extraction, in A. Gruen, O. Kuebler & P. Agouris, eds, 'Automatic Extraction of Man-Made Objects from Aerial and Space images', Birkhauser Verlag, Basel, Boston, Berlin,pp.199-210 (1995),.
[2] C. Brenner and N. Haala, *Rapid acquisition of virtual reality city models from multiple data sources*, in 'IAPRS, Vol 32, Part 5', Hakodate, Japan (1998b).
[3] A. Brunn and U. Weidner, *Extracting buildings from digital surface models*, in 'ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects', Stuttgart, Germany, pp. 27-34 (1997).
[4] R. Collins, A. Hanson, E. Riseman and H. Schultz, *Automatic extraction of buildings and terrain from aerial images*, in A. Gruen, O. Kuebler & P. Agouris, eds, 'Automatic Extraction of Man-Made Objects from Aerial and Space Images', Birkhauser Verlag, Basel, Boston, Berlin, pp. 169-178 (1995).
[5] E. Ebner, W. Eckstein, C. Heipke and H. Mayer, *Automatic Extraction of GIS Objects From Digital Imagery*, Vol. 32, B3-2W5, ISPRS, Munich, Germany (1999).
[6] P. Fua and A. Haanson, *Extracting generic shapes using model-driven optimization*, in 'DAPRA Image Understanding Workshop', Morgan Kaufmann Publishers, Cambridge, Massachusetts, PP. 994-1004 (1988).
[7] A. Gruen, E. Baltsavias and O. Henricson, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Vol. II, Birkhauser Verlag, Basel, Boston, Berlin (1997).
[8] C. Jaynes, A. Hanson, E. Riseman, *Model-based surface recovery if building in optical and rang images, in* 'Workshop on Semantic Modelling for the Acquisition of Topographic Information from Images and Maps, SMATI', pp. 212-227 (1997).
[9] B. Ameri, D. Fritsch, *3-D reconstruction of polyhedral-like building models,* Archive for photogrammetry and remote sensing, Vol 32, Part 3-2W5', Munich, Germany, pp.15-20 (1999).
[10] N. Haala and C. Brenner, *Extraction of buildings and trees in urban environments*, ISPRS Journal of Photogrammetry and Remote Sensing 54(2-3), 130-137 (1999).
[11] M. J. P. M Lemmens, *Structure-Based Edge Detection*, Ph.D. thesis, Technical University of Delft (1996).
[12] H. G. Maas, *Closed solutions for the determination of parametric building models from invariant moments of airborne laser scanner data*, in "IAPRS, Vol 32, Part W5", Munich, Germany, pp. 193-199 (1999).
[13] N. Haala, K.-H. Anders, *Acquisition of 3D urban models by analysis of aerial images, digital surface models and existing 2D building information, in* 'SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III', Orlando, Florida, pp. 212-222 (1997).
[14] S. Nebiker, A. Carosio, *Automatic extraction and Structuring of objects from scanned topographic maps-an alternative to the extraction from aerial and space images*, in A. Gruen, O. Kuebler & P. Agouris, eds, 'Automatic Extraction of Man-Made Objects from Aerial and Space Images', Birkhauser Verlag, Basel, Boston, Berlin (1995),.
[15] P. Axelsson, Interactive 3D extensions of 2D map data using mono images, in : ISPRS Workshop on 3D Reconstruction and Modelling of Topographic Objects", Stuttgart, Germany, pp. 145-150 (1997).
[16] F. Lang, W. Forstner, *3D-City Modelling with a digital one-eye stereo system*, in "IAPRS, Vol XXXI, Part 4", Wien (1996)
[17] F. Samadzadegan, *Automatic 3D object recognition and reconstruction based on artificial intelligence and information fusion concepts*, Ph.D. thesis, University of Tehran (2002).
[18] K. B. Atkinson, *Close Range Photogrammetry and Machine Vision.* Wittles Publishing (1996).
[19] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Addison-Wesley Publishing (1993).
[20] R. Gonzales and R. Woods, *Digital image processing*, Addison-Wesley Publishing, Reading, Massachusetts (1993).
[21] P. J. Besl and R. C. Jain, *Segmentation through variable order surface fitting*, IEEE Transaction of Pattern Analysis and Machine Intelligence, 10(2), 167-192 (1988).
[22] H. J. Zimmermann, *Fuzzy Set Theory*, and Its Applications., Kluwer Academic Publishers (1993).

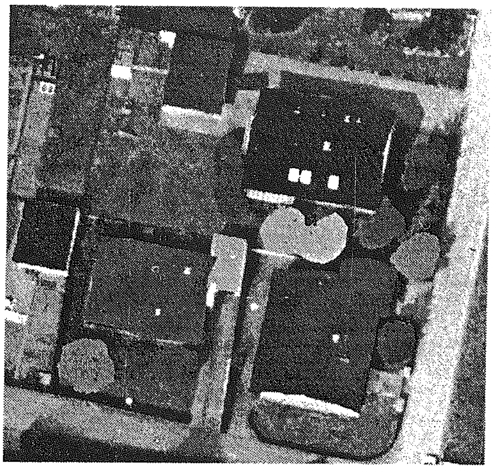


Figure (9) Final extracted regions in image space.

Fig. 10 indicates the extracted vectors for the sub-regions that are superimposed over the digital image.

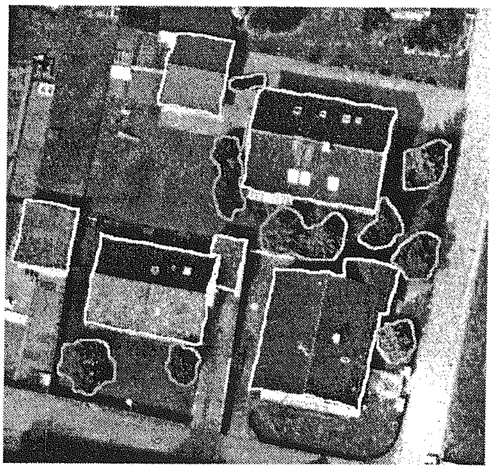


Figure (10) Extracted boundaries.

Conclusion

We believe the foregoing sections and the presented test results have demonstrated a promising and comprehensive solution to a complicated problem and the evaluation of our OE method has indicated its high potentials for a multi layer extraction of the 3D GIS objects. It should be emphasized, however, that in the preceding sections the main intention was to express the general structure of the proposed OE strategy. The principle feature of this strategy is not so much its individual modules that perform different tasks, but the methodology itself that governs the entire system. Our methodology is based on these premises: (1) Simultaneous fusion of all available information for the object extraction for a more reliable extraction operation is considered to be necessary. In our case the extracted multi layer information were limited to the three RTS components. However, they can be extended to include other possible descriptive attributes if they are available. (2) Because of the fuzzy behaviour of the objects, a rigorous and crisp modelling approach for extraction is avoided.

Reference

[1] E. Baltsavias, S. Mason and D. Stallmann, *Use of DTMs/DSMs and orthoimages to support building*

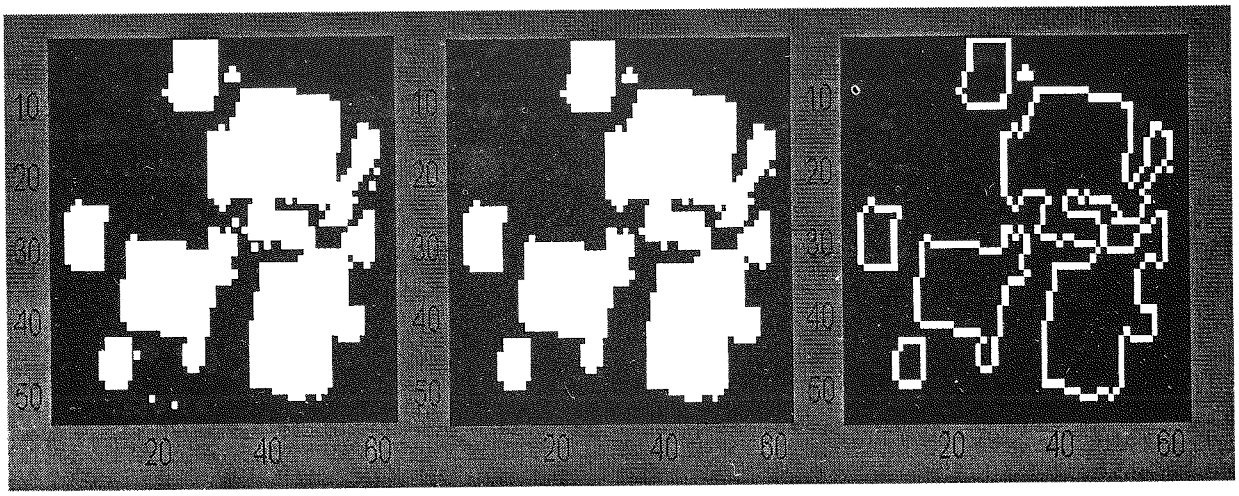


Figure (7) Initial 3D Regions, 3D Regions after refinement, Boundary of 3D Regions.

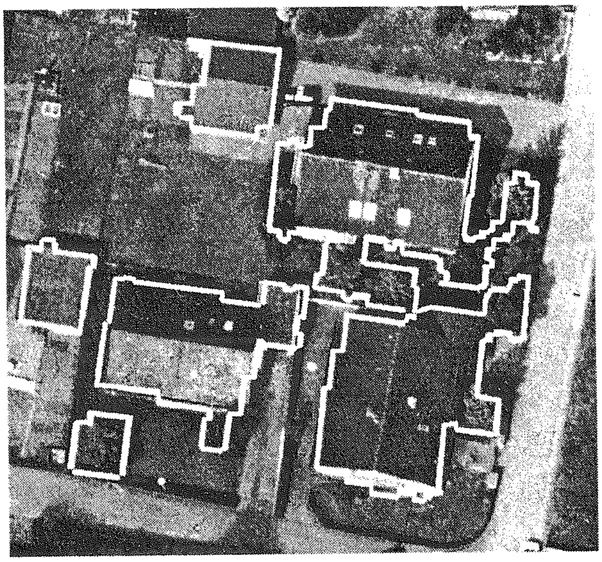


Figure (8) The Superimposed boundaries of the extracted objects in image space.

Fig. 8 demonstrates the superimposed boundaries of the objects that are transformed into the image space. Within the 2D regions in the image space, a radial search from the regions center of gravity is initiated. The fuzzy based region growing process is conducted by means of the parameters K and H, given by Eq. 3, evaluated both in image and object spaces. As already mentioned, it is assumed that a single object should demonstrate a homogenous variation of textural information. The same assumption is also extended to the relief variation in object space. That is a single object should also expose a uniform height variation. If an abrupt change in texture or height variation is encountered, it has been taken as a signal showing the presence of a new object.

Fig. 9 presents the final extracted regions in image spaces. As Fig. 9 shows our fuzzy reasoning strategy has successfully identified the presence of sub-regions within the initial regions and hence the 2D regions are subdivided accordingly to separate segments. In order to make the multi layer information consistent, the corresponding boundaries in the object space is derived by means of space intersection and therefore 3D boundaries are also modified.

Fig. 7 shows the initial segmented, refined and the extracted boundaries after the clean-up operation. As Fig. 7 shows due to the high proximity of the objects (i.e. building, car and tree), these features are not distinct as individual 3D objects. This example clearly demonstrates the incapability of single layer information for 3D object segmentation.

This defect is compensated by incorporating the relevant information in image space. The corresponding areas in image space for the 3D objects are determined by the inverse solution of collinearity condition equations given by:

$$x = -f\frac{a_1(X-X_0)+b_1(Y-Y_0)+c_1(Z-Z_0)}{a_3(X-X_0)+b_3(Y-Y_0)+c_3(Z-Z_0)}$$

$$y = -f\frac{a_2(X-X_0)+b_2(Y-Y_0)+c_2(Z-Z_0)}{a_3(X-X_0)+b_3(Y-Y_0)+c_3(Z-Z_0)} \qquad (4)$$

where X,Y,Z denote the ground coordinates of the 3D object points; x,y are their corresponding image coordinates; X_0,Y_0,Z_0 are the ground coordinates of the camera projection center; $a_1,\ldots,c_3$ are the direction cosines expressing the image coordinate orientation with respect to the object coordinate system; and f is the calibrated focal length of the camera.

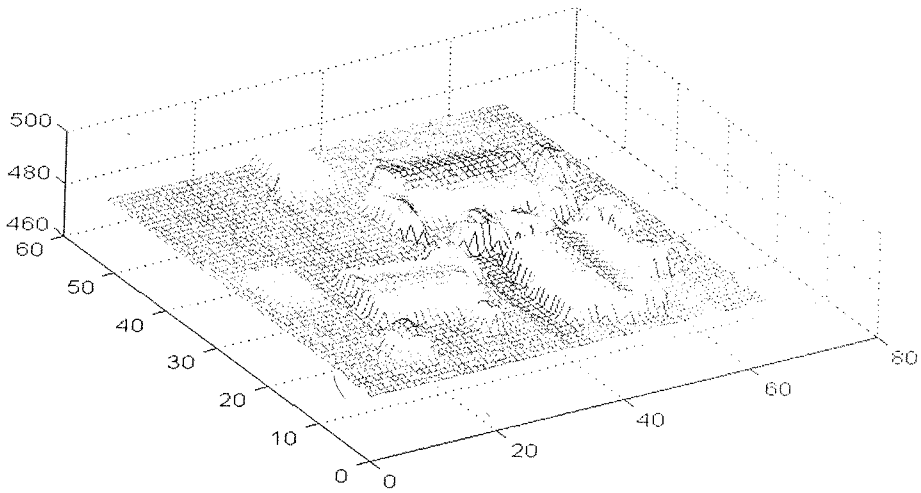


Figure (6) The automatically generated DSM for the image patch.

To segment the individual 3D objects the morphological operators, given in Table 2., are applied.

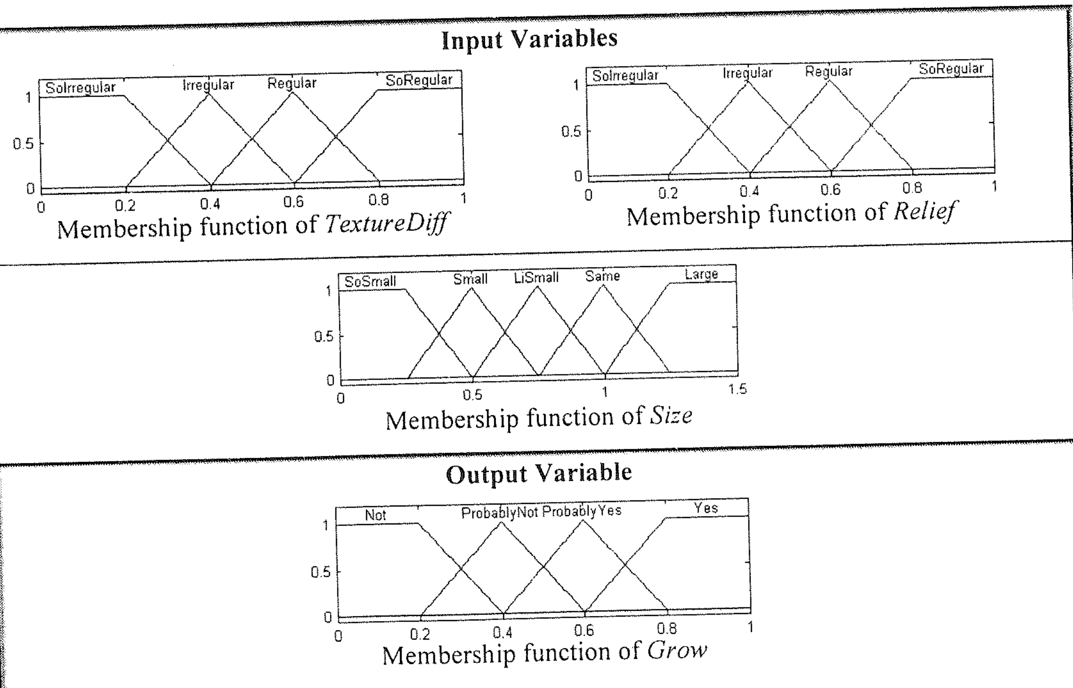
Table (2) Morphological operators used for the region extraction.

Opening	Closing
$A \circ B = (A \ominus B) \oplus B$	$A \bullet B = (A \oplus B) \ominus B$
Top-Hat	Boundary Extraction
$A * B = A - (A \circ B)$	$\beta(A) = A - (A \ominus B)$
Region Filling	
$X_k = (X_{k-1} \oplus B) \cap A^c \quad \therefore k = 1,2,\ldots$ and $X_0 = p$	

output membership functions and the corresponding rules for the three parameters of texture, relief and size of the region.

Operational stages

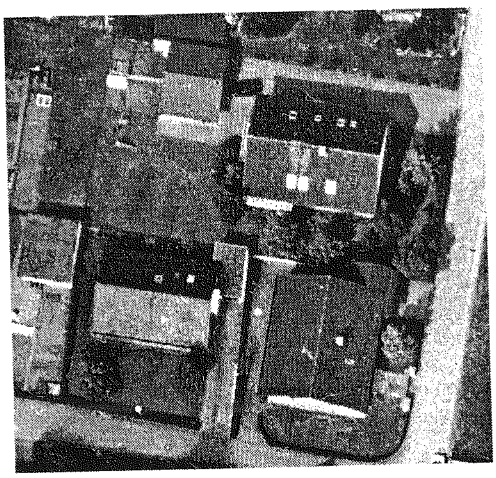
To expose different operational stages, a sub section of the digital image is selected (Fig. 5). The process was initiated with automatic DSM generation. The overall view of the generated DSM is given in Fig. 6.



Some sample rules:

IF *Area* Is *So Small* **AND** *TextureDiff* Is *So Small* **AND** *CollorDiff* Is *SoSmall* **THEN** *Grow*
IF *Area* Is *So Small* **AND** *TextureDiff* Is *Small* **AND** *CollorDiff* Is *Small* **THEN** *Grow*
IF *Area* Is *Small* **AND** *TextureDiff* Is *Small* **AND** *CollorDiff* Is *Small* **THEN** *ProbablyGrow*
IF *Area* Is *Large* **AND** *TextureDiff* Is *SoLarge* **AND** *CollorDiff* Is *SoLargel* **THEN** *NotGrow*
IF *Area* Is Large **AND** TextureDiff Is Big **AND** CollorDiff Is Big **THEN** ProbablyNotGrow

Figure (4) Membership functions and some sample rules for region growing process.



Figeur (5) The image patch used for the OE evaluation.

based region growing in the image space. After the fuzzy based consistency check, the region is subdivided into two uniformly varying sections and subsequently will be treated as two different objects. This object subdivision in image space immediately demands the corresponding 3D structural modification in the object space. This is achieved by a forward projection of the newly generated boundaries into the object space.

Evaluation of the proposed OE strategy

To realize the proposed OE method, a complete software package system was designed and implemented. It is important to mention that the system enjoys a modular design, i.e. within the general scope of the implemented methodology, individual OE modules such as region growing, relief and textural analysis, etc. can be improved parallel with the related future algorithmic developments. That is, while the general implemented strategy remains unchanged, individual tasks may be performed by a variety of other existing and yet to come algorithms.

To assess the capabilities of the proposed OE method a sample pair of scanned colour aerial photographs of an urban area in the city of Engen (Germany) was selected (Fig. 3). The selected area was suitable for the evaluation of the proposed OE method because the required complexities (e.g. proximities of different objects: building, car and tree) were available in the image. The test was conducted for the multi layer extraction of three different objects of buildings, trees and cars (BTC).

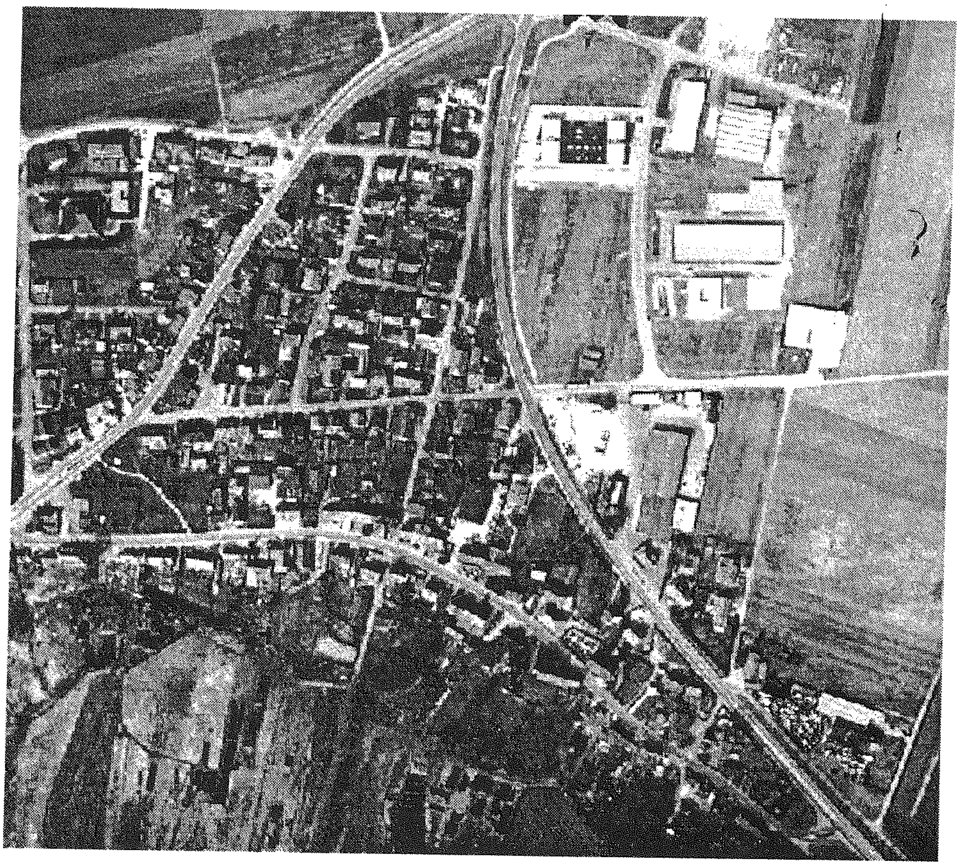


Figure (3) Digital aerial image of the test data set.

System setup

Before the system operation is started it is necessary to set up the fuzzy reasoning parameters. For the fuzzy-based region growing operations the membership functions and rules are determined using an experienced human operator. Fig. 4 shows the input and the

The output of this process will be binary data with the values one and zero denoting the 3D objects and the background respectively. This stage is then followed by the binary cleaning and the opening morphological operators. In this way only the objects of interest will remain and insignificant objects and artefacts are excluded from the extracted regions.

Object Extraction Based on Simultaneous Fusion of RTS Information

The 2D region extraction process is started by an inverse projection of the extracted 3D regions into the multi spectral image space using the pre computed values of the exterior orientation parameters of the imaging sensors. This upward projection generates preliminary 2D regions in all spectral bands. As mentioned earlier, these 2D regions may require boundary modification operations. Thus, it is necessary to perform a consistency check. It is assumed that a region that belongs to a single object demonstrates a uniform variation of the spectral values for all pixels included in the region. For example for a 2D region that belongs to a tree, the fluctuation of the values of the spectral components of the green colour should remain relatively uniform for all pixels on the region. This means that if the spectral variation exceeds a certain level, the possibility of the presence of a second object in the region is signalled. Therefore, the region's texture can be regarded as an indicator for the consistency evaluation. The textural analysis for a 2D region is carried out according to the values of: H and K denoting the mean and the Gaussian curvatures [21]:

$$H = \frac{I_{uu} + I_{vv} + I_{uu}I_u^2 - 2I_uI_vI_{uv}}{2(1 + I_u^2 + I_v^2)^{3/2}} \quad , \quad K = \frac{I_{uu}I_{vv} - I_{uv}^2}{(1 + I_u^2 + I_v^2)^2} \tag{3}$$

where $I_u, I_v, I_{uu}, I_{vv}, I_{uv}$ are differential derivative of the intensity values.

Taking into account the complexities and the fuzziness behaviour associated with these consistency checks, a fuzzy based region growing approach is adapted as follows: The region growing starts from the pixel located in the centre of the gravity of the 2D regions. For all neighbouring pixels, based on a fuzzy reasoning strategy and Mamdani inference type, a consistency check is carried out [22]. The linguistic variables to be fed into the fuzzy reasoning module are: (1) the pixels fluctuation values in all spectral components represented by the computed values of H and K, (2) the values of H and K as the relief descriptors and (3) the size of the region. The first two items contribute to the spectral and relief consistency checks respectively. That is, if for all spectral bands the grey scale variation of a pixel and its corresponding relief variations in object space is within a predefined value it is assumed that the encountered pixel is possibly consistent with the present state of the region. The third item is used to exclude the objects that are smaller than a predefined size. By the region growing process the regions undergo one of the following steps: (a) the region remains unchanged if it satisfies the consistency criteria, (b) the region is subdivided into two or more regions if consistency criteria are not satisfied, (c) different regions are merged if they are consistent. The linguistic variables and labels for the fuzzy region growing process are given in Table 1.

Table (1) Linguistic variables and labels for the fuzzy region growing.

		Linguistic Variable	Linguistic Labels
Input	Dimension	Area	*SoSmall , Small , Medium , Large , SoLarge*
	Texture	TextureDiff	*Small , Medium , Large*
	Color	ColorDiff	*Small , Medium , Large*
Output	Grow	Grow	*NotGrow , ProbablyNotGrow , ProbablyGrow , Grow*

As an example, suppose that a building with its nearby tree in the object space is extracted and classified as a single 3D object. This inevitable miss classification is revised by the fuzzy

Object Extraction methodology

The proposed object extraction method is designed to perform two sequential procedures, namely: (a) the extraction of all objects of interest in the object space, irrespective of their identity, i.e., 3D region extraction, and (b) the extraction of the corresponding 2D regions in the image space by means of a region growing process. Thus, 3D object extraction is governed by the following sequential steps:

1- Preliminary 3D regions are extracted from the underlying DSM by utilizing morphological operators. The DSM is generated by an automatic image matching method. Description of our DSM generation algorithm is beyond the scope of this paper and is reported elsewhere (see [17]). The 3D regions that are extracted from the DSM may be quite close and thus morphological operators may fail to isolate them as 3D individual objects and hence they may be erroneously classified as a single 3D object. This defect is resolved by exploiting other information available in the data set. That is, the relief and textural information. To utilize these information, the extracted 3D regions are projected upward into all available spectral spaces. This leads to the generation of the preliminary 2D regions.

2- From the center of gravity of each 2D region, a region growing process starts by which a consistency check is carried out on each 2D region using all potential information content of the object: These are roughness, texture and size (RTS). The last two attributes are evaluated in image space for all available spectral spaces. The first attribute is the roughness characteristics which is extracted from the 3D regions generated in the previous stage. Our region growing operation uses a fuzzy reasoning scheme that incorporates RTS attributes. It is assumed that a single object should somehow demonstrate a uniform texture and roughness. Using these attributes the fuzzy region growing process performs a consistency check in the region. In this way a boundary modification can be conducted for the already extracted objects. If a new sub region in a 2D segment is detected, the corresponding boundary information is transferred into the object space and the respective 3D object is split accordingly. If on the other hand, two individual objects during the region growing process are found to have similar textural and roughness characteristics, they are merged to form a single object.

The following sections expose more details regarding the proposed methodology.

3D regions extraction based on morphological operators

Having produced a general DSM for the area under investigation, it is now possible to find the individual 3D regions for each object. In this study we have adapted morphological operators for extraction and refinement of the 3D regions [20]. The adapted process may be outlined as follows: in the first step the initial regions are extracted by a top-hat morphological operator:

$$\mathrm{Re}\,gions = DSM - (DSM \circ b) \tag{1}$$

where b is the structuring element function, and $\circ$ denotes the opening operator given by:

$$f \circ b = (f \ominus b) \oplus b \tag{2}$$

$$\therefore (f \oplus b)(s,t) = \min\{f(s-x,t-y) - b(x,y) | (s-x),(t-y) \in D_f ; (x,y) \in D_b\}$$
$$\therefore (f \ominus b)(s,t) = \min\{f(s+x,t+y) - b(x,y) | (s+x),(t+y) \in D_f ; (x,y) \in D_b\}$$

where $\ominus$ and $\oplus$ denote the grey scale *Erosion* and *Dilation* operators and D_f, D_b are the domains of f and b, respectively.

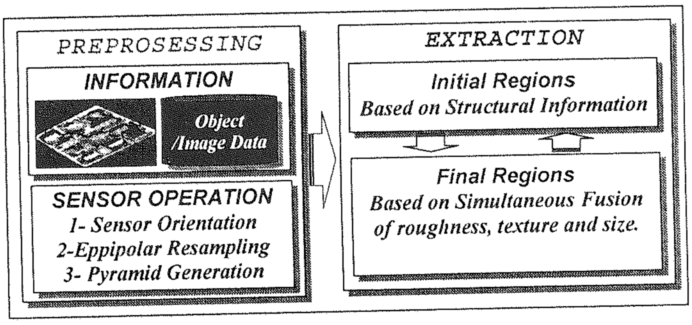


Figure (1) Extraction work flow.

Having stated the general working principle of the proposed OE method, in the following sections detailed treatments of the main individual modules that govern the OE process are presented.

Sensor Orientation

Information fusion of the segmented layers in image and object spaces are accomplished by establishing the image/object geometric relationships. To cater for majority of imaging platforms, such as: terrestrial, aerial or space borne imageries acquired by different imaging sensors, comprehensive geometric transformations are included in the OE system (see Fig. 2). The transformation equations include: the standard collinearity condition equations, direct linear transformation (DLT), projective and polynomial equations. For the satellite imageries, the transformation equations also incorporate the satellite orbital parameters into a dynamic form of collinearity condition equations. A detailed treatment of these mathematical models is given in [17].

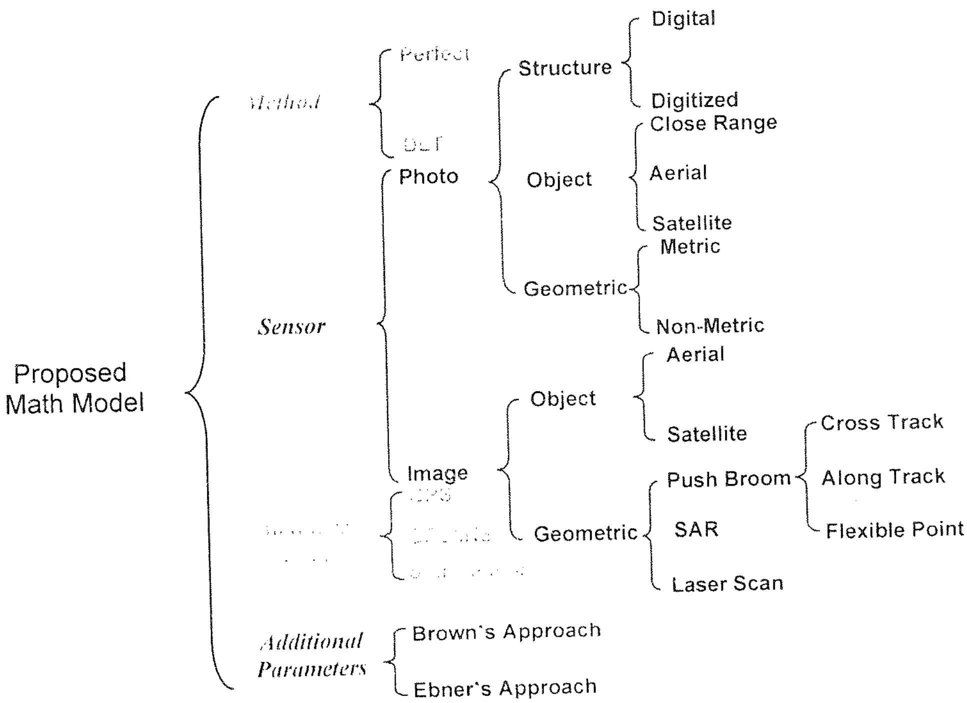


Figure (2) Geometric transformation for image/object registration.

orthographic projection of an aerial image assuming that the rectangles are the building boundaries. In other cases, however, the objects are extracted by means of the morphological operators enacted over a dense underlying topographic surface. Detection of buildings is then carried out by the subtraction of a dense grid-based digital surface model (DSM) from the computed DTM. Brunn and Weidner [3] and Ameri and Fritsh [9] may be mentioned as examples for this approach. Haala and Brenner [10] also uses DSM by incorporating height isolines to segment the digital elevation model which is equivalent to a local height thresholding approach. Lemmens [11] worked on building detection in irregularly distributed laser scanner data sets. He also applied a threshold approach. Brunn and Weidner [3] and Mass [12] use the roughness of the DSM surface measured by differential geometric quantities as an additional criterion for the discrimination of buildings and vegetation. 2D GIS data sets or existing digital map data have also been incorporated for the 3D GIS objects acquisition. These data can be used very efficiently in those regions for which objects have already been generated. 2D information extracted from the GIS data [13] or the existing digital map data [14] are either projected into image spaces [15] or DSM ([14],[10]) to delimit the corresponding regions. Some other investigators, on the other hand, preferred an interactive solution for the object extraction. The operation is performed by marking the interested regions by a boundary polygon or just a single click inside the area. This is then followed by a region growing approach for the object extraction (see for example, Lange and Forstner [16]).

To exploit more fully all available information that contribute to the extraction process, we propose an object extraction strategy which makes use of the object information inherent both in image and object spaces. In object space, by means of morphological operators, 3D objects are extracted from the automatically generated underlying digital surface model. In image space the radiometric and textural information are analyzed simultaneously with the relief variation in object space using a fuzzy-based region growing process. These information layers are fused using a rigorous object/image spaces registration scheme. Thus, effectively, the extracted objects include geometrically fused multiple layers. These extracted multi layer information provide valuable data source for those applications which perform automatic object recognition operation [17].

The overview of the proposed 3D OE methodology

The overall strategy for our proposed 3D object extraction operations may be expressed, with reference to Fig. 1, by the following interrelated procedures.

Pre-processing: To facilitate the object extraction operations, a set of pre-processing modules are initially applied to the input data. These are mainly fundamental radiometric and geometric corrections such as the grey scale filtering, histogram modification, determination of the sensor attitude and altitude parameters, epipolar resampling and the generation of the image pyramids [18],[19]. **Extraction:** In this stage a preliminary inspection is carried out to locate and extract all 3D objects that exist in the entire area covered by the image patches irrespective of the objects identity. This is achieved based on digital surface model (DSM) of the underlying area. The morphological operators are then applied to the generated DSM to delimit and isolate the individual 3D objects. In the next step for each 3D object an upward projection into the image space is performed to determine the corresponding 2D region in the image space. The final decision for each individual object's boundary is made in the image space by a fuzzy-based region growing approach. Any modification of the object boundaries as an outcome of the region growing process will result a corresponding modification in the 3D boundaries in the object space.

The Design and Implementation of a Novell Method for Automatic 3D Object Extraction in Computer Vision

F. Samadzadegan
Assistant Professor

A. Azizi
Assistant Professor

Department of Surveying and Geomatics Engineering
Faculty of Engineering, University of Tehran

C. Lucas
Professor
Department of Electrical Engineering
Faculty of Engineering, University of Tehran

Abstract

Three dimensional object extraction has been an area of major interest in computer vision for quite a long time. However, most of the existing methods for automatic object extraction employ parametric methods and hence object's fuzziness behaviour is basically neglected. These methods, thus, do not take into account the extraction complexities and may fail to reach a satisfied reliability level in complex situations. In this paper an approach for 3D object extraction (OE) is formulated which takes into account simultaneously relief variation of the object as well as its corresponding radiometric behaviour in image space. The proposed method is implemented based on the following strategy: (a) for a more reliable extraction of objects, the underlying 3D digital surface model is generated, (b) morphological operators are then applied to delimit the individual 3D objects, (c) This 3D structural information layer is accurately fused to corresponding 2D regions in image space by a rigorous geometric registration process, (d) the textural information and the size of the regions on image space as well as the roughness values for the relief variations in object space are simultaneously analyzed to modify the initially generated 2D regions. Because of the fuzzy behaviour associated with the texture, size and relief attributes, the region analysis is performed by a fuzzy based 2D region growing approach. Thus, the proposed object extraction methodology takes advantage of all object's potential information content, inherent in image and object spaces, using a fuzzy logic reasoning strategy. The proposed methodology is evaluated for the extraction of three different object classes of buildings, cars and trees, using a portion of overlapping digital aerial photos of an urban area. The visual inspection of the extracted objects demonstrates promising results.

Keywords

Extraction, Information fusion, Fuzzy logic, Morphological operators, Region growing

Introduction

The idea of having a fully automatic three-dimensional object extraction (OE) system to replace the human operator has been one of the main aspirations and the final goal for computer vision investigators [1]-[12].

The existing methods for the object extraction are mainly formulated using parametric approaches. Extraction operation by these methods normally takes place in image space using classical supervised and unsupervised techniques [7]. The extraction potentials of these methods may be greatly enhanced by incorporating additional information such as the objects structural boundaries. Joynes et al. [8] detect linear and rectangular structures in an