Semi-automatic Objects Recognition Process Based on Fuzzy Logic

Federico Prandi and Raffaella Brumana

Politecnico di Milano, Dept. B.E.S.T.
Piazza Leonardo da Vinci 32, 20133 Milano
{federico.prandi,raffaella.brumana}@polimi.it

Abstract. Three dimensional object extraction and recognition (OER) from geographic data has been one of most important topics in photogrammetry for a long time. Today, the capability of being able to rapidly generate high-density DSM increases the provision of geographic information. However the discrete nature of the measuring makes it more difficult to correctly recognize and extract 3D objects from these surfaces. The proposed methodology wants to semi-automate some of the operations required for clustering of geographic objects, in order to perform the recognition process. Fuzzy logic allows using, in a mathematical process the uncertain information typical of human reasoning. In this paper we present an approach for detecting objects based on fuzzy logic. In a first phase only the structural information are extracted and integrated in the fuzzy reasoning process in order to have a more generic treatment. The recognition algorithm has been tested with different data sets and different objectives.

Keywords: Objects Recognition, DSM, Fuzzy logic, disaster management.

1 Introduction

Three dimensional object extraction and recognition (OER) from geographic data has been one of most important topic in photogrammetry and remote sensing for a long time. However, most of the existing methods for automatic extraction and recognition of objects from data are based on a range of different information and make use of parametric methods. Within these systems object's vagueness behavior is basically neglected [1].

Manual intervention is still needed to reconstruct 3D models introducing a critical bottleneck to the modeling of geographic objects. Aerial photogrammetry has been, and still is, one of the preferred ways to retrieve three-dimensional information on objects, being it very well understood and since it delivers accurate results. The major drawback is that automation of the measurement process is closely related to image understanding which is a problem hard to solve.

Experience and daily practice make it possible for our brain to automatically interpret what we see. Human recognition takes advantage of a variety of acquired information rather than relying on a single descriptor of an object. Further, human perception has a tremendous potential for learning and it deals perfectly with the fuzziness of the real world. Whenever it is required to identify objects within

K. Sithamparanathan et al. (Eds.): PSATS 2010, LNICST 43, pp. 343–353, 2010.

geographic data an interpretation made by a human operator may represent an easier option. However if the process were performed by a computer, it would be very likely that none of the objects were identified.

Fuzzy logic gives the possibility to use, within a mathematical process, uncertain information which is typical of human reasoning. The purpose of this paper is to use 3D information contained in the DSM to automatically detect and recognize topographic objects in complex scenes. Furthermore, we try to use only LiDAR or Image Matching DSM. Due to this limitation our goal becomes even more challenging. As illustrated in the future work section it is possible to extend the methodology to others data sets such as multispectral or high resolution satellite images.

The proposed approach can be divided into five major steps:

- Pre-processing (data acquisition, interpolation, matching).
- DSM normalization.
- Extraction and definition of object's structural descriptor.
- Fuzzy reasoning process (membership function, IF-THEN Rules, inference process).
- Object detection.

It is clear that the problem of automatically extracting objects is still far from being solved. Now, after about two decades of research on the topic of automated recognition and reconstruction of man-made objects, there are still no fully automatic systems. The variety of methods available and the analysis of their advantages and weaknesses can provide quite a handful of hints for other scientists on how to approach the problem of 3D object extraction.

1.1 Related Works

As seen above, the research issues involved in the generation of 3D topographic objects for 3D GIS are very wide. They range from ICT theory of conceptual models to virtual reality and software development, passing through the development of data acquisition sensors, automation in data extraction and data analysis. Each one of these issues is related to many others: for instance in order to develop a given conceptual model it is important to be aware on how data is acquired. Similarly, when developing an algorithm for feature extraction, one has to understand which is the model to be used with those features.

In fact there is not a universal automatic or semi-automatic approach and the process of 3D reconstruction is often manual. The main task required to automatically generate a DTM from data is to divide terrain points from non-terrain points. Further it may be necessary to classify non-terrain points as belonging to buildings, vegetation or other objects (e.g. bridges), depending on the specific application.

Automatic approaches to generate DTM and building reconstruction based on radiometric images have since long been a challenging research topic. In the literature it is acknowledged that the issues related to segmentation and classification are not only an interesting research topic, but they are also very important in practice. Multiple, largely complementary, sensor data, such as color or multi-spectral aerial images and range data from laser scanners or SAR, have been used to ensure robustness and better performances in 3D ORR (Object Recognition and Reconstruction). For example, color infrared (CIR) aerial images can be combined to laser scanner data DTMs for feature extraction purposes [2].

If multispectral imagery is available, the classification approach is the most convenient way to detect building areas or urban regions [3]. Everything that exceeds a certain threshold in the nDSM (normalized Digital Surface Model) will be included, and vegetation will be then excluded by further interpreting the NDVI (Normalized Difference Vegetation Index).

In many cases however multispectral information is missing and therefore techniques to segment the DSM are adopted. Different filtering strategies have been proposed, based on deviations from a parametric surface, slope threshold, clustering, etc.

A slope-based filtering using mathematical morphology has been presented in [4] and it defines a slope threshold as the maximum allowed height difference between two points as a function of their spatial distance. Another interesting approach has been discussed in [5]: the segmentation is carried out by combining region growing with a principal component analysis (PCA). A three-stage framework has been implemented for a complete, robust and automatic classification of LiDAR data. This is composed by a region-growing technique to identify regions with a step edge along their border, a grouping of connected sets of pixels on the basis of an 8-classes partition of the height gradient orientation and a rule based scheme applied to the classification of the regions [6]. Another study [7] focuses on automatic extraction of a DTM from a high-resolution DEM produced by image correlation in urban or rural areas based on a hybrid approach. The study combines complementary aspects of both TIN-based and segmentation-based techniques.

Once the locations of features have been identified, either through automatic segmentation or manual digitization, the feature extraction process can be started. The different approaches available depend on whether image data or laser scanner data have to be processed, mainly due to the differences in the nature of the data.

The reconstruction of man-made objects is a task of major concern nowadays, and a rich variety of approaches has been proposed in recent years. A comprehensive overview is given in the proceedings of the Ascona Workshops [8].

However, despite the progress that has been made with scanning systems and digital image acquisition, achieving automatic processing of the resulting datasets is at a very early research stage. For example, nowadays laser scan data is mostly used to produce digital terrain models which can be obtained from the original (measured) point cloud through interpolation algorithms which are do not differ significantly from photogrammetric DTM modules that have been in use for the last three decades. Only in specialized applications, such as the derivation of DTM in wooded areas or the surveying of power lines, there are first approaches capable to exploit specific properties of laser scan datasets to achieve automatic extraction.

2 Classification Based on Fuzzy Logic

With the term recognition we refer to the process that assigns a label (e.g. "building" or "tree") to the result of a segmentation, in particular a region, based on properties (descriptors) of the region. Segmentation is the process that partitions the spatial domain of an image, or other raster datasets, like digital elevation models, into mutually exclusive parts, called regions.

Fuzzy logic can provide the link to connect computational theories with human perception. It provides a simple way to get to a specific conclusion based upon imprecise, uncertain, ambiguous, vague or missing input information. When dealing with simple 'black' and 'white' answers is no longer satisfactory, a *degree of membership* (as suggested by Prof. Zadeh in 1965) becomes the way to tackle a range of different problems. The natural description of problems, in linguistic terms, rather than in terms of relationships between precise numerical values, is the other advantage of fuzzy logic theory.

In fact fuzzy logic introduces *linguistic variables* for each descriptors characteristic of an object and *linguistic labels* to describe the fuzzy sets on the range of all possible values that those linguistic variables may be equal to.

When applied to raster images, fuzzy classification estimates the contribution of each class within each pixel. The theory assumes that a pixel is not an indecomposable unit in the image analysis. Consequently, it works on the principle of "one pixel—several classes" to provide more information about the pixel, unlike hard classification methods which perform poorly when extracting information.

The research works dealing with classification and feature extraction [9, 10] have provided several demonstrations of the capability of such fuzzy approaches.

3 Proposed Methodology

In the proposed fuzzy process (Fig. 1), each pixel is transformed into a matrix of degrees of membership representing the fuzzy inputs. A minimum-reasoning rule is then applied to infer the fuzzy outputs. Finally, a defuzzification step is applied to extract features [11].

The main components of the fuzzy recognition process are as follows:

- A database, which defines the membership functions of the fuzzy sets.
- A rule base, which contains fuzzy if-then rules.
- A fuzzy reasoning procedure, which performs inference operations on the rules.

The process should account for i) all available descriptors of an object (such as: 3D structure, textural information and spectral responses), ii) a fuzzy description of object properties and a fuzzy inference strategy for object recognition, iii) learning capabilities to be able to modify imprecise model descriptions and increase the potential of recognition in particular if new and unrecognized objects are encountered.

The proposed method for object extraction requires some preliminary steps that consist in: i) to locate and separate all 3D objects from the terrain and ii) to analyze and generate all the geometric properties that can describe the objects. Then, the challenge is to develop a segmentation procedure in connection with an inference process for object recognition.

A rule base should include observations of important descriptors. Moreover, it reflects the fact that people may formulate similar "fuzzy statements" to characterize how they perceive how objects appear, for instance, within aerial color images.

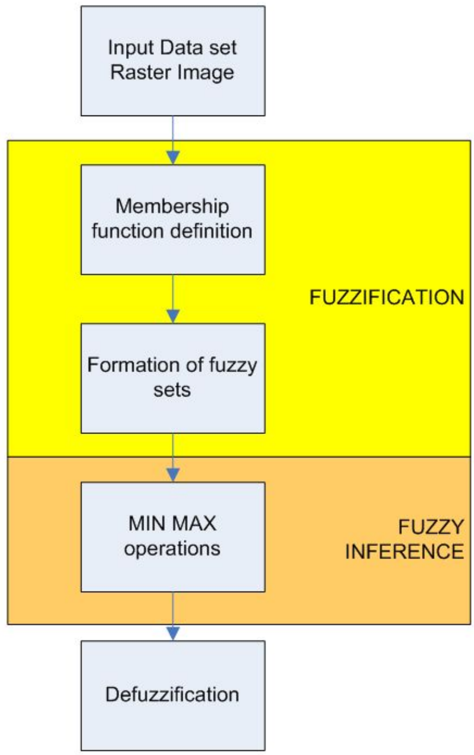


Fig. 1. Implementation flowchart of the steps required for fuzzy recognition process

The fuzzy inference AND or OR operators combine the membership values of the inputs in each rule for the antecedent of that rule. The MIN reasoning rule, applied on the matrix of produced fuzzy inputs, will consider, for each class, the membership degrees provided by the different fuzzy sets. Furthermore it will pick out the minimal membership degree to represent the class extent in the pixel. Then a MAX operation will be performed, as a result of each rule, to extract the element with the highest value (fuzzy output) and the corresponding class of that feature will be considered as associated fuzzy class to that pixel (fig. 2).

More specifically the steps of the fuzzy recognition process will be:

- Input raster data and structural descriptor variables are introduced.
- Membership functions are defined by using results from human heuristic knowledge.
- Definition of fuzzy logic inference rules.
- Performing raster data classification.

The OER process starts with the extraction of the Structural Descriptors (SD). Then those descriptors are analyzed to derive SD membership functions. In the next stage, the recognition operation is performed with the application of the sample if-then fuzzy rules and the MIN-MAX inference process.

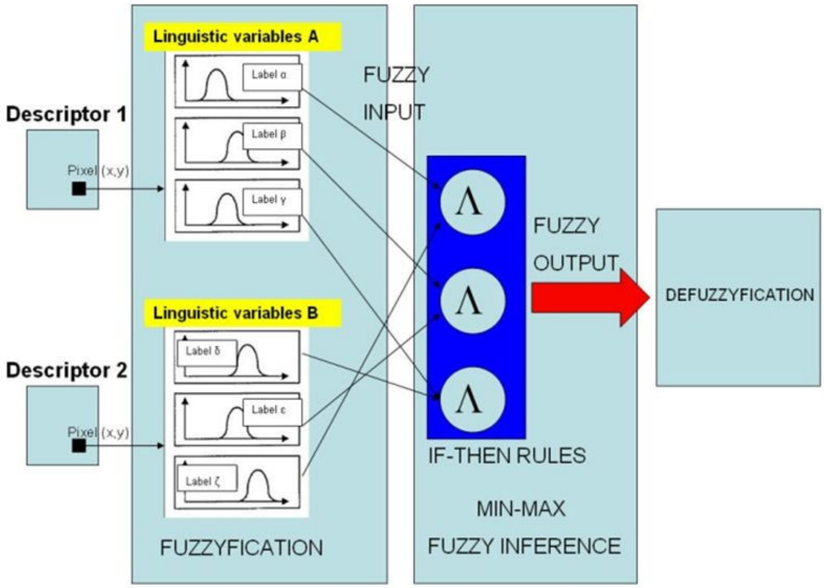


Fig. 2. Architecture of the explicit fuzzy method illustrated for two descriptors, two variables with three classes

The membership function parameters are determined through human heuristic knowledge. It should be noted that this method is more appropriate to situations where there is a clear linear ordering in the measurement of the fuzzy concept for instance when dealing with concepts such as tallness, heat, time, etc.

Interval estimation is a relatively simple way of acquiring the membership function and it results in membership functions that are "less fuzzy" (i.e. the spread is narrower) when compared to other methods.

The trapezoidal membership function, with maximum equal to 1 and minimum equal to 0, is used in this work. Special cases, including symmetrical trapezoids and triangles, reduce the number of parameters to three. The trapezoidal functions are modeled with four parameters (α,β,γ,δ,).

$$\mu_A(x,\alpha,\beta,\gamma,\delta)=\begin{cases}0 & if\ x<\alpha\\ \dfrac{x-\alpha}{\beta-\alpha} & if\ \alpha\le x<\beta\\ 1 & if\ \beta\le x<\gamma\\ \dfrac{\delta-x}{\delta-\gamma} & if\ \gamma\le x<\delta\\ 0 & if\ x<\delta\end{cases}$$

The linguistic variables, which are variables that assume in linguistic terms values of the object's structural descriptors, have to be defined in fuzzy logic. Linguistic variables are associated to each input structural descriptor and for each linguistic variables some linguistic labels has been assigned. This assignment is mostly a mixture of expert knowledge and examination of the desired input–output data.

Through definition of linguistic variables and membership function parameters calculation the input to the fuzzy recognition process is performed. Sufficient overlap of neighboring membership functions is taken into account to provide smooth transition from one linguistic label to another.

Table 1. Linguistic variables and labels for the fuzzy-based object recognition process

Structural descriptor	Linguistic labels
Height	*Very Low, Low, Medium, Tall, Very Tall*
Area	*Very Small, Small, Medium, Large, Very Large*
Gradients on segment borders	*Very Flat, Flat, Steep, Very Steep*
Relief	*Very Irregular, Irregular, Regular, Very Regular*
Height range in a point neighborhood	*Very Low, Low, Medium, Tall, Very Tall*

As mentioned earlier, the object recognition potentials can be enhanced through the simultaneous fusion of the SD parameters extracted. For this reason our recognition strategy is based on the concept of information fusion. The descriptors are used simultaneously within the recognition engine to perform the object recognition process.

Because of the wide variety of clustering cases we have decided to use a versatile tool for the management of fuzzy clustering process. Since many spatial phenomena are inherently fuzzy or vague or possess indeterminate boundaries, fuzzy logic has been applied in many GIS scenarios, including fuzzy spatial analysis, fuzzy reasoning, and the representation of fuzzy boundaries.

In a second step the membership values which have been identified have to be combined to get to a final decision (inference process). This component of the fuzzy recognition process consists in the definition of a set of rule bases, which contain fuzzy if-then rules.

Formulation of fuzzy rules requires a profound observation of the integrative impact of the descriptors on the recognition of an object. This again depends on the experience of an operator but also on the complexity of an object. In the experiments we observed, when dealing with partition walls it is sufficient to use a relatively small number of fuzzy rules, whereas when it identifying trees the process requires a higher number of rules, probably due to their more complex shapes and variety in terms of appearance.

By taking into account the geometric properties, it is possible to overcome some problems typical of the use of image spectral characteristics. However a better result can be achieved if the descriptors are not necessarily limited to the structural descriptor values. The process of information fusion may also include other types of descriptors if they are available. If instead there are only one or two descriptors available (e.g. only spectral, or spectral and structural), the recognition process can still be executed, albeit being more prone to less reliable results. The insertion of others descriptors in the fuzzy recognition process is always possible once the membership function and the if-then fuzzy base rules are defined.

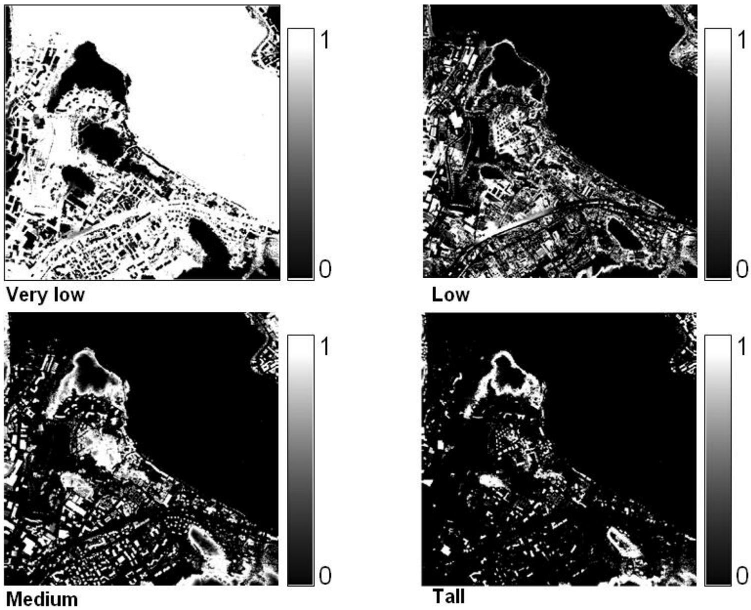


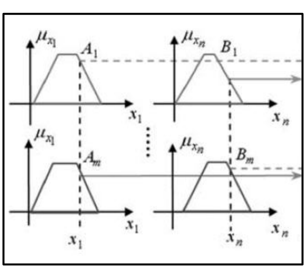
Fig. 3. Membership function's grids (as many as the linguistic variables associated to the descriptor) calculated for the structural descriptor "Height". The values are between zero and one.

If-then rules are statements that make fuzzy logic useful. A single fuzzy if-then rule can be formulated according to:

IF x is A; THEN y is B

Where A and B are linguistic variables defined by fuzzy sets on the range of all possible values of x and y, respectively. The antecedent may integrate several inputs using logical AND and OR. Fuzzy reasoning with fuzzy if-then rules enables linguistic statements to be treated mathematically.

Given the rules and inputs, the degree of membership to each of the fuzzy sets has to be determined. By combining the individual membership functions trough the simple rules and after the aggregation process we obtain the final result and we can then defuzzify it.



$$\left.\begin{aligned} m_1 &= \min(\mu_{A_i}(x_1), \mu_{B_i}(x_n)) \\ \\ m_n &= \min(\mu_{An}(x_1), \mu_{Bn}(x_n)) \end{aligned}\right\} \mu_c = \max(m_1, m_2, \ldots\ldots, m_n)$$

Fig. 4. MIN-MAX reasoning structure

Fig. 5. Results of fuzzy buildings recognition (Above), the grey level are proportional to the output fuzzy value of the inference process. We can qualitatively compare the results to the aerial image (below) of the area.

Our approach for fuzzy object recognition follows the MIN-MAX concept. The membership values in the premise part are combined according to minimum values (so-called min- inference) to extract the value of each rule, where m_i is the consequent for each equation rules and μAi(xi) the value of the membership functions for each linguistic variable of the premise part of the rule.

The final output (fig. 4) is then obtained by the MAX of the consequent equations rules and is calculated by:

$$\mu_c = \max(m_1, m_2, ..., m_n)$$

The qualitative results (Fig. 5) highlight a good classification of the buildings objects; in particular we have an optimal results in case of isolated buildings or urban areas. Instead more complex, with lack of recognition, is the case of building near to areas characterized by vegetation where large trees are also classified as buildings. This problem can be overcome using a spectral descriptor such as NDVI index which allows an efficient separation between objects and vegetation.

The investigations presented here have given a first demonstration of the capability of this approach. The recognition process could identify approximately 80% of the buildings objects within the area used for the test.

4 Conclusions and Future Works

In object recognition, human interaction remains an important part of the workflow even though the amount of work to be performed by the human operators can be reduced considerably in the global extraction process. Many automatic and semi-automatic methods proposed in the literature focus either on the reconstruction process or on feature extraction once the objects has been recognized by an operator.

The core of the system presented is an approach to recognize object's primitives through use of fuzzy logic theory. This allows analyzing the data to extract a maximal amount of information, through an explicit process that uses structural information of objects and integrates them within a fuzzy reasoning process.

In our approach detection, classification and modeling of objects is based exclusively on structural description, without additional information like GIS data.

Further, in this way we have been able to assess i) the versatility of the fuzzy recognition process in different challenging clustering situations, ii) the efficacy to recognize 3-D structural information iii) the capability to solve difficult problems by using the property to benefit from uncertain or vague concepts, typical of human thinking and language.

Despite this further investigation and developments might be carried out. With regard to data used as input, different object descriptors, such as textural or multispectral properties, could be considered in the recognition process. Moreover the generation of structural descriptors has to be improved by considering the sensitivity of the algorithms to coarse input.

Another important improvement is to provide learning capability through neural-networks. The learning capability of neural networks can be introduced in the fuzzy recognition process by taking adaptable parameter sets into account thus moving towards a neuro-fuzzy approach.

Finally data acquired from different sensors, such as high density airborne laser scanning, high resolution satellite images, or from other technologies such as Unmanned Aerial Vehicles (UAV), could be used to improve the recognition process and their effectiveness can be assessed.

References

1. Samadzadegan, F., Azizi, A., Hahn, M.T., Lucas, C.: Automatic 3D object recognition and reconstruction based on neuro-fuzzy modelling. ISPRS Journal of Photogrammetry & Remote Sensing 59, 255–277 (2005)
2. Hahn, M., Stätter, C.: A scene labeling strategy for terrain feature extraction using multi-source data. IAPRS 32 Part 3/1, 435–443 (1998)
3. Knudsen, T., Olsen, B.: Automated Change Detection for Updates of Digital Map Databases. Photogrammetric Engineering & Remote Sensing 69(11), 1289–1296 (2003)
4. Vosselmann, G.: Slope based filtering of laser altimetry data. In: IAPRS, Amsterdam, The Netherlands, vol. 33(B3), pp. 935–942 (2000)
5. Roggero, M.: Object segmentation with region growing and principal component analysis. IAPRS 34(3A), 289–294 (2002)
6. Forlani, G., Nardinocchi, C., Scaioni, M., Zingaretti, P.: Complete classification of raw LIDAR data and 3D reconstruction of buildings. Pattern Analysis Application 8, 357–374 (2006)
7. Baillard, C.: A Hybrid Method for Deriving DTMs from Urban DEMs IAPRS, XXXVII Part B3b Commission III p.109 (2008)
8. Workshops on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona/Switzerland in 1995, 1997 and 2001
9. Mohammadzadeh, A., Tavakoli, A., Zoej, M.J.V.: Automatic Linear Feature Extraction of Iranian Roads from High Resolution Multi-spectral Satellite Imagery. IAPRS XXXV part B3, 764 (2004)
10. Wuest, B., Zhang, Y.: Region Based Segmentation Of Quickbird Imagery Through Fuzzy Integration. IAPRS WG VII/4 XXXVII Part B7, 491 (2008)
11. Melgani, F., Al Hashemy, B.A.R., Taha, S.M.R.: An Explicit Fuzzy Supervised Classification Method for Multispectral Remote Sensing Images. IEEE Transactions On Geoscience And Remote Sensing 38(1) (2000)