j is *j*th class.

If the classifiers in the ensemble are not of identical accuracy, then it is reasonable to attempt to give the more importance to better classifiers in making the final decision. The label outputs can be represented as degrees of support for the classes in the following way:

$$d_{i,j} = \begin{cases} 1 & \text{if } \mathrm{D_i} \textit{ labels } \mathrm{x} \text{ in } \mathrm{w_j} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The discriminant function for class w_j obtained through weighted voting is:

$$g_i(x) = \sum_{i=1}^{L} b_i d_{i,j} \tag{3}$$

Where, b_i is the Weight of D_i classifier.

2.1 Optimum Classifier Selection (OCS)

The performance of MCS essentially depends on the complementarity (diversity) of constituent classifiers (Ship and Kuncheva, 2002). The complementarity can be obtained by diversifying the feature representation, classifier structure, and the training data of constituent classifiers. A strategy is to overly generate a large ensemble of candidate classifiers and then select a subset for good complementarity (Giacinto and Roli, 2001).

Optimum classifiers selection (OCS) from large ensemble is significant in two respects. First, the limited computing source of MCS demands that a small number of classifiers are to be combined. Second, in many cases, the combination of a subset of classifiers may give higher accuracy than combining all the classifiers at hand. In pattern recognition applications, we can easily generate a large number of classifiers by varying pre-processing, feature extraction, learning or classification algorithms, and training data. Running and combining all these classifiers is obviously not a good choice, and the selection of a subset should give better performance-to-cost ratio (Sharkey, 2000).

Given a set of candidate classifiers, a validation dataset and an appropriate selection criterion, the task of OCS is reduced to searching the space of classifier subsets to find a subset that give optimal criterion on the validation dataset. The validation data, the selection criterion, and the search algorithm are all influential to the combination performance of MCS. For selection from large ensemble, efficient search algorithms are needed to overcome the combinatorial explosion of search space. Recently, a few works have contributed to MCS design using classifier selection. (Giacinto, Roli, 2001) clustered the candidate classifiers according to interdependency and selected one classifier from each cluster. Roli, et al., also used heuristic search for classifier selection (Hao and Liu, 2003). In regard of the selection criterion, a number of classifier diversity measures have been studied (Ship, Kuncheva, 2002).

3. PROPOSED METHOD FOR DETERMINATION OF OPTIMUM CLASSIFIERS SELECTION

The goal of proposed classifier selection strategy is to select a subset of k classifiers from a given set of K (K>k) candidates, to achieve the best combination performance for classification of LIDAR data in complex urban area for extraction of road objects. Given a selection criterion (herein the classification accuracy of combination on validation data), classifier selection is reduced to a combinatorial search problem. Figure 1., shows the general structure of proposed methodology for optimum subset selection of classifiers for constructing a multiple classifier system.

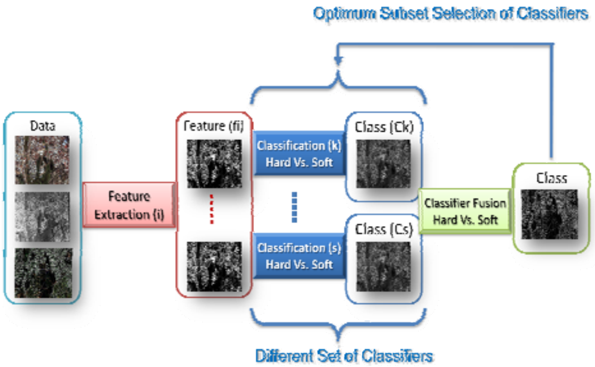

Figure 1. Flowchart of proposed method

In our proposed GA based optimization method, a selected subset of classifiers is represented by a binary string called a chromosome, with a *1/0* at position *i* denoting the presence/absence of classifier *i*. A number of chromosomes, called a population, evolve from generation to generation by selection, crossover, and mutation, with hope that the criterion measures (called fitness functions) of the chromosomes improve. The selection of chromosomes to survive to the next generation is based on the fitness functions such that the chromosomes with higher fitness have more chance to survive. The crossover and mutation operations enlarge the variation of population so as to increase the chance of escaping from local optima. After a number of generations, the chromosome of highest fitness in the population gives the solution of classifier selection (Hao, Liuo, 2003).

The use of genetic algorithms for MCS optimization requires the determination of five fundamental issues, namely: 1- Chromosome Representation, 2- Objective function, 3- Natural selection, 4- recombination and, 5- Termination criteria.

- **Chromosome Representation:** Using a bit string encoding scheme for chromosome string, the validity of MCS is encoded as shown in Figure 2. The aim of coding is to create a representation of existence (value 1) or extinction (value 0) of each one of MCS's classifiers.

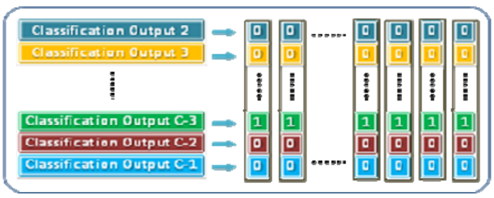

Figure 2. Chromosome in proposed GA based method for determination of optimum classifiers

- **Objective Function:** The accuracy of fusion of any combinations of classifiers can be summarized by contemplating the correctness of the detected object as defined in (Heipke, 1997).

$$correctness = \frac{TP}{TP + FP} \tag{4}$$

Where TP denotes the number of true positives, which is the number of pixels found in both the reference and detected data sets. FP is the number of false positives, which is the number of pixels that were detected but did not exist in the reference data set. Correctness is the ratio of the number of relevant pixels extracted to the total number of relevant pixels and detected irrelevant pixels retrieved.

- **Natural Selection:** In our approach, we used the roulette wheel selection method which has been developed by Holland, This is done by assigning each string a wedge on a roulette wheel whose size is proportional to the string's fitness. In this way a 'fit' string is more likely to be chosen than an 'unfit' string.

- **Recombination:** Genetic operators in recombination are two basic types of operators: crossover and mutation. Crossover: Depending on a predefined probability value ($Prc;\ 0 \leq Prc \leq 1.0$), the MCS parameter values of the parental individuals will be combined through a uniform crossover algorithm. Mutation: In our method, a one-point mutation algorithm is implemented. The mutation will be carried out depending on the mutation probability ($Prm;\ 0 \leq Prm \leq 1.0$). To mutate the value of a MCS parameter, the value of a randomly chosen position of the binary-coded MCS parameter is changed. This means that if the value at this position is 0 (i.e. lack of the term), it will be changed into 1 (signifying the presence of the term) and vice versa.

- **Termination Criteria:** In our method, if the mean or standard deviation of the population's cost (RMSE) reaches a certain level, the optimization process is terminated.

4. EXPERIMENT AND REULTS

To assess the capabilities of the proposed MCS method a sample LIDAR data of an urban area of city of Castrop-Rauxel which is located in the west of Germany, was selected (Figure 3). The selected area was suitable for the evaluation of the proposed classification strategy because the required complexities (e.g. proximities of different objects: building, tree and road) were available in the image. The pixel size of the range images is one meter per pixel.

Figure 3. Digital image area

This reflects the average density of the irregularly recorded 3D points which is fairly close to one per m^2. Intensity images for the first and last pulse data have been also recorded and the intention was to use them too in the experimental investigations. Figure 4 shows first- and last- pulse range and intensity images from the Ickern area.

4.1 Results of single classifiers

The classification process initiated with feature extraction operation.Twelve features have proved to be suitable to distinguish the main object classes buildings, vegetation (trees, bushes etc.), roads and terrain objects. Table 1 shows these features. Here N is the number of grey levels and P is the normalized symmetric co-occurrence matrix of dimension N x N. V is the normalized grey level difference vector of dimension N.

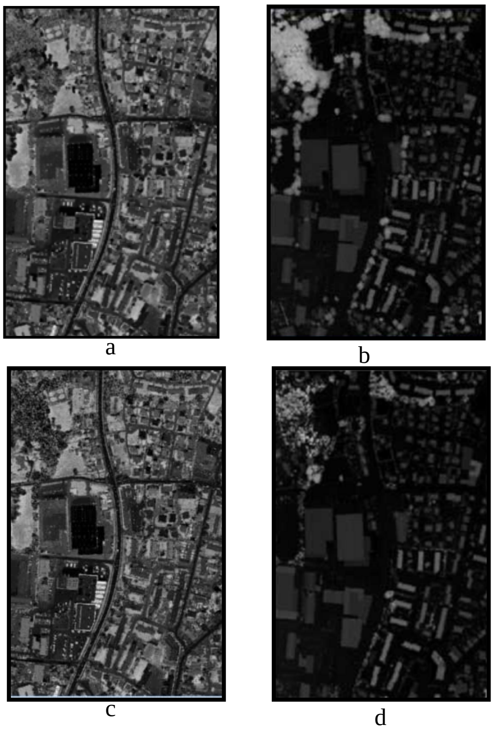

Figure 4. Data set, a) first pulse intensity, b) first pulse range, c) last pulse intensity, d) last pulse range

By generation feature space for each data layer of LIDAR data, the classification of these information computed based on two methods of Maximum Likelihood and Minimum Distance.

Maximum likelihood classification assumes that the statistics for each class in each band are normally distributed and calculates the probability that a given pixel belongs to a specific class. Unless a probability threshold is selected, all pixels are classified. Each pixel is assigned to the class that has the highest probability (i.e., the "maximum likelihood"). The Minimum Distance classification uses the mean vectors of each ROI and calculates the Euclidean distance from each unknown pixel to the mean vector for each class. All pixels are classified to the closest ROI class unless the user specifies standard deviation or distance thresholds, in which case some pixels may be unclassified if they do not meet the selected criteria.

Figure 5 shows the outputs of single classifiers. In this figure, “max” and “min” denote Maximum Likelihood and Minimum Distance classifiers. In this figure “F”, “L”, “I” and “R” denotes respectively First Pulse, Last Pulse, Intensity and Range. For each classifier that has been shown in this figure, correctness measures have been computed and these results have been shown for asphalt road class in Table 2.

Results of single classifiers show that Minimum Distance on the first pulse intensity has best correctness measure with 79.67 and Maximum Likelihood on the first range produce worse result. Two classification methods produce weak result on the range of LIDAR data than intensity like 5.83 for range and 51.22 for intensity.

4.2 Results of classifier fusion and selection

In this step, best combinations of 2, 3, 4, 5, 6 and 7 classifiers are selected and fused by Majority Voting (MV). In the end of this experiment, best combination of classifiers has selected by Genetic Algorithm. Figure 6(h) shows result of fusion without selection (by all of 8 classifiers) and figure 6(g) shows result of fusion of selected classifiers by GA. Comparison between result of table 2, 3 shows that multiple classifier system improved correctness measures for road class rather than single classifiers. Best result for correctness produced by fusion of classifiers that selected by GA. 87.37 is the best measure which produced by fusion and selection and improved best result of single classifiers to 8%.

Table1: Different features that used for classification of ALS data set.

Name of Features	Equation	Name of Features	Equation
Homogeneity	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i+j)^2}$	Standard Deviation	$\text{var}\,iance\_i = \sum_{i,j=0}^{N-1} P(i,j)\times(i-Mean\_i)^2$
Contrast	$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$	Correlation	$\sum_{i,j=0}^{N-1}(P(i,j)\times(i-Mean\_i)\times (j-Mean\_j))/\sqrt{(\text{var}\_i\times\text{var}\_j)}$
Dissimilarity	$\sum_{i,j=0}^{N-1} P_{i,j}\,\lvert i-j\rvert$	GLDV Angular Second Moment	$\sum_{i,j=0}^{N-1} V(k)^2$
Mean	$Mean\_i = \sum_{i,j=0}^{N-1} i\times P(i,j)$	GLDV Mean	$\sum_{i,j=0}^{N-1}(V(k)\times K)$
Angular Second Moment	$\sum_{i,j=0}^{N-1} P(i,j)^2$	GLDV Contrast	$\sum_{i,j=0}^{N-1}(V(k)\times K^2)$
Entropy	$\sum_{i,j=0}^{N-1} P_{i,j}\times(-\ln P_{i,j})$	GLDV Entropy	$\sum_{i,j=0}^{N-1}(-V(k)\times\ln(V(k)))$

Table 2. correctness measures for single classifiers

classifiers	FIc_1	FRc_1	LIc_1	LRc_1	FIc_2	FRc_2	LIc_2	LRc_2
Asphalt road	34.96	5.83	51.22	8.13	**79.67**	61.79	78.86	58.54

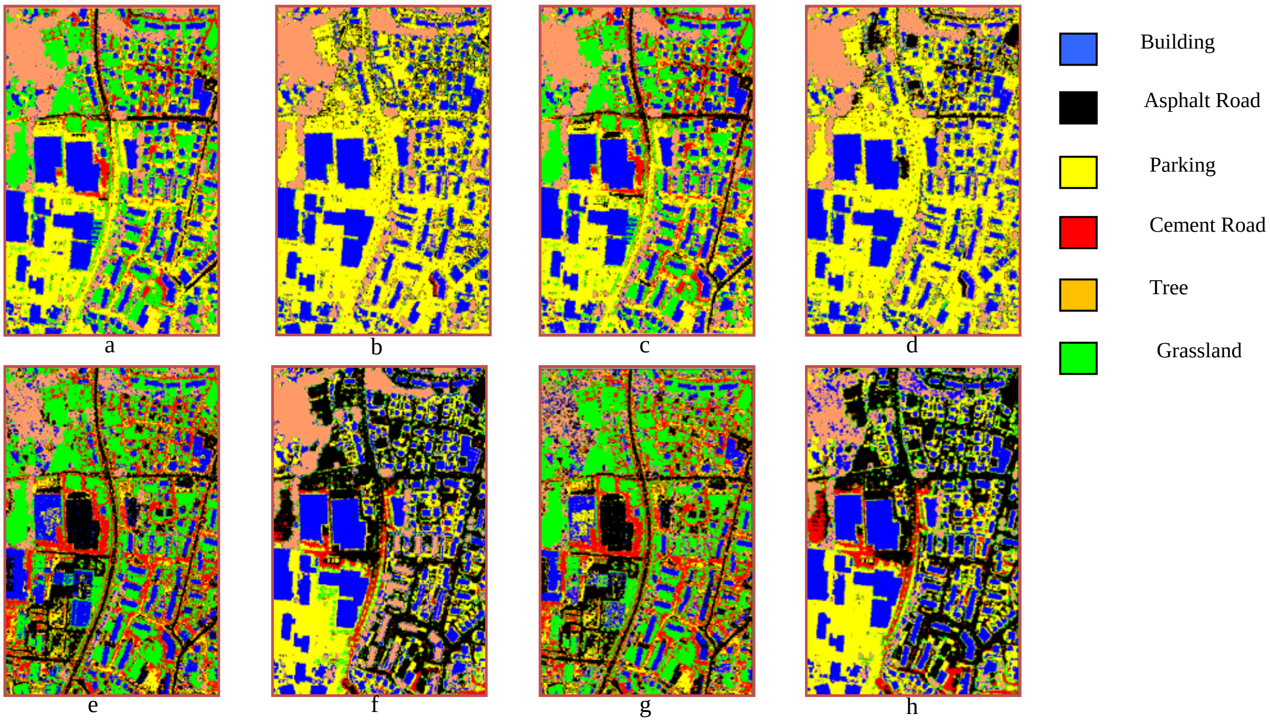

Figure 5. Result of classification, a) FImax , b) FRmax, c) LImax, d) LRmax, e) FImin, f) FRmin, g) LImin, h) LRmin

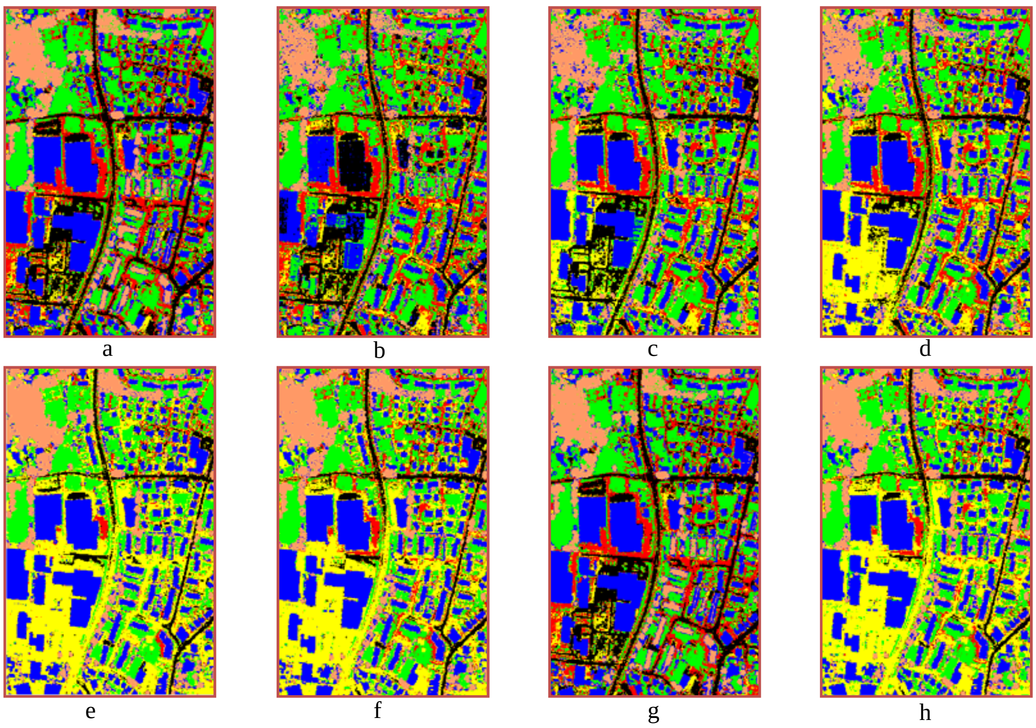

Figure 6. Result of fusion with N best combination of classifiers, a) 2 classifiers, b) 3 classifiers ,c) 4 classifiers ,d) 5 classifiers ,e) 6classifiers , f) 7classifiers , g) Best result from GA for Asphalt road class, h) result of fusion without selection and by all 8 classifiers

Table3 .Result of combination with different best selected classifiers

Combinations	***Optimum Subset of Classifiers***	***Correctness***
Combination of 2 of n	FIc_2, FRc_2	87.37
Combination of 3 of n	FIc_2, FRc_2, LIc_2	80.15
Combination of 4 of n	$LIc_1, FIc_2, LIc_2, FRc_2$	79.80
Combination of 5 of n	$LRc_1, FIc_2, FRc_2, LIc_2, LRc_2$	69.11
Combination of 6 of n	$FRc_1, FRc_2, LRc_1, LRc_2, FIc_2, LIc_2$	55.28
Combination of 7 of n	$FRc_1, LIc_1, LRc_1, FIc_2, LIc_2, FRc_2, LRc_2$	52.84
Combination of 8 classifier without selection	$FIc_1, FRc_1, LIc_1, LRc_1, FIc_2, LIc_2, FRc_2, LRc_2$	50.404
Best combination by GA	FIc_2, FRc_2	**87.37**

5. CONCLUSION

In this paper an optimum multiple classifier system presentd for classification and extraction of road objects by involving both height and intensity information of LIDAR data. Optimum subset of classifiers determined through an evolutionary strategy. The obtained results showed that there is not any guaranty to improving the accuracy of classification by any combination of classifiers in a multiple classifier system. Although we got promising results from proposed optimum classifier fusion method for classification of objects and extraction of roads in different urban areas, but there are still available open questions in this research work. These questions are around the optimum determination of features, type of classification techniques and the potential of other methodologies in classifier fusion.

6. REFERENCES

Wiedemann, C., Ebner, H., 2000. Automatic completion and eval- uation of road networks. International Archives of Photogrammetry
and Remote Sensing, vol. 33, Part B3/2, pp. 979–986.

R.Gecen, G.Sarp,2008. ROAD DETECTION FROM HIGH AND LOW RESOLUTION SATELLITE IMAGES. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing, China.

Grote, A., Butenuth, M., Heipke, C., 2007. Road extraction in suburban areas based on normalized cuts. In: International Archives of Photogrammetry and Remote Sensing, Vol. 36, Part 3/W49A, pp. 51-56.

Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H., 1999. Automatic road extraction based on multi-scale, grouping, and context. Photogrammetric Engineering and Remote Sensing 65 (7), 777– 785.

Hinz, S., 2004. Automatic road extraction in urban scenes – and beyond. In: International Archives of Photogrammetry and Remote Sensing, Vol. 35, Part B3, pp. 349-355.

Alharthy, A. Bethel, J., 2003. Automated Road Extraction From LIDAR Data. In: Proceedings of ASPRS, Anchorage, Alaska, unpaginated, CD-ROM.

Mao, J. H, Zeng, Q. H, Lai, J. Z, Liu, X. F. Filtering lidar points by fusion of intensity measures and aerial images. The international archive of the photogrammetry, remote sensing and spatial information science, Vol.XXXVII.part B3b, Beijing, China, 2008.

Clode, S., Kootsookos. P, Rottensteiner.F, 2004a. The Automatic Extraction of Roads from LIDAR Data. In IAPRSIS, , Vol. XXXV-B3, pp. 231 – 236.

Kncheva, L., 2004. Combining Pattern Classifiers methods and algorithms. A john Wiley&sons, INC.publication, Hoboken, New jersey.canada.

Biggio, B., Roli, F., Fumera, G. 2008. Adversarial Pattern Classifier Using Multiple Classifier System and Randomisation. 12th Joint IAPR International Workshop on Stractural and Syntatic Pattern Recognition(SSPR2008), Orlando, USA.

Lam L, Suen CY. Application of majority voting to pattern recognition: an analysis of its behaviour and performance. IEEE Transactions on Systems, Man, and Cybernetics 1997; 27(5): 553-568.

Giacinto, G. Roli F, 2001. Design of effective neural network ensembles for image classification purposes. Image Vision and Computing Journal, 19(9-10):669–707.

Sharkey, A.J.C, Sharkey, N.E, 1997. Combining diverse neural nets. The Knowledge Engineering Review, 12(3):231–247.

Shipp, C.A, Kuncheva, L, 2002. Relationship between combination methods and measures of diversity in combining classifiers. Information Fusion, 3(2):135–148.

Hao, H. Liu, C. Sako, H, 2003. Comparison of Genetic Algorithm and sequential search methods for classifier subset selection.Proceeding of the seventh International Conference on document Analysis and Recognition (ICDAR).

Heipke, C., Mayr, H., Wiedemann, C., & Jame, 1997. Evaluation of automatic road Extraction. In International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, volume XXXII-2-3W3, pp. 47–56.