

A Multi-Agent strategy for automatic 3D object recognition based on the fusion of Lidar range and intensity data

Farhad Samadzadegan, Fatemeh Tabib Mahmoudi

Dept. of Geomatics, Faculty of Engineering, University of Tehran, P.O Box: 11365-4563, Tehran, Iran
{samadz, fmahmoudi}@ut.ac.ir

Abstract. Three dimensional object recognition and extraction from Lidar and other airborne or space borne data have been an area of major interest in photogrammetry for quite a long time. Therefore, many researchers have been trying to study and develop automatic or semi-automatic approaches for object extraction based on sensory data in urban areas. Lidar data have proved to be a promising data source for various mapping and 3D modeling of objects. But, according to the complicated relationships between objects in urban areas, especially buildings and trees, the performance of obtained results from ordinary object recognition algorithms based on Lidar data, is still dependent on several assumptions and simplifications. In this paper a multi-agent strategy has been proposed for automatic building and tree recognition based on the fusion of textural and spatial information extracted from Lidar range and intensity data. In this multi-agent methodology, two different groups of object recognition agents are defined for building and tree recognition in parallel and the algorithm has two different operational levels based on the types of contextual information. According to the difficulties in the field of building and tree detection based on the textural descriptors or spatial context, using the communicational behaviors and other capabilities of the multi-agent systems can be so useful in the field of 3D object recognition in dense urban areas. The evaluation of obtained results of the proposed methodology confirms the high capabilities of Lidar data and this multi-agent algorithm to decrease the conflicts in the field of building and tree recognition in parallel.

Keywords: Multi-Agent, Lidar Data, Information Fusion, Spatial Descriptors, Textural Descriptors

1 Introduction

The increasing demand for generation of three dimensional city models and their update, led to research efforts which aim to set up automatic tools for the extraction and recognition of 3D man-made and natural objects. The spectrum of application areas dealing with 3D city models is huge: environmental planning and monitoring, telecommunication, location based services, navigation, virtual reality, cadastre etc. Therefore, many researchers have been trying to study and develop automatic or semi-automatic approaches for object extraction in urban areas [1, 5, 13, 15]. Due to similarities between contextual information and complex relationships between objects in an urban area, the most complicated problems in the field of object recognition are related to the 3D objects such as buildings and trees.

Light Detection and Ranging (Lidar) is an active remote sensing technology that directly provides 3D coordinates of objects on the surface. These coordinates can be converted to surface and terrain models by algorithms with a high degree of automation [11]. Lidar data has proved to be a promising data source for 3D object recognition and modeling. But, due to the difficulties in the case of detecting buildings and trees, the performance of obtained results from ordinary object recognition algorithms based on Lidar data is still dependent on several assumptions and simplifications. So, intelligent techniques such as multi-agent systems may be regarded as a powerful tool in order to facilitate 3D object recognition based on Lidar data.

Intelligent Agent is a valuable concept in distributed artificial intelligence with strong abilities for solving complicated problems in different applications. Therefore, some of the researchers have proposed proper agent-based methodologies to solve the difficulties in the field of automatic object recognition. On the other side, agents cannot have complete information on their environment. Thus, to solve complicated problems or reach goals, agents must work with other agents. In a multi-agent model, several agents with specific goals and tasks are deployed, and they are trying to reach the main goal together [21]. Therefore, definition of proper social abilities and communication, coordination and negotiation between agents may improve overall results in solving complicated problems by multi-agent applications.

In this paper, a proper multi-agent methodology has been proposed for building and tree recognition based on the fusion of textural and spatial context extracted from Lidar range and intensity data.

Proposed multi-agent algorithm is an intelligent fusion technique in order to fuse different types of Lidar data on the level of their contextual information.

2 Literature Review

According to the remotely sensed data capabilities, many researchers have been proposing different algorithms for automatic or semi-automatic object recognition based on different types of sensory data such as aerial photographs or high resolution satellite imagery and Lidar data. Dealing with digital imagery, whenever there is no stereo pair in the study area, automatic object recognition process based on the mono imagery has some difficulties. Therefore, some researchers have proposed complex methodologies such as Hough transformation or snake models in order to automatic or semi-automatic object recognition based on mono imagery [6, 10]. On the other side, some of the proposed 3D objects recognition algorithms are based on digital surface model generation from stereo images. In these methodologies, after generating DSMs based on proper stereo matching techniques, relief information can be analyzed in order to improve object recognition results [3, 19, 16]. However, analyzing digital values of a DSM may be very useful for automatic object recognition, but generating DSM based on stereo matching techniques has a lot of difficulties. In addition, this kind of DSM doesn't have enough details for distinguishing between different objects. Therefore, after coming Lidar technology, many researchers have been utilizing Lidar data solely or with combination by digital imageries in order to facilitate automatic 3D object recognition processes.

One of the important strategies in automatic object recognition based on Lidar products is terrain filtering and generating a normalized surface of all 3D objects such as buildings and trees and then, detecting a specified object among all other 3D objects on that normalized surface [13, 17, 12, 4, 15, 14].

In this strategy, a surface containing all 3D objects in the scene is produced by subtracting the generated digital terrain model from the main Lidar DSM. This new surface is called Normalized Digital Surface Model. Most of the proposed object recognition algorithms utilize different classification methods to discriminate buildings or trees among other objects in the nDSM [4, 8, 14]. According to the similarities between contextual information of urban objects such as buildings and trees and complex spatial relationships such as vicinity between them, automatic building detection among trees has major difficulties. So, the researchers have been utilizing different solutions to improve their classification results and decrease object recognition problems. The surface roughness extraction and textural descriptor measurements based on Lidar products are some of the popular proposed methodologies for 3D object recognition facilitation [13].

In addition, differences between the first and last return Lidar products have been proposed for vegetation removal in some of the methodologies [1, 14]. On the other side, the fusion of the Lidar range data with digital imagery has a high potential in order to discriminate between buildings and trees. Therefore, many object recognition algorithms have been proposed based on the fusion of Lidar and imagery [9, 17, 14, 5].

The intelligent techniques also have high potential in order to facilitate automatic 3D object recognition based on sensory data. A multi-agent algorithm in [21] has been proposed for automatic building and tree detection based on the fusion of Lidar range DSM and a color infrared aerial photograph. In this multi-agent algorithm, two different groups of object recognition agents are defined for building and tree detection in parallel.

3 Proposed Methodology

According to the high capabilities of intelligent agents for solving complex problems, proper definition of a multi-agent system is useful to decrease major difficulties in the field of automatic building and tree recognition among other objects in dense urban areas. Therefore, in this research, a multi-agent system has been proposed for automatic object recognition based on the fusion of contextual information extracted from Lidar range and intensity data. The general process of this proposed algorithm can be divided into two main phases, pre-processing and multi-agent object recognition. In the pre-processing stage, all of the range and intensity Lidar data should be prepared for building and tree recognition facilitation. Then in the multi-agent object recognition stage, the pre-defined agents perform proper analysis based on textural and spatial information in order to separate buildings among trees.

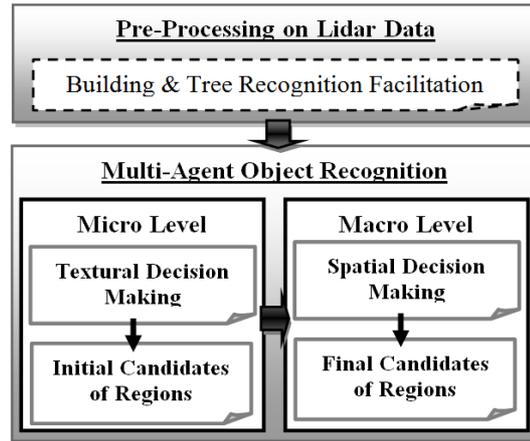


Fig. 1. The work flow of the proposed methodology

3.1 Pre-Processing on Lidar Data

Before utilizing Lidar data in multi-agent object recognition algorithm, different pre-processing operations should be performed based on types of recognizing objects. Removing terrain structures may be facilitating automatic building and tree recognition. For this reason, morphological grayscale reconstruction algorithm based on geodesic dilation is performed for terrain structure elimination [2].

$$\delta_I^{(1)}(J) = (J \oplus B) \wedge I \quad (1)$$

$$M_r = \delta_I^{(1)}(J) \circ \delta_I^{(1)}(J) \circ \dots \circ \delta_I^{(1)}(J) \quad (2)$$

In the above equations, I is the Lidar range DSM and J is the result of subtracting a height offset from it. \oplus is the dilation operator with elementary isotropic structuring element B . \wedge stands for the point wise minimization between I and J .

$$nDSM = LidarDSM - M_r \quad (3)$$

Then, M_r as the morphological reconstructed image is subtracted from original Lidar range DSM to generate the normalized DSM containing all 3D objects in the scene. This is the main pre-processing stage for building and tree recognition in order to eliminate terrain structures and normalized DSM generation based on Lidar range products.

The elimination of the majority of trees also decreases some difficulties in the field of automatic building recognition. On the other side, the normalized differences between the first and last return Lidar data can indicate the locations of the majority of trees and vegetation areas. Thus, subtracting these normalized difference indexes from nDSM of Lidar range data and original Lidar intensity products facilitates automatic building recognition based on Lidar range and intensity data, respectively. So, this is the second pre-processing operation just for automatic building recognition based on Lidar data.

$$PreparedRangeData = nDSM - NDI_{Range} \quad (4)$$

$$PreparedIntensityData = LidarIntensity - NDI_{Intensity} \quad (5)$$

In (4) and (5), NDI_{Range} and $NDI_{Intensity}$ are the normalized difference indices between the first and last return Lidar range and intensity products, respectively. Using the above mentioned pre-processing operations, all of the first and last return Lidar range and intensity data can be ready for object recognition. So, the multi-agent object recognition stage of the proposed methodology can be started.

3.2 Multi-agent Object Recognition

As depicted in Fig.2, this proposed multi-agent methodology has two different groups of object recognition agents. One group contains building recognition agents and another one contains tree recognition agents. In addition, each of the object recognition agent groups has two different kinds of agents based on the types of contextual information, textural and spatial. A coordinator agent and a yellow page are also defined in this multi-agent system in order to facilitate the coordination and communication between different recognition agents.

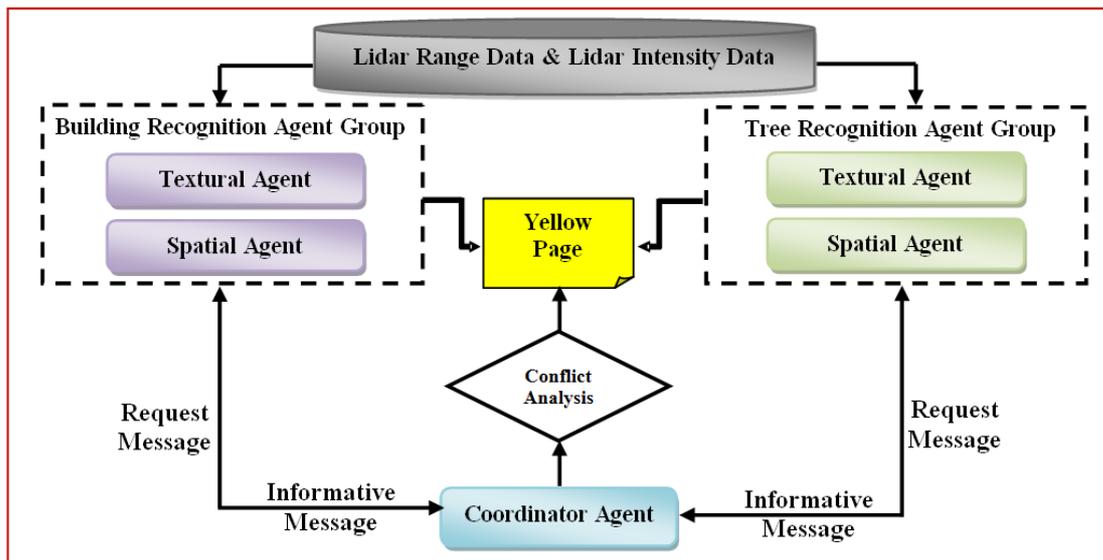


Fig. 2. The structure of the proposed multi-agent system

In general, the process of this multi-agent algorithm can be divided into the two main levels based on the types of contextual information. In the first operational level that is called micro level object recognition, textural descriptors have been utilized by the object recognition agents to generate candidates of building and tree regions. The second operational level that is called macro level object recognition is based on spatial information calculated for each candidate of building and tree regions in order to improve the overall results of the system.

3.2.1 Micro Level Process; Textural Decision Making

In the micro level operations, the textural object recognition agents perform some proper actions in order to recognize buildings or trees based on Lidar data. For this reason, the knowledge layer of these object recognition agents should be structured based on the optimum textural information. Therefore, different kinds of texture generation and analysis methods should be applied separately on the Lidar data. Then, after performing the diversity analysis on their results, the optimum texture features will be selected.

Texture descriptors can be measured based on the gray value relationships between each pixel and its neighboring pixels in a local window or in the global image. So many researchers have been utilizing different texture analysis methods in their proposed object recognition algorithms [22].

Using Gabor filter bank is so effective for texture analysis. Therefore, some researchers used this method in their segmentation and feature extraction algorithms [7, 20]. Gabor filters are linear and local and based on the below equation, their kernel is composed of a Gaussian and a cosine function. By convolving Gabor kernel and Lidar data, a new space generates for texture measurements.

$$g_{\lambda,\theta,\phi,\sigma,\gamma}(x,y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \phi\right) \quad (6)$$

On the other side, one of the important strategies in texture measurements is using the statistical methods. Some of the statistical methods directly use the histogram of the pixel gray values in the image and other ones utilize the relationships between the pixel pairs with a random orientation for statistical texture measurements [18]. The gray level co-occurrence matrix, GLCM, is one of the most popular statistical methods for texture measurements in the co-occurrence matrix space. In this method, by definition a local neighboring window and proper orientation selection, the relationships between the gray values of the pixels transforms to the co-occurrence matrix space.

Another effective statistical method in the texture analysis is computing central moments with proper orders based on the relationships between neighboring pixels.

$$\xi_{p,q} = \sum_m \sum_n (m - \bar{m})^p (n - \bar{n})^q f(m,n) \quad (7)$$

In (7), $f_{(m,n)}$ is an image as a 2D function and its central moment with the order of $(p+q)$ and in the centre of (\bar{m}, \bar{n}) is depicted with $\xi_{p,q}$.

Variance operators are one of the useful statistical methods for texture analysis based on the differences between the pixels in a local neighboring window. Semi-variogram as a geo-statistical operator is also in the family of variance operators. The operation of semi-variogram is based on the directional distances between each pixel and its neighboring pixels in a local window.

$$\gamma_{(h)} = \frac{1}{2n_{(h)}} \sum_{i=1}^{n_{(h)}} \left((I(x_i) - I(x_i + h))^2 \right) \quad (8)$$

In the above equation, $n_{(h)}$ is the number of pixels who are in the distance lag h from the central pixel x_i in the local neighboring window.

The main objective of applying different texture operators on Lidar data is providing a valuable knowledge-base for textural object recognition agents. Therefore, it is necessary to perform another process to select optimum texture features in order to structure agent's knowledge layer. Thus, the proper diversity analysis process is performed based on the comparison between the properties of the different texture descriptors.

After preparation the optimum texture features and provide agent's knowledge in the micro level object recognition process, the main operational behavior of recognition agents will be started. In this stage, each of the object recognition agents perceives the first and last return Lidar range and intensity data and uses some reasoning rules based on their textural knowledge. Finally, the building recognition agent generates a binary image of building candidate regions and the tree recognition agent generates such a binary image of tree candidates, in parallel.

According to the certain nature of Lidar products and the similarities between textural descriptors of different objects, the building and tree candidates may have some conflicts with each other. The coordinator agent analysis both of the generated binary images to detect conflicts among them and sends proper messages to the textural object recognition agents for solving these conflicts. However, doing especial reasoning based on the spatial descriptors calculated for each of the candidate regions can improve the overall object recognition results.

3.2.2. Macro Level Process; Spatial Decision Making

In the macro level object recognition process, some of the suitable spatial descriptors should be calculated for each of the building and tree candidate regions. For instance, rectangularity as a useful spatial attribute indicates how well a region can be described by a rectangle. On the other side, by comparison the area of the region to the area of the smallest convex polygon, another valuable shape descriptor is calculated which is called solidity. After calculating the proper spatial descriptors for each of the initial candidate of building and tree regions, the knowledge layers of the pre-defined spatial object recognition agents will be provided. Both of the spatial building and tree recognition agents use their knowledge-base in order to modify and improve the overall results of the multi-agent system.

Because of the complex spatial relationships between buildings and trees in an urban area, it may be possible to merge a building region with some very close trees. In addition, it may be also possible that a large building area splits into several smaller regions and recognized by the textural building recognition agent in micro level process. Therefore, the pre-defined spatial building and tree recognition agents perform some proper actions in order to solve these problems based on their

spatial knowledge. The first task of these agents is performing the specified split and merges operation on the building candidate regions and tree candidate regions. This operation may modify the overall results of the building and tree recognition.

After that, the coordinator agent again analyses the conflicts between the two binary images of modified building and tree regions. If there are still any conflicted regions, both of the spatial building and tree recognition agents are requested to solve their conflicts based on the proper spatial reasoning.

4 Experiments and Results

The proposed multi-agent methodology was implemented on the Lidar range and intensity DSM data in an urban area in Stuttgart, Germany. As depicted in Fig .3, this urban area has a lot of buildings with variety of shapes and sizes. In addition, the vicinity of tall trees with some of the building regions causes major difficulties in recognizing buildings among trees.

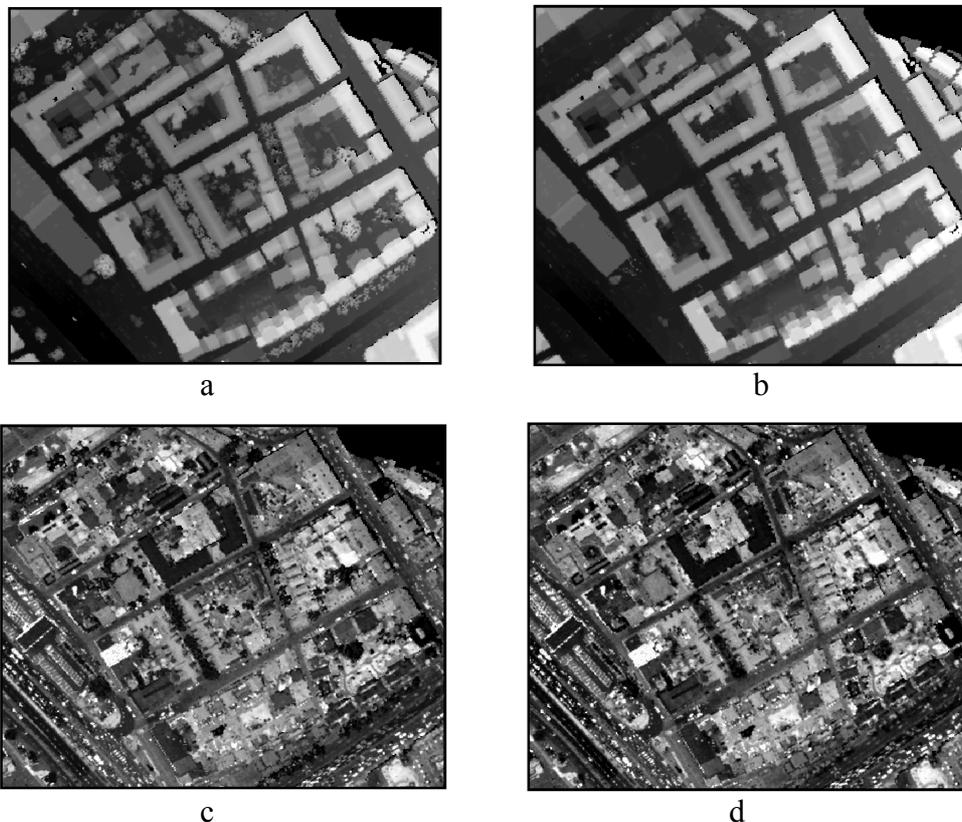


Fig. 3. a. First return Lidar range data (FPR), b. Last return Lidar range data (LPR), c. First return Lidar intensity data (FPI), d. Last return Lidar intensity data (LPI)

As mentioned above, removing terrain structures and majority of trees facilitate automatic building and tree recognition. Then, generating nDSM and subtracting NDI from it is necessary for proper texture measurements based on Lidar range data. In the case of intensity data, normalized difference indices are directly subtracted in order to remove the majority of trees and vegetation areas.

For automatic tree recognition, generating the nDSM based on the range data is sufficient for proper texture measurements, and the intensity data should be entered directly.

After performing the pre-processing operations, different texture measurement methodologies should be applied on Lidar data to generate a rich texture feature space for building and tree recognition, separately. In Figure.4, some textural descriptors based on the prepared data for building recognition and tree recognition have been shown separately in the selected region of the study area.

According to the different properties of various texture features, selecting a set of optimum features is so effective in order to structure the knowledge-base of the recognition agents. Therefore, diversity analysis has been performed for optimum feature selection.

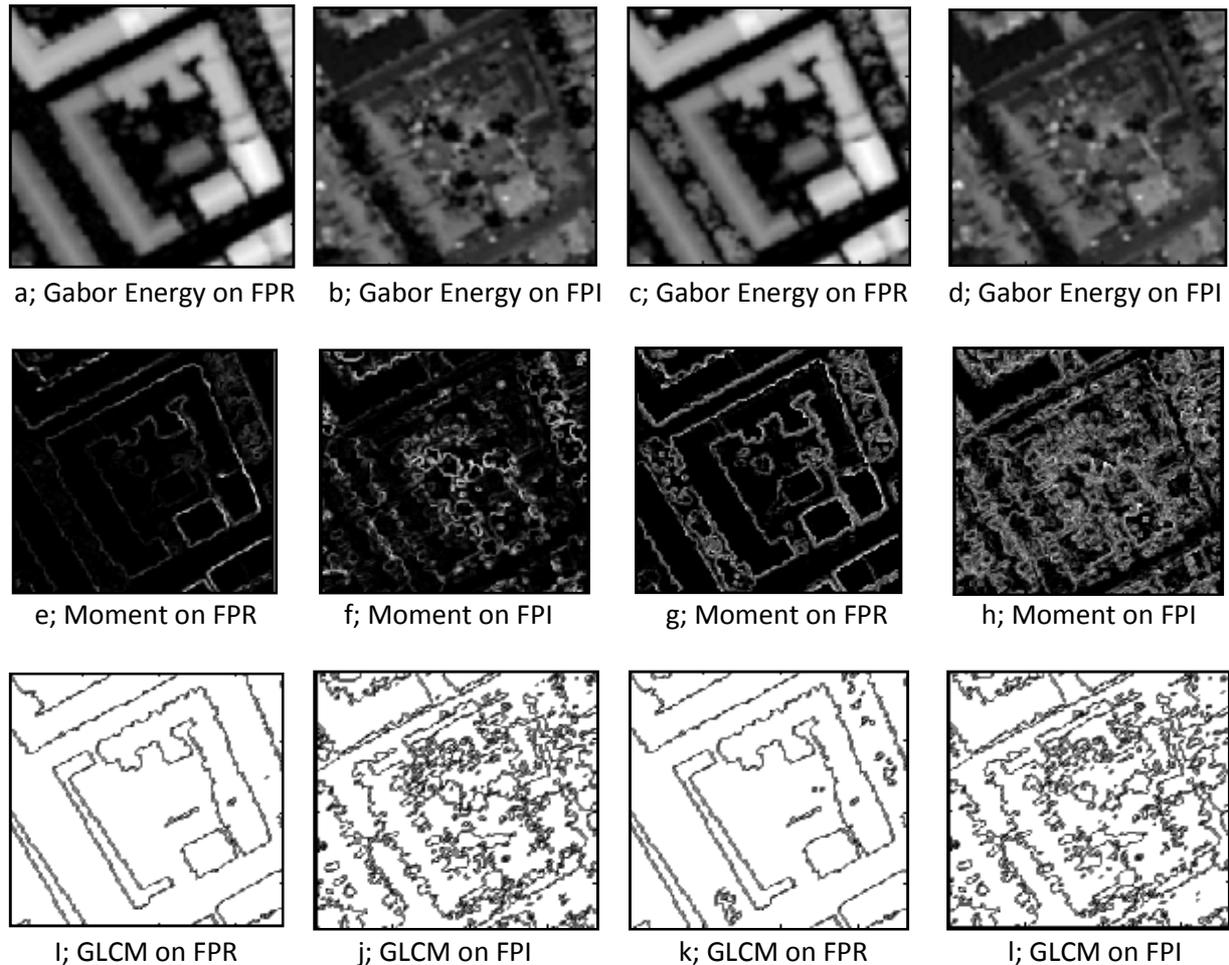


Fig. 4. Different textural descriptors for automatic building recognition based on range & intensity data, (a , b, e, f, i , j), Different textural descriptors for automatic tree recognition based on range & intensity data, (c , d, g , h, k , l)

Then, each of textural building and tree recognition agents performs proper reasoning for generating a binary image of the candidate regions based on the textural knowledge. But, because of the similarities between textural information of different urban objects and the certain nature of Lidar data, the candidate of building and tree regions have conflicts with each other. Therefore, calculating spatial descriptors for each of the candidate regions and performing some reasoning by the spatial agents is necessary to improve the overall object recognition results.

As depicted in Fig .5, the results of the building and tree recognition based on spatial reasoning in the macro level process have been improved and the majority of wrongly detected areas have been corrected.

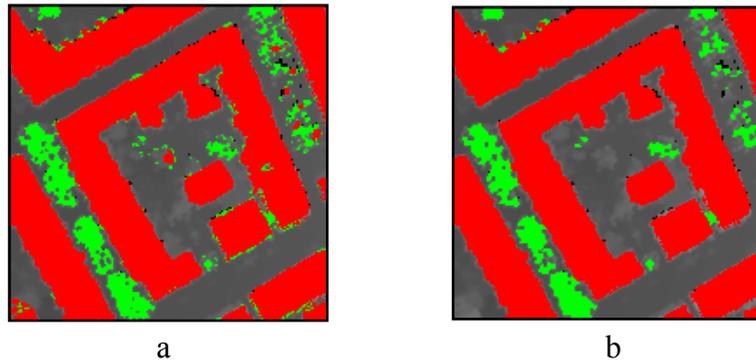


Fig. 5. a. The results of the micro level building and tree recognition, b. The results of the macro level building and tree recognition

5 Accuracy Assessment

For the quantitative evaluation of the obtained results from the proposed multi-agent methodology in this paper, some of the building and vegetation areas are manually digitized and then compared with their corresponding results from multi-agent algorithm.

Then, the total number of pixels in each digitized area and its corresponding agent-based recognized area are determined. In addition, the number of wrongly detected pixels, false positive, in the results of the recognition agents and the number of pixels those weren't recognized by agents, true negative, will be determined in each of the micro and macro level object recognition results.

Table 1. Quantitative evaluation of the obtained results from the proposed multi-agent algorithm

	Object	Digitized Areas	Agent-Based Areas	True Positive	False Positive	True Negative	Accuracy
Micro Level Results	Building	40361	34020	33416	604	6927	82.8%
	Tree	2662	1597	1397	200	1256	52.5%
Macro Level Results	Building	40361	35028	34807	221	5554	86.2%
	Tree	2662	1982	1784	198	878	67%

According to the quantitative evaluation, the results of the textural decision making by the building and tree recognition agents have some wrongly detected areas. So, after using spatial descriptors in macro level operations, most of the mistakes will be corrected and the overall accuracy of the results will be improved.

6 Conclusion

According to the difficulties in the field of automatic 3D object recognition from Lidar data and because of the high capabilities of multi-agent systems in solving problems a proper multi-agent methodology has been proposed for building and tree recognition in parallel, in this paper. In this methodology an information fusion technique is utilized based on the textural and spatial context extracted from Lidar range and intensity data. Using high capabilities of multi-agent systems, one can solve most of the main problems in the field of automatic building and tree detection in complex urban areas.

However, this method still needs some more modifications in the field of definition of the agents and the contextual information which is used in recognition process. Moreover, incorporating of different source of information, e.g. digital imagery, in the different stages of recognition process, could efficiently modify the potential of the proposed methodology.

Reference

- [1] Alharthy, A. and Bethel, J. Heuristic filtering and 3d feature extraction from Lidar data. In Proceeding of the ISPRS technical commission III symposium 2002, Photogrammetric computer vision, September, Austria. Graz. (2002).
- [2] Arefi, H. and Hahn, M. A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data. ISPRS WG III/3, III/4, V/3 Workshop, Laser scanning 2005, 12-14 September, the Netherlands, 120-125. (2005).
- [3] Baltsavias, E. and Mason, S. And Stallmann, D. Use of DTMs/DSMs and orthoimages to support building extraction. In: A. Gruen and O. Kubler and P. Agouris, eds. Automatic Extraction of Man-Made Objects from Aerial and Space Images, Birkhauser, Basel, 199–210. (1995).
- [4] Brunn, A. and Weidner, U. Hierarchical Bayesian nets for building extraction using dense digital surface models. ISPRS Journal of Photogrammetry & Remote Sensing, 53 (1998), 296–307. (1998).
- [5] Demir, N. Poli, D. and Baltsavias, E. Extraction of buildings and trees using Images and Lidar data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B4. Beijing. 313-318. (2008).
- [6] Gonzales, R. and Woods, R. Digital image processing. Massachusetts: Addison-Wesley publishing. (1993).
- [7] Grogorescu, S.E. and Petkov, N. and Kruijzinga, P. Comparison of texture features based on Gabor filters. IEEE TRANSACTIONS ON IMAGE PROCESSING, 11(10), 1160-1167. (2002).
- [8] Haala, N. and Brenner, C. Extraction of buildings and trees in urban environments. ISPRS Journal of Photogrammetry and Remote Sensing, 54 (2-3), 130-137. (1999).
- [9] Matikainen, L. and Kaartinen, H. and Hyypä, J. Classification tree based building detection from laser scanner and aerial image data. ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007, 12-14 September, Finland, 280-287. (2007).
- [10] Mayunga, S.D. and Zhang, Y. and Coleman, D.J. Semi-automatic building extraction utilizing quickbird imagery. In: U. Stilla, F. Rottensteiner, S. Hinz, eds. CMRT05: Object Extraction for 3D City Models, Road Databases, and Traffic Monitoring - Concepts, Algorithms, and Evaluation, IAPRS, Vol. XXXVI, part3/W24, 29-30 August, Austria, Vienna. 131-136. (2005).
- [11] Morsdorf, F. Lidar remote sensing for estimation of biophysical vegetation parameters. Thesis (PhD). Zurich University. (2006).
- [12] Neidhart, H. and Sester, M. Extraction of building ground plans from Lidar data. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B2. Beijing. 405-410. (2008).
- [13] Rottensteiner, F. and Jansa, J. Automatic extraction of buildings from Lidar data and aerial images. Commission IV, WG IV/7. (2001).
- [14] Rottensteiner, F. and Trinder, J. and Clode, S. and Kubik, K. Building detection using Lidar data and multi-spectral images. In: C. Sun and H. Talbot and S. Ourselin and T. Adriaansen, eds. VIIth Digital Image Computing: Techniques and Applications, 10-12 December, Sydney, 673-682. (2003).
- [15] Samadzadegan, F. Automatic 3D object recognition and reconstruction based on artificial intelligence and information fusion techniques. Thesis (PhD). Dept. of Geomatics, faculty of engineering, University of Tehran. (2002).
- [16] San, D.K. and Turker, M. Automatic building extraction from high resolution stereo satellite images. Commission VII. (2005).
- [17] Sohn, G. and Dowman, I. Building extraction using Lidar DEMs and Ikonos images. Commission III, WG III/3. (2002).
- [18] Tuceryan, M. and Jain, A.K. Texture analysis. In: C. H. Chen and L. F. Pau and P. S. P. Wang, eds. The Handbook of Pattern Recognition and Computer Vision. World Scientific Publishing Co, 207-248. (1998).
- [19] Vozikis, G. Urban data collection: an automatic approach in remote sensing. Application of high resolution remote sensing data. Vienna University of technology, Austria: Institute of Photogrammetry and Remote Sensing, 101-110. (2004).
- [20] Wei, H. and Bartels, M. Gabor Wavelets and Gaussian Models to Separate Ground and Non-ground for Airborne Scanned LIDAR Data. School of Systems Engineering. The University of Reading. Reading RG6 6AY, UK. (2005).

- [21] Yoon, T. Object recognition based on multi-agent spatial reasoning. Thesis (PhD). Ohio State University. (2008).
- [22] Zhang, Y. Optimization in building detection in satellite images by combining multispectral classification and texture filtering. ISPRS Journal of Photogrammetry & Remote Sensing, 54(1999), 50-60.(1999).