CLUSTERING OF LIDAR DATA USING PARTICLE SWARM OPTIMIZATION ALGORITHM IN URBAN AREA

F. Samadzadegan, S. Saeedi

Dept. of Geomatics, Faculty of Engineering, University of Tehran, Tehran, Iran, (samadz, ssaeedy)@ut.ac.ir

KEY WORDS: Clustering, LIDAR, Particle swarm optimization, Urban Area, Object Extraction

ABSTRACT:

One of the fundamental steps in the transformation of the LIDAR data into the meaningful objects in urban area involves their segmentation into consistent units through a clustering process. Nevertheless, due to the scene complexity and the variety of objects in urban area, e.g. buildings, roads, and trees, it is clear that a clustering using only a single cue will not suffice. Considering the availability of additional data sources, like laser range and intensity information in both first and last echo, more information can be integrated in the clustering process and ultimately into the recognition and reconstruction scheme. Multi dimensionality nature of LIDAR data with a dense sampling interval in urban area generates a huge amount of information. This amount of information has produced a lot of problems for finding global optima in most of traditional clustering techniques. This paper describes the potential of a Particle Swarm Optimization (PSO) algorithm to find global solutions to the clustering problem of multi dimensional LIDAR data in urban area. It is a kind of swarm intelligence that is based on social-psychological principles and provides insights into social behaviour, as well as contributing to engineering applications. By integrating the simplicity of the *k*-means algorithm with the capability of the PSO algorithm, this paper presents a robust and efficient clustering method which can overcome the problem of trapping to local optima of k-means technique. This algorithm successfully applied to clustering of several LIDAR data sets in different urban area with different size and complexities. The experimental results demonstrate that PSO based clustering technique produces much better outputs in terms of both accuracy and computation time than other traditional clustering techniques.

1. INTRODUCTION

LIDAR has become a standard data collection technology for capturing object surface information in 3D modelling of urban area. Although, these systems provide detailed valuable geometric information, but still require interpretation of the data for object extraction and recognition to make it practical for mapping purposes. During last decade, several numbers of clustering and filtering techniques have been developed to recognition of the objects for 3D city modelling or removing the "artefacts" of bare terrain (Maas, 2001; Tao and Hu, 2001; Filin, 2002; Rottensteiner and Briese, 2002; Sithole and Vosselman, 2003; Samadzadegan, 2004; Vosselman et al., 2004; Lodha et al., 2007; Sampath and Shan, 2008).

In order to improve the performance of 3D object extraction process, additional data has been considered. First, LIDAR systems register two echoes of the laser beam, the first and the last pulse, generally corresponding to the highest and the lowest object point hit by the laser beam. First and last pulse data will especially differ in the presence of vegetation (Kraus, 2002). Moreover, LIDAR systems record the intensity of the returned laser beam which is mainly in the infrared part of the electromagnetic spectrum. Nevertheless, the intensity image is under-sampled and, thus, very noisy. In addition, an extra powerful source of information is visible images. Digital images can provide necessary information due to their intensity and spectral content as well as their high spatial resolution which is still better than the resolution of laser scanner data.

Therefore, in the context of 3D object extraction in urban area, various type of information has been fused to overcome the difficulties of classification and identification of complicated objects (Lim and Suter, 2007; Morgan and Habib, 2002; Vosselman et al., 2004). Collecting this information, extremely enlarge the size of data sets and proportionally the dimension of feature spaces in clustering process. As a result, most of traditional clustering techniques that have been applied properly in other applications with standard data and feature space dimension are not efficient enough for object extraction process from LIDAR data (Lodha et al., 2007).

k-means is one of the most popular clustering algorithms, which could be consider for handling the massive datasets. The algorithm is efficient at clustering large data sets because its computational complexity only grows linearly with the number of data points (Kotsiantis and Pintelas, 2004). However, the algorithm may converge to solutions that are not optimal. In this paper, an alternative clustering method to solve the local optimum problem of the *k*-means algorithm is presented. The applied method adopts the particle swarm optimization algorithm as it has proved to give a more robust performance than other intelligent optimisation methods for a range of complex problems.

2. BASIC CONCEPTS IN DATA CLUSTERING

Historically the notion of finding useful patterns in data has been given a variety of names including data clustering, data mining, knowledge discovery, pattern recognition, information extraction, etc (Ajith et al., 2006). Data clustering is an analytic process designed to explore data in discovering of consistent patterns and/or systematic relationships between variables, and then to validate the findings by applying the detected patterns to new subsets of data.

Generally, clustering algorithms can be categorized into partitioning methods, hierarchical methods, density-based methods, grid-based methods, and model-based methods. An

excellent survey of clustering techniques can be found in (Kotsiantis and Pintelas, 2004). In this paper, the focus will be on the partitional clustering algorithms because they are more popular than other clustering algorithms in clustering of remotely sensed data (Harvey et al., 2002; Omran et al., 2005; Wilson et al., 2002). Partitional clustering algorithms divide the data set into a specified number of clusters and then evaluate them by some criterion. These algorithms try to minimize certain criteria (e.g. a square error function) and can therefore be treated as optimization problems.

The most widely used partitional algorithm is the iterative k-means approach (Kotsiantis and Pintelas, 2004). The k-means algorithm starts with k centroids (initial values are randomly selected or derived from a priori information). Then, each pattern in the data set is assigned to the closest cluster (i.e. closest centroid). Finally, the centroids are recalculated according to the associated patterns. This procedure is repeated until convergence is achieved.

It is known that the *k*-means algorithm may become intent at local optimal solutions, depending on the choice of the initial cluster centres. Genetic algorithms have a potentially greater ability to avoid local optima than is possible with the localised search employed by most clustering techniques. Maulik and Bandyopadhyay (2004) proposed a genetic algorithm-based clustering technique, called GA-clustering, that has proven effective in providing optimal clusters. With this algorithm, solutions (typically, cluster centroids) are represented by bit strings.

The search for an appropriate solution begins with a population, or collection, of initial solutions. Members of the current population are used to create the next generation population by applying operations such as random mutation and crossover. At each step, the solutions in the current population are evaluated relative to some measure of fitness, with the fittest solutions selected probabilistically as seeds for producing the next generation. The process performs a generate-and-test beam search of the solution space, in which variants of the best current solutions are most likely to be considered next.

3. CLUSTERING OF LIDAR DATA USING PSO ALGORITHM

Swarm Intelligence (SI) is an innovative distributed intelligent paradigm for solving optimization problems that originally took its inspiration from the biological examples by swarming, flocking and herding phenomena in vertebrates. These techniques incorporate swarming behaviours observed in flocks of birds, schools of fish, or swarms of bees, and even human social behaviour, from which the idea is emerged (Omran et al., 2002, 2005; Paterlini and Krink, 2005; Pham et al., 2006; Wu and Shi, 2001). Data clustering and Swarm intelligence may seem that they do not have many properties in common. However, recent studies suggest that they can be used together for several real world data clustering and mining problems especially when other methods would be too expensive or difficult to implement.

Clustering approaches inspired by the collective behaviours of ants have been proposed by Wu and Shi (2001), Labroche et al. (2001). The main idea of these approaches is that artificial ants are used to pick up items and drop them near similar items resulting in the formation of clusters. Omran et al. (2002) proposed the first PSO-based clustering algorithm. The results of Omran et al. (2002, 2005) show that PSO outperformed k-means, FCM and other state-of-the-art clustering algorithms.

More recently, Paterlini and Krink (2005) compared the performance of *k*-means, GAs, PSO and Differential Evolution (DE) for a representative point evaluation approach to partitional clustering. The results show that GAs, PSO and DE outperformed the *k*-means algorithm. Pham et al. (2006) used the bee algorithm for clustering of different data sets. The obtained results of their work show that their proposed bee algorithm has better performance than both of standard *k*-mean as well as GA-based method. In general, the literature review of recent techniques in the background of clustering shows that the swarm-based clustering algorithm performs better than the *k*-means algorithm.

Regarding to the recent researches in the context of clustering of massive LIDAR data and unique potential of PSO algorithm in solving complex optimization, a clustering algorithm based on Particle Swarm Optimization (PSO) for outperforming k-mean method is developed in this paper.

3.1 Particle Swarm Optimization

The PSO method is one of optimization methods developed for searching global optima of a nonlinear function (Kennedy and Eberhart, 1995; Kennedy et al., 2002). It is inspired by the social behaviour of birds and fish. The method uses group of problem solutions. Each solution consists of set of parameters and represents a point in multidimensional space. Each potential solution is called particle and the group of particles (population) is called swarm. Particles move through the search domain with a specified velocity in search of optimal solution. Each particle maintains a memory which helps it in keeping the track of its previous best position. The positions of the particles are distinguished as personal best and global best. PSO has been applied to solve a variety of optimization problems and its performance is compared with other popular stochastic search techniques like Genetic algorithms, Differential Evolution, Simulated Annealing etc. (Kennedy et al., 2002; Omran and Engelbrecht, 2005).

The working of the PSO may be described as: For a D-dimensional search space the position of the *i*th particle is represented as $X_i = (x_{i1}, x_{i2}, ... x_{iD})$. Each particle maintains a memory of its previous best position $P_{best} = (p_{i1}, p_{i2}... p_{iD})$. The best one among all the particles in the population is represented as $P_{gbest} = (p_{g1}, p_{g2}... p_{gD})$. The velocity of each particle is represented as $V_i = (v_{i1}, v_{i2}, ... v_{iD})$. In each iteration, the *P* vector of the particle with best fitness in the local neighbourhood, designated *g*, and the *P* vector of the current particle are combined to adjust the velocity along each dimension and a new position of the particle is determined using that velocity. The two basic equations which govern the working of PSO are that of velocity vector and position vector given by:

$$v_{id}(t+1) = wv_{id}(t) + c_1 r_{1d}(t)\big(p_{id}(t) - x_{id}(t)\big) + $$
$$c_2 r_{2d}(t)\big(p_{gd}(t) - x_{id}(t)\big) \qquad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \qquad (2)$$

The first part of equation (1) represents the inertia of the previous velocity, the second part is the cognition part and it tells us about the personal experience of the particle, the third part represents the cooperation among particles and is therefore named as the social component. Acceleration constants c_1, c_2 and inertia weight w are the predefined by the user and r_1, r_2 are the uniformly generated random numbers in the range of [0, 1].

3.2 PSO Clustering

For the purpose of this paper, define the following symbols: N_d denotes the input dimension, i.e. the number of parameters of each data vector; N_o denotes the number of data vectors to be clustered; N_c denotes the number of cluster centroids (as provided by the user), i.e. the number of clusters to be formed; z_p denotes the p-th data vector; m_j denotes the centroid vector of cluster j; n_j is the number of data vectors in cluster j; C_j is the subset of data vectors that form cluster j.

In the context of clustering, a single particle represents the N_c cluster centroid vectors (Merwe, and Engelbrecht, 2003; Esmin et al., 2008). That is, each particle X_i is constructed as follows:

$$X_i = \left(m_{i1}, \dots, m_{ij}, \dots, m_{iN_c} \right)$$

where m_{ij} refers to the j-th cluster centroid vector of the i-th particle in cluster C_{ij}. Therefore, a swarm represents a number of candidate clusters for the current data vectors. The fitness of particles is easily measured as the quantization error (Merwe, and Engelbrecht, 2003).

$$J_e = \frac{\sum_{j=1}^{N_c}\left[\sum_{\forall z_p \in C_{ij}} d(z_p, m_j) / |C_{ij}| \right]}{N_c} \quad (3)$$

where d is defined in equation (1), $|C_{ij}|$ is the number of data vectors belonging to cluster C_{ij}.

Using the standard *gbest* PSO, data vectors can be clustered as follows:

1. Initialize each particle to contain N_c randomly selected cluster centroids.
2. For $t = 1$ to t_{max} do
   i. For each particle i do
   ii. For each data vector z_p
      a. calculate the Euclidean distance $d(z_p, m_{ij})$ to all cluster centroids C_{ij}
      b. assign z_p to cluster C_{ij} such that distance $d(z_p, m_{ij}) = \min_{\forall c=1,\dots N_c} \{ d(z_p, m_{ic}) \}$
      c. calculate the fitness using equation (3)
   iii. Update the global best and local best positions
   iv. Update the cluster centroids.

where t_{max} is the maximum number of iterations. The population-based search of the PSO algorithm reduces the effect that initial conditions has, as opposed to the K-means algorithm; the search starts from multiple positions in parallel.

However, the K-means algorithm tends to converge faster (after less function evaluations) than the PSO, but usually with a less accurate clustering (Omran et al., 2005). Merwe, and Engelbrecht, showed that the performance of the PSO clustering algorithm can further be improved by seeding the initial swarm with the result of the K-means algorithm (Merwe, and Engelbrecht, 2003). The hybrid algorithm first executes the K-means algorithm once. In this case the K-means clustering is terminated when (1) the maximum number of iterations is exceeded, or when (2) the average change in centroids vectors is less that 0.0001 (a user specified parameter). The result of the K-means algorithm is then used as one of the particles, while the rest of the swarm is initialized randomly. The gbest PSO algorithm as presented above is then executed.

However, sometimes a particle that does not represent a good solution is going to be evaluated as if it did. For instance, suppose that one of the particle clusters has a data vector that is very close to its centroid, and another cluster has a lot of data vectors that are not so close to the centroid. This is not a very good solution, but giving the same weight to the cluster with one data vector as the cluster with a lot of data vectors can make it seem to be. Furthermore, this equation is not going to reward the homogeneous solutions, that is, solutions where the data vectors are well distributed along the clusters (Esmin et al., 2008). To solve this problem Esmin et al., proposed the following equation, where the number of data vectors belonging to each cluster is taken into account on the calculation of the quality (Esmin et al., 2008):

$$F = \left\{ \sum_{j=1}^{N_c} \left[\left(\sum_{\forall z_p \in C_{ij}} d(z_p, m_{ij}) / |C_{ij}| \right) \times \left(|C_{ij}| / N_o \right) \right] \right\} \quad (4)$$

where *No* is the number of data vectors to be clustered.

To take into account the distribution of the data among the clusters, the equation can be changed to:

$$F' = F \times \left(|C_{ik}| - |C_{il}| + 1 \right) \quad (5)$$

Such that, $|C_{ik}| = \max_{\forall j=1,\dots,N_c}\{|C_{ij}|\}$ and $|C_{il}| = \min_{\forall j=1,\dots,N_c}\{|C_{ij}|\}$

4. EXPERIMENTAL INVESTIGATIONS

The airborne LIDAR data used in the experimental investigations have been recorded with TopScan's Airborne Laser Terrain Mapper system ALTM 1225 (TopScan, 2004). The data are recorded in area of Rheine in Germany. Two different patches with residential and industrial pattern were selected in the available dataset. The selected areas were suitable for the evaluation of the proposed classification strategy because the required complexities (e.g. proximities of different objects: building and tree) were available in the image (Figure 1. a, b). The pixel size of the range images is one meter per pixel. 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 (1) shows first- and last-pulse range images from the Rheine area. The impact of the trees in the first- and last- pulse images can be easily recognized by comparing the two images of this figure.

The first step in every clustering process is to extract the feature image bands. The features of theses feature bands should carry useful textural or surface related information to differentiate between regions related to the surface. Several features have been proposed for clustering of range data. Axelsson (1999) employs the second derivatives to find textural variations and Maas (1999) utilizes a feature vector including the original height data, the Laplace operator, maximum slope measures and others in order to classify the data. In the following experiments we restrict to five types of features:

- LIDAR range data
- The ratio between first and last pulse range images
- Top-Hat filtered last pulse range image
- Local height variation which is computed using a small window (3*3) around a data sample.
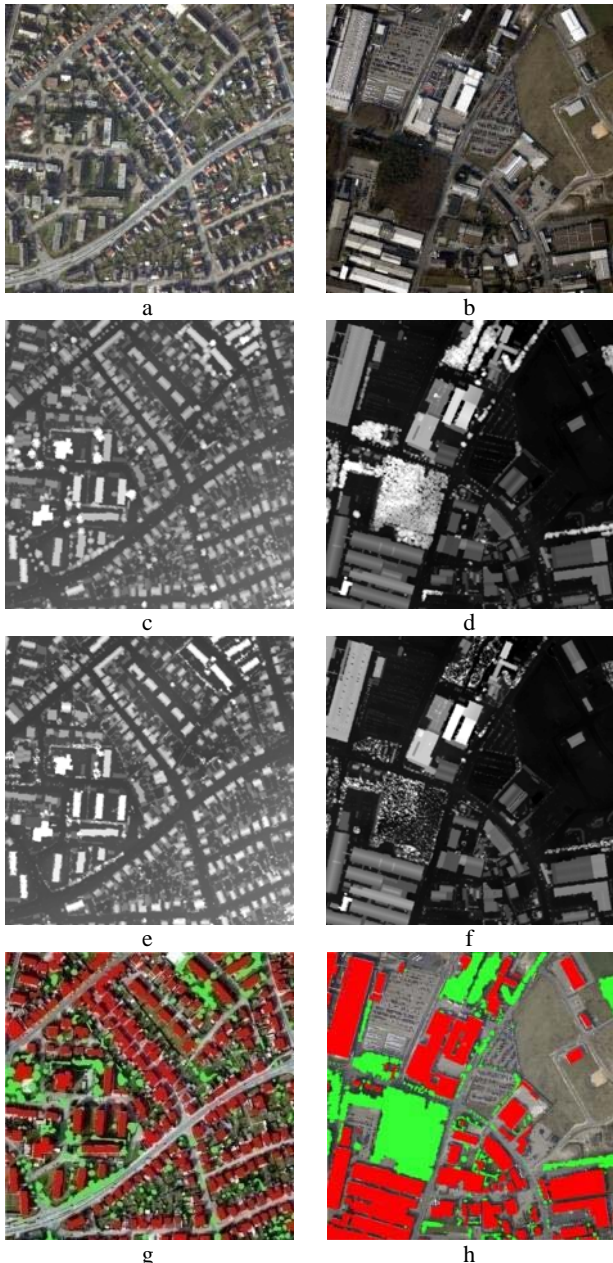- Last pulse Intensity

Figure 1. a) Aerial image of residential evaluation area. b) Aerial image industrial evaluation area. c) First LIDAR range data of residential area. d) First LIDAR range data of industrial area. e) Second LIDAR range data of residential area. f) Second LIDAR range data of industrial area. g) Overlaid manual digitization objects in residential area. h) Overlaid manual digitization objects in industrial Area

Evaluation of these two algorithms for clustering of the data sets into three clusters (ground, tree, and building) is depicted in Figure 2. Figures (2c) and (2d) show the *k*-means clustering results and Figures (2e) and (2f) show the PSO clustering algorithm results in two evaluation areas. Building class regions are highlighted in red and vegetation class regions in green colour in Figure 2. Visual inspections shows that vegetation class is directly associated with vegetation, in particular trees, bushes or forest and the B-class is mainly associated with building regions.
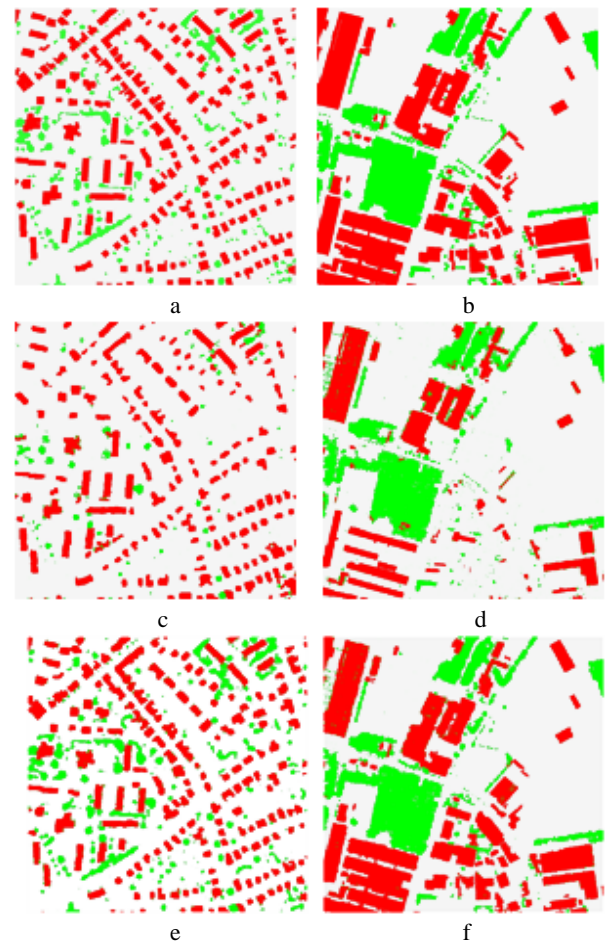

Figure 2. Clustering results. a) Manual Digitization objects in residential area. b) Manual Digitization objects in industrial area. c) Clustering results of k-means in residential area. d) Clustering results of *k*-means in industrial area. e) Clustering results of PSO algorithm in residential area. f) Clustering results of PSO algorithm in industrial area.

For both data sets, averages over 30 simulations are given. All algorithms are run for 300 function evaluations, and the PSO algorithms used 15 particles. For PSO, $w = 0.72$ and $c_1 = c_2 = 1.49$. These values were chosen to ensure good convergence (Omran et al., 2005).

4.1 Accuracy Assessment

Comparative studies on clustering algorithms are difficult due to lack of universally agreed upon quantitative performance evaluation measures. Many similar works in the clustering area use the classification error as the final quality measurement; so in this research, we adopt a similar approach (Zhong and Ghosh, 2003).

In this paper, confusion matrix used to evaluate the true labels and the labels returned by the clustering algorithms as the quality assessment measure. Given some ground truth the relation between the "true" classes and the classification result can be quantified. With the clusters the same principle can be applied. Mostly a much bigger number of clusters is then related to the given ground truth classes to examine the quality of the clustering algorithm. From the confusion matrix we calculate the *Kappa Coefficient* (Cohen, 1960). The accuracy measurements showing above, namely, the overall accuracy, producer's accuracy, and user's accuracy, though quite simple

337

to use, are based on either the principal diagonal, columns, or rows of the confusion matrix only, which does not use the information from the whole confusion matrix. A multivariate index called the Kappa coefficient (Tso and Mather, 2009) has found favour. The Kappa coefficient uses all of the information in the confusion matrix in order for the chance allocation of labels to be taken into consideration. The Kappa coefficient is defined by:

$$\hat{k} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \times x_{+i})} \qquad (6)$$

In this equation, $\hat{k}$ is the Kappa coefficient, r is the number of columns (and rows) in a confusion matrix, xii is entry (i, i) of the confusion matrix, xi+ and x+i are the marginal totals of row i and column j, respectively, and N is the total number of observations (Tso and Mather, 2009).

Table 1 shows the confusion matrix and Kappa coefficient of *k*-means and PSO algorithms clustering in residential dataset. The Error matrix and Kappa coefficient of k-means and PSO algorithms clustering in industrial dataset presented in Table 2.

Table 1. Error matrix and Kappa coefficient of *k*-means and PSO algorithms clustering applied to residential area.

		Reference Data			
		Building	Tree	Ground	Total
K-Means	Building	64338	1551	338	66227
	Tree	3561	58692	5930	68183
	Ground	54341	10509	290740	355590
	Total	122240	70752	297008	490000
Kappa coefficient = 0.6927					

		Reference Data			
		Building	Tree	Ground	Total
PSO algorithms	Building	114673	2473	6463	123609
	Tree	3022	61356	5193	69571
	Ground	3466	6803	286551	296820
	Total	121161	70632	298207	490000
Kappa coefficient = 0.898					

Table 2. Error matrix and Kappa coefficient of *k*-means and PSO algorithms clustering in industrial area.

		Reference Data			
		Building	Tree	Ground	Total
K-Means	Building	26878	2168	1108	30154
	Tree	187	3707	105	3999
	Ground	16443	12879	139025	168347
	Total	43508	18754	140238	202500
Kappa coefficient = 0.584					

		Reference Data			
		Building	Tree	Ground	Total
PSO algorithms	Building	38061	1316	2416	41793
	Tree	1108	13831	1763	16702
	Ground	4345	3616	136044	144005
	Total	43514	18763	140223	202500
Kappa coefficient = 0.842					

By comparing the counts in each class a striking difference to the PSO algorithm result has to be observed. For the two classes of major interest in this study, the Building class and Tree class, the differences are quite significant. Visual interpretation indicates that the Building class of *k*-means and include building areas but also regions related to roads which supports the smaller number of counts of the PSO method to be more precise (Figure 3). Similarly the higher number of counts for the Tree class indication (3D) vegetation regions (trees, bushes)

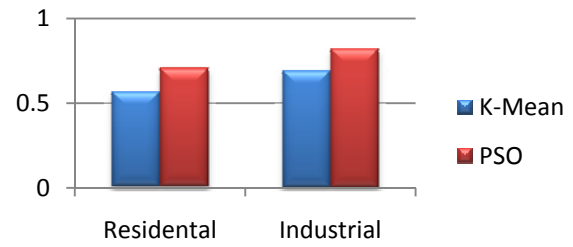obtained with the PSO algorithm method is supported by visual interpretation.

Figure 3. The value of Kappa coefficient, in two situations of clustering by kemean and PSO.

5. CONCLUSION

This paper presented the capability of a new clustering method based on the particle swarm optimization algorithm in object extraction form LIDAR data. The method employs the PSO algorithm to search for the set of cluster centres that minimizes a given clustering metric. One of the advantages of this method is that it does not become trapped at locally optimal solutions. This is due to the ability of the PSO algorithm to perform local and global search simultaneously. Experimental results for different LIDAR data sets have demonstrated that the PSO algorithm method produces better performances than those of the *k*-means algorithm.

One of the drawbacks of the artificial PSO algorithm is the number of tunable parameters it employs. A possible line of research, therefore, is to find ways to help the user choose appropriate parameters.

6. REFERENCES

Ajith, A., Crina G., Vitorino R., (Eds.) (2006). Swarm Intelligence in Data Mining, Studies in Computational Intelligence (series), Vol. 34, Springer-Verlag, ISBN: 3-540-34955-3, 267 p., Hardcover.

Axelsson, 1999 P. Axelsson (1999), Processing of laser scanner data-algorithms and applications, ISPRS Journal of Photogrammetry and RS 54 (2–3), pp. 138–147. Article | PDF (1269 K) | View Record in Scopus | Cited By in Scopus (113)

Camazine, S., Deneubourg, J., Franks, N.R., Sneyd, J., Theraula, G. and Bonabeau, E. (2003). Self-Organization in Biological Systems, (Princeton University Press, Princeton).

Cohen, J. 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20(1):37–46.

Esmin, A., Pereira, D., and Araujo, F.(2008). Study of Different Approach to Clustering Data by Using the Particle Swarm Optimization Algorithm, IEEE Congress on Evolutionary Computation (CEC 2008) .

Filin, S. 2002, "Surface Clustering from Airborne Laser Scanning Data." International Archives of Photogrammetry and Remote Sensing, Vol. 34/3 A, Graz, Austria, 2002.

Harvey N. R., J. Theiler, S. P. Brumby, S. Perkins, J. J. Szymanski, J. J. Bloch, R. B. porter, M. Galassi, A. C. Young, (2002). "Comparison of GENIE and conventional supervised classifiers for multispectral image feature extraction," IEEE Trans. on Geoscience and RS, vol. 40, no. 2, pp. 393-404.

Jain, A.K. , M.N. Murty, P.J. Flynn, (1999). Data Clustering: A Review, ACM Computing Surveys, Vol. 31, No. 3, Sep. 1999.

Kennedy, J., Eberhart, RC., (1995). Particle Swarm Optimization, Proceedings of the IEEE International Joint Conference on Neural Networks, Vol. 4, pp 1942–1948.

Kennedy, J., Eberhart, RC., Shi, Y., (2002). Swarm Intelligence, Morgan Kaufmann, 2002.

Kotsiantis S. and Pintelas P., (2004). Recent advances in clustering: a brief survey, WSEAS Transactions on Information Science and Applications 1:73-81.

Kraus, K., (2002). Principles of airborne laser scanning. Journal of the Swedish Society for Photogrammetry and Remote Sensing 2002:1 53–56.

Kuncheva L., (2004). Combining pattern classifiers, John Wiley & Sons.

Labroche N, Monmarche N, and Venturini G., (2002). Visual Clustering based on the Chemical Recognition System on Ants. In Proceedings of the European Conf. on AI, 2002.

Lim, E. H. and Suter, D., (2007). Conditional Random Field for 3D Point Clouds with Adaptive Data Reduction. In Proceedings of the 2007 international Conference on Cyberworlds (October 24 - 26, 2007). International Conference on Cyberworlds. IEEE Computer Society, Washington, DC, 404-408. DOI= http://dx.doi.org/10.1109/CW.2007.24

Lodha, S. K., Fitzpatrick, D. M., and Helmbold, D. P., (2007). Aerial Lidar Data Classification using AdaBoost. In Proceedings of the Sixth international Conference on 3-D Digital Imaging and Modeling (August 21 - 23, 2007). 3DIM. IEEE Computer Society, Washington, DC, 435-442. DOI= http://dx.doi.org/10.1109/3DIM.2007.10

Maas, H.-G., (1999). Close solutions for the determination of parametric house models from invariant moments of airborne laserscanner data, International Archives of Photogrammetry and Remote Sensing, 32(3-2W5): 193-199.

Maas, H.-G., (2001). The suitability of Airborne Laser Scanner Data for Automatic 3D Object Reconstruction, Third Int. Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona, Switzerland.

Maulik U. and Bandyopadhyay S., (2004). "Genetic Algorithm Based Clustering Technique", Pattern Recognition, vol. 33, no. 9, pp. 1455-1465, 2004.

Merwe, D., Engelbrecht, A., (2003). DataClustering using Particle SwarmOptimization. Congress on Evolutionary Computation, 2003. Proceedings of IEEE Congress on Evolutionary.

Morgan, M., Habib, A., (2002). Interpolation of LiDAR data and automatic building extraction. ASPRS Annual Conference, CD-ROM, Washington, DC, April 19–25.

Omran M, Salman A, and Engelbrecht AP., (2002). Image Classification using Particle Swarm Optimization. In Conf. on Simulated Evolution and Learning, vol. 1, pp 370–374, 2002.

Omran M., A. Engelbrecht and A. Salman, (2005). Particle Swarm Optimization Method for Image Clustering, Pattern Recognition and Artificial Intelligence Vol. 19, No. 3, pp 297–321.

Paterlini S and Krink T., (2006). Differential Evolution and PSO in Partitional Clustering. Computational Statistics and Data Analysis, 50(2006):1220–1247, 2005.

Pham, D.T., Ghanbarzadeh, A., Koç, E., Otri , S., Rahim , S. and Zaidi, M., (2006). The Bees Algorithm – A novel tool for complex optimisation problems. In: Proc. of the 2nd Virtual

International Conference on Intelligent Production Machines and Systems (I*PROMS-06), Cardiff, UK, 2006, 454-459.

Rottensteiner, F., Briese, Ch., (2002). "A new method for building extraction in urban areas from high-resolution LIDAR data", International Archives of Photogrammetry and Remote Sensing, Vol. 34, Part 3A, Graz, Austria

Samadzadegan, F., (2004) Object extraction and recognition from LIDAR data based on fuzzy reasoning and information fusion techniques, ISPRS- Turkey.

Sampath, A.,Shan, J., (2008). Building Roof Segmentation and Reconstruction from Lidar Point Clouds Using Clustering Techniques, International Archives of Photogrammetry and Remote Sensing, Vol. XXXVII. Part B3a, 279-284.

Sithole, G., Vosselman, G., (2003), Report: "ISPRS Comparison of Filters",Department of Geodesy, Faculty of Civil Engineering and Geosciences, Delft Univ. of Tech.

Tao, C., Hu, Y., (2001). A review of post-processing algorithms for airborne LiDAR data. CD-ROM Proceedings of ASPRS Annual Conference, April 23–27, St. Louis, USA.

Tso, B., and P. M. Mather. 2009. Classification Methods for Remotely Sensed Data. Taylor & Francis Group.

Vosselman, G., B. Gorte, G. Sithole and Rabbani, T., (2004). Recognising Structure in Laser Scanner Point Clouds, International Conference NATSCAN "Laser-Scanners for Fores and Landscpae Assessment – Instruments, Processing Methods and Applications, ISPRS working group VIII/2, Freiburg im Breisgau, Germany.

Wilson H. G., B. Boots, A.A. Millward, (2004). "A comparison of hierarchical and partitional clustering techniques for multispectral image classification," in Proceedings of the International Geoscience and Remote Sensing Symposium, pp. 1624-1626, Toronto, Canada, 2002.

Wu B and Shi Z., (2001). A Clustering Algorithm based on Swarm Intelligence. In Proceedings of the International Conference on Info-tech and Info-net, pages 58–66, 2001.

Zhong S. and J. Ghosh, (2003). A comparative study of generative models for document clustering. In SIAM Int. Conf. Data Mining Workshop on Clustering High Dimensional Data and Its Applications, San Francisco, CA, May 2003.