Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 26 (2005) 1201–1220
Pattern Recognition Letters
www.elsevier.com/locate/patrec

An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery

J.B. Mena *, J.A. Malpica

Department of Mathematics (Geodesy), Polytechnic School, Alcala University Aptdo. 20, Alcalá de Henares, E-28871 Madrid, Spain

Received 2 October 2003; received in revised form 22 October 2004
Available online 10 December 2004

Abstract

In this paper an efficient method for automatic road extraction in rural and semi-urban areas is presented. This work seeks the GIS update starting from color images and using preexisting vectorial information. As input data only the RGB bands of a satellite or aerial color image of high resolution is required. The system includes four different modules: data preprocessing; binary segmentation based on three levels of texture statistical evaluation; automatic vectorization by means of skeletal extraction; and finally a module for system evaluation. In the first module the color image is rectified and geo-referenced. The second module uses a new technique, named Texture Progressive Analysis (TPA), in order to obtain the segmented binary image. The TPA technique is developed in the evidence theory framework, and it consists in fusing information streaming from three different sources for the image. In the third module the obtained binary image is vectorized using an algorithm based on skeleton extraction techniques and morphological methods. The result is an extracted road network which is defined as a structural set of elements geometrically and topologically corrects. The fourth module is an evaluation of the procedure using a popular method. Experimental results show that this method is efficient in extracting and defining road networks from high resolution satellite and aerial imagery.

Keywords: Road extraction; Segmentation; Vectorization; Evaluation; Texture analysis; Evidence theory; Skeleton; Mathematical morphology

1. Introduction

Cartographic object extraction from digital imagery is a fundamental operation for GIS update. However the complete automation of extraction processes is still an unsolved problem. In fact,

* Corresponding author. Tel.: +91 880 40 47; fax: +34 171 15032.
E-mail address: juan.mena@uah.es (J.B. Mena).

0167-8655/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.patrec.2004.11.005

many works on this topic have been presented (Mena, 2003), but the manual intervention of the operator in extracting, defining and validating cartographic objects for GIS update is still needed.

Nevertheless, important advances have been achieved. In (Mayer, 1999 and Baltsavias et al., 2001) many of these approaches can be found. The first paper presents the state of the art on automatic object extraction techniques, where the majority of the work in this area is discussed. In the second very interesting papers on cartographic object extraction and GIS update can be consulted. Sowmya and Trinder (2000) is also very important for us because it presents a review about some of the approaches used in knowledge representation and modeling for machine vision, which is a fundamental topic in automatic extraction matter.

Recently, some automatic systems for serve as support in GIS update tasks have been developed. Between them we can find Heipke et al. (2000), where different aspects of image analysis are discussed and a framework is provided for scene interpretation which is based on the integration of image analysis and a GIS data model. This work includes two examples concerned with the combined extraction of roads and trees, and with the multitemporal interpretation and monitoring of moorland. Likewise, Eidenbenz et al. (2000) presents the project ATOMI which aims to update vector data of road centerlines and building roof outlines from 1:25000 maps, fitting in to the real landscape, improve the planimetric accuracy to 1 m and derive height information with 1–2 m accuracy. Also important for us is the work Wallace et al. (2001). It presents the research project Automatic Linear Feature Identification and Extraction (ALFIE), which uses a geographical information system built around an object oriented geospatial database. Likewise, Bückner et al. (2002) and Gerke (2002) describe the system named Geo Automatic Image Data Analyser (GeoAIDA), which have been developed at the Institute of Communication Theory and Signal Processing of Hannover, and it allows an intelligent, concise and flexible control of a scene interpretation by utilizing a semantic scene description. The system produces a hierarchic, pictorial description of the results as well as the structural context of the identified objects including the associated attributes.

However, any of previous works achieves the complete automation in the cartographic process. In fact, this problem is bigger that many researches have opted for semi-automatic extraction methods. Between them Baumgartner et al. (2002) presents a prototype system for semi-automatic extraction of road axes and a study on its efficiency for operational use. The core of this system is a road tracker based on profile matching, which is enhanced with a graphical user interface that guides the operator through the whole data acquisition process. Likewise, in Zhao et al. (2002) a semi-automatic method to create and update road maps in urban and suburban area using high resolution satellite images is proposed. In this research road mask is defined as a mask of road pixels, which are discriminated from others using a commercial remote sensing software.

Mainly focusing on GIS update is the framework presented in Ohlof et al. (2000), which is applied in many cartographic military systems of OTAN countries. This paper reports on the results of two projects conducted for the AmilGeo of the German Federal Armed Forces. The first project consists of to establish an operational workflow to update existing Vmap Level 1 data using commercially available satellite imagery. The second project is a study focused on the generation and update of Vmap Level 2 data using both satellite and airborne imagery. Also very interesting, although outside of the military environment, Bonnefon et al. (2002) presents a complete process to update and upgrade geographic linear features in GIS, such as roads, railways and little rivers, with methods as automatic as possible and with a quality evaluation.

About the phases members of the general automatic object extraction process, Markov and Napryushkin (2000) proposes a detailed and efficient sequence based on segmentation and classification for solving the remote sensing data interpretation problem in the GIS framework. Although this sequence requires manual operation in some steps, it is very important for our work. It is the following: (1) designing the list of geometrical objects to be extracted, conditional splitting the

initial raster image into n classes. (2) Selecting on the initial image the samples determining accordingly each class and specifying its features. Each sample contains the information on the certain type of objects. The set of geometric figures of each sample is the basis for marking a separate vector layer of digital thematic map in vector GIS. (3) Classification of initial raster image with use of the supervised classifier based on Bayesian decision rule. The subsequent actions are carried out for every ith sample, where $i = 1, 2, \ldots, n$. (4) Transformation of the classified image in two level raster one describing the ith class. In the obtained two level image every pixel of selected class accepts value of unity, others pixels of the image are treated as background and accept zero value. (5) Vectorization of the binary image of ith class. This phase is carried out on the basis of plane pass algorithms. As result the set of geometrical objects is generated. (6) Correction of the errors and exporting the obtained vector data into vector GIS.

In the automatic road extraction topic, another important reference also is Amini et al. (2002). In this work a new approach for automatic extraction of main roads in large scale imagemaps is proposed. The paper describes how the gray scale imagemap is converted to a simplified imagemap using morphological algorithms. The proposed method consists of two stages. In the first stage, the simplified imagemap is segmented and converted to a binary image. In the second stage, the resolution of the simplified image is reduced through the Wavelet transform and the skeleton of roads is extracted.

The support on GIS information and the use of context information are also common characteristics in many works on road extraction, including our system. In this field Ruskoné and Airault (1997) and Mayer (1999) are two of the most indicative works; and other important works also are: Agouris et al. (1998) where a road extraction method which is governed for a fuzzy system is proposed; Jeon et al. (2000) where curvilinear structures in SAR images together with digital maps are analyzed; McKeown et al. (1999a) and Fabre et al. (2001) where context information into the study of data fusion in hyperspectral image processing is used; and Hinz et al. (2001) which proposes the generation of models in the context information analysis.

Data fusion is other important aspect in our work on automatic extraction. In this matter Hellwich and Wiedemann (2000) offers an approach to the combined extraction of linear as well as surface objects from multisensor image data based on a feature and object level fusion. In this case, data sources are high resolution panchromatic digital orthoimages, multispectral image data, and interferometric SAR data. Very interesting also are McKeown et al. (1999b) and Peddle and Ferguson (2002). The last work proposes three methods for optimizing the process of data fusion, relative to the specification of user defined inputs, based on different levels of empirical testing and computational efficiency. These methods are: the exhaustive search by recursion, the isolated independent search, and the sequential dependent search. Likewise, Fabre et al. (2001) uses pixel fusion in order to elaborate a classification method at level pixel. This paper proposes a formalism of modeling of the sensor reliability to the context that leads to two methods of integration: the first one amount to integrate this further information in the fusion rule as degrees of trust, and the second models the sensor reliability directly as mass function according to evidence theory, which can be studied in Kohlas (1995) Kohlas and Besnard (1995a,b), Kohlas (1997) and Bauer (1997).

Finally, a fundamental topic in automatic object extraction also is the knowledge representation and modeling for computer vision. In this field, Mayer (1999) presents an exhaustive study and detailed analysis on automatic object extraction. This paper defines criteria in order to establish a model of knowledge. The model comprises: the derivation of characteristic properties from the function of objects, three dimensional geometry and material properties, scales and levels of abstraction/aggregation, and local and global context. The strategy consists of grouping, focusing on different scales, context based control and generation of evidence from structures of parts, and fusion of data and algorithms. Likewise Sowmya and Trinder (2000) presents a review of the approaches used in knowledge representation and modeling for machine vision, and give examples

of their applications in research for image understanding of aerial and satellite imagery. More recently Dell'aqua and Gamba (2001) offers a fuzzy approach to the analysis of airborne synthetic aperture radar images of urban environments. In particular, it shows how to implement structure extraction algorithms based on fuzzy clustering unsupervised approaches. Another reference is Andersen et al. (2002) where LIDAR sensing geometry is explicitly modeled in the domain of scan space three dimensional, analogue to two dimensional image space. Here, prior models for object configurations take the form of Markov marked point processes, where pair wise object interactions depend upon object attributes. Given the complexity of the distribution used, inferences are based upon dependent samples generated via Markov chain Monte Carlo simulation. Also in Zhang and Baltsavias (2002a,b) a concept for road network reconstruction from aerial images using knowledge based image analysis is presented. In contrast to other approach, the proposed approach uses multiple cues about the object existence, employs existing knowledge, rules an models, and treats each road subclass differently to increase success rate and reliability of the results. Finding 3D edges on the road and specially the road borders is a crucial component of their procedure and is the focus of theses papers.

1.1. Hypotheses of this work

Restricting the input image type as well as the cartographic object to extract, we limit the automatic extraction problem and, subsequently, good results could be obtained. In this line, following hypotheses and input requirements have been selected for our system on automatic extraction and GIS update:

1. *Linear cartographic objects.* Linear objects like roads or river have been selected. Therefore the system is not valid for building, trees or superficial shapes, though the segmentation phase could be applied for these objects. Specifically, our objective is the automatic road extraction, including the geometrical and topological definition of the road network.
2. *Aerial or satellite color images*. RGB bands of an aerial or satellite color image have been only selected as input data. Therefore, in our system the multitemporal, multispectral or stereoscopic analysis are excluded. Although these techniques improving the quality of results, we want to offer a simple system on road extraction which provides reasonable outputs starting from a minimum of input requirements.
3. *High resolution imagery.* An appreciable width of the roads in the input image is needed for our method, since the texture analysis is used in order to obtain a binary segmented image. Therefore Ikonos, Quick Bird and high resolution aerial imagery can be used. Landsat, Spot and low resolution aerial imagery are excluded.
4. *Rural and semi-urban areas.* In this work only rural and semi-urban areas are considered, because in urban areas frequently the texture analysis of the ground can not be applied.
5. *Automatic system.* By means of our system reasonable outputs are obtained when the algorithm is applied starting from images of mentioned characteristics.

1.2. Work sequence

For solving the automatic road extraction problem, the following flow of work has been accepted for us based on Ohlhof et al. (2000); Markov and Napryushkin (2000); Wallace et al. (2001) and Amini et al. (2002):

1. Data preprocessing.
2. Object extraction and classification.
3. Raster vector conversion.
4. Validation of results.
5. Assignment of attributes.
6. Update of existing vector data.

In this line, we proposes a new method for automatic road extraction and GIS update which starting from a minimum of input data provides the geometrical and topological definition of the road network from high resolution imagery.

2. Objectives and modules of the automatic road extraction system

In order to expedite the integration of our algorithm in other object extraction existing systems, the following goals have been chosen:

1. Acquisition of initial GIS information and image preprocessing.
2. Automatic road network detection by means of segmentation based on texture progressive analysis.
3. Geometrical and topological definition of road network based on skeleton extraction methods.
4. Evaluation, validation and storage in GIS of graphic elements which have been obtained.

These objectives are achieved by means of our algorithm which comprises the following modules according to different levels of knowledge:

1. Data preprocessing.
2. Low level: segmentation.
3. Mid level: vectorization.
4. High level: evaluation and validation.

3. First module: Data preprocessing

As external input data, only the RGB bands of a high resolution image is needed for our algorithm. In this paper an Ikonos image in order to clarify the system description have been chosen. In (Dial et al., 2001) several aspects about the use of Ikonos imagery in automated road extraction are analyzed.

3.1. Orthoimage

The first task consists in obtaining the rectified and georeferenced image. The rectification process is indispensable because it avoids the accumulation of errors when big surfaces are analyzed. However the inclusion of geodesic coordinates can be later applied. In fact, in this paper only image coordinates are used.

3.2. Reference cartographic elements

Since updating the GIS is the objective of our system, preexisting roads into the vector GIS database is supposed. Between them, those which correspond with the region in the image represented are selected in order to serve as training area for the segmentation phase.

It also is possible the automatic selection of the training set by means of unsupervised classification and result validation (Ruskoné and Airault, 1997). Nevertheless, we have opted for considering the GIS stored information, since using GIS database the validation process of the training area can be omitted. In fact, the literature presents some algorithms which aim at correcting the error due to noise and generalization in cartographic registers of databases. Examples of them can be found in (Rellier et al., 2002; Willrich, 2002). Subsequently, a minimum of accuracy in stored data could be supposed.

3.3. Median filter

In order to smooth the image noise, a median filter is applied considering the window 3×3 for each pixel (Heijden, 1995). This operation could be substituted for the filtering technique in (Bhattacharya et al., 2001) described, which is based on texture properties. In this technique the texture unit comprising of eight neighborhood elements is decomposed into two separable texture units, namely cross texture unit and diagonal texture unit, of four elements each. For each pixel cross and diagonal texture matrix is evaluated using several types of combinations of cross and diagonal texture units. Using the median technique with 3×3 window best result in the reduction of noise in satellite data have been obtained.

Fig. 1 shows the initial orthoimage after applying the median filter, as well as the training set which have been taken from the GIS database.

Fig. 1. Original orthoimage and reference cartographic elements.

4. Second module: Segmentation process (low level of knowledge)

Images segmentation represents a first step in many tasks that pattern recognition or computer vision has to deal with. There are many papers about segmentation of images using color, see Skarbek and Koschan (1994) for an early survey and Cheng et al. (2001) for a more recent one. Several authors are applying different techniques for color in order to improve the final result of the segmentation, for example, Park et al. (1998) presents a new algorithm based in mathematical morphology which performs a clustering in 3D color space; Zugaj and Lattuati (1998) proposes the fusion of several characteristics such as region and edges; fuzzy techniques are applied in (Moghaddamzadeh and Bourbakis, 1997; Yang et al., 2002); and Markov Random Fields (MRF) are used for clustering in (Jayanta, 2002). Of the all possibilities for studying segmentation of color images, we are going to focus in this paper on color texture, since it fits the profile for dealing with context information so important in the psychological characteristics of objects image recognition. Several authors have also studied color textures; among them Song et al. (1996) who have obtained the fractal dimension of color texture with the correlation among the color bands for feature extraction. Likewise, Conci and Proenca (1997) have also used color and fractal dimension for color texture segmentation. Krishnamoorthi and Bhattacharyya (1997) have put forward an orthogonal polynomial basic color texture model for the unsupervised segmentation of images with interesting results. The algorithms for image segmentation using color have been used in wide range of applications such as: clothing (Chang and Wang, 1996); automated surveillance (Paschos and Valavanis, 1999); images retrieval from a large database (Belongie et al., 1998; Zhong and Jain, 2000); classification of weed species (Burks et al., 2000a,b). In the latter case 93% accuracy is achieved by using only saturation and hue.

In this paper color texture means using the interweaving of color information in the three bands with different order statistics and color coordinate systems. In order to obtain the greatest amount of information from the different order statistics we have relied on the theory of evidence (Shafer, 1976). This theory has been used as a fusion technique.

This module is divided in several sections. Thus, Section 4.1 deals with different order statistics. Section 4.2 dwells on the relationship between the RGB and HSI systems, and how order statistics is applied. The theory of evidence will be introduced in Section 4.3 to fuse the information coming from the three sources of information in the same image.

4.1. The sources

Mathematically, each image can be thought of as a set of points in a three dimensional euclidean space, where each pixel $\boldsymbol{x}$ is represented as a vector. Consequently, in this representation the vector's three coordinates could be its RGB or HSI corresponding values.

The training set is represented in Fig. 1 by several pixels under the area of interest. Let us call $(\boldsymbol{m}_t, \Sigma_t)$ the mean and covariance matrix of the training set, respectively.

In order to segment the image three order statistics have been used: in the first order statistic Mahalanobis distance between pixels for detection and the training set are used. Therefore this phase is based on the comparison between a vector

(pixel) and distribution of vectors (training set). In the second step the multidimensional distribution of the three bands in a window centred in each pixel is used. In this phase each pixel is not yet associated with a vector, but with its corresponding distribution ($\boldsymbol{m}_x$, Σ_x) which is compared with the distribution ($\boldsymbol{m}_t$, Σ_t) of the training area. Now the comparison is accomplished by using both distributions and therefore the Mahalanobis distance can not be yet used. In its place, the Battacharyya distance is applied. This step can be classified as one and a half order statistic. Finally, in the third step second order statistic is applied using cooccurrence matrices in order to analyze the color relations that interweaves the three bands of the image. Here, the different color levels between each pair of adjacent pixels is considered. Then a distribution for each pixel is built which is compared with the corresponding to the training set similarly to the previous step.

4.1.1. Source 1: First order statistics

Since the pixels are also points in a three dimensional space, the simplest way of classifying pixels in the image, whether it belongs or not to the area of interest, would be to calculate the distance between the pixels for classification and the training pixels. To see how far the new pixel is from the pixels in the area of interest it would be necessary to see how far the pixel in study $\boldsymbol{x}$ is from $\boldsymbol{m}_t$ of the training set. The Mahalanobis distance d (Fukunaga, 1990) will be used instead of the euclidean distance, since the former gives a better approximation thanks to the introduction of the covariance matrix. Its equation is:

$$d = \sqrt{(\boldsymbol{m}_t - \boldsymbol{x})^t \Sigma_t^{-1} (\boldsymbol{m}_t - \boldsymbol{x})} \tag{1}$$

where $\boldsymbol{x}$ represents the studied pixel and ($\boldsymbol{m}_t$, Σ_t) are the mean and covariance matrix of the training set. The result of applying the Mahalanobis distance to road detection in the image from figure is presented in Fig. 3a. This last image represents the values obtained from the Mahalanobis distance, which have been normalized on a scale from 0 to 255. The brighter the pixel the closer it will be to the training texture. In this case, several pixels from the interest area were taken as training pixels.

4.1.2. Source 2: One and a half order statistics

Instead of comparing ($\boldsymbol{m}_t$, Σ_t) to isolated pixels, as it was done in the first order from above, the group of pixels in the training set from this order statistic is going to be compared to the group of pixels around the pixel $\boldsymbol{x}$ currently analyzed. This fact is going to be established by comparing the distribution of pixels in the training area, with the distribution of the pixels in an area around the pixel $\boldsymbol{x}$. In order to do that 5×5 window has been selected around the pixel $\boldsymbol{x}$, which allows to calculate ($\boldsymbol{m}_x$, Σ_x) mean and covariance matrix for the distribution of the pixels in the window. We have called this process one and a half order statistics, since it is going to represent an intermediate state between the image obtained with the Mahalanobis distance and the one that it is going to be obtained with the cube texture later on. If distribution of pixels in the window are close enough to the distribution of pixels in the training set then the central pixel in the window could be considered as belonging to the same group of the pixels in the training set. These distributions are compared by using the Bhattacharyya distance b (Fukunaga, 1990):

$$b = \frac{1}{8}(\boldsymbol{m}_t - \boldsymbol{m}_x)^t \left(\frac{\Sigma_t + \Sigma_x}{2}\right)^{-1} (\boldsymbol{m}_t - \boldsymbol{m}_x) + \frac{1}{2}\log \frac{\left|\frac{\Sigma_t + \Sigma_x}{2}\right|}{\sqrt{|\Sigma_t||\Sigma_x|}} \tag{2}$$

where Σ_x is the covariance matrix of the group of pixels in the window drawn around the pixel in study $\boldsymbol{x}$. The result of applying the distance between distributions, using function $\exp(-b)$, to the real image of Fig. 1 is presented in Fig. 3b.

4.1.3. Source 3: Second order statistics, the texture cube

Pixels that are close together tend to be more related than pixels that are far away from each other. Statistics for pairs of pixels could be studied through many different techniques. We have sought the one that provides the greatest amount of information with the minor algorithm complexity in terms of computing time.

The technique of cooccurrence matrices has proven its possibilities in many practical applications of textures. A long time ago Weszka et al. (1976) showed that cooccurrence matrices gave better results than spatial frequency methods for ground classification. Gagalowicz (1987) has demonstrated the power of cooccurrence matrices for synthetic textures. He has also studied color and third order statistics, obtaining very complicated synthetic textures although it requires a great amount of computing times for processing the images.

In order to enhance robustness of the method, only the pixels that are less than three standard deviations from the mean were considered for the calculation of the mean and covariance matrix of the training set. This iterative process was thought to achieve the correct depuration of training set distribution, and it is based on criteria of statistical tolerance. In fact, some pixels of the training set could not properly be set to represent the texture for some reasons, thus corrupting the corresponding distribution. To remedy this, a value between 2 and 3 standard deviation are chosen, as statistics text books recommend. However, in our case an exclusion of excessive number of candidate pixels was not feasible because the distribution of training set would result too restrictive for our search of image segmentation under many texture types. Therefore, by increasing this threshold to 3 std we were able to achieve the exclusion of outliers pixels for the training set without restricting excessively the corresponding distribution.

In order to study the second order a model of color texture, similar to the one in (Mao and Jain, 1992) has been created. We have tried to preserve the greatest amount of information through out the whole process, basically by using as few parameters as possible. Actually there are only two, the size of the window and the threshold for the plausibility that will be discussed in the following section. These two parameters are fixed for all the images so there is no need of tuning them for each picture, in that sense the process is intervention free and therefore automatic.

For each pixel $\boldsymbol{x}$, a three-pixel per edge cube is considered, as the one represented in Fig. 2. This

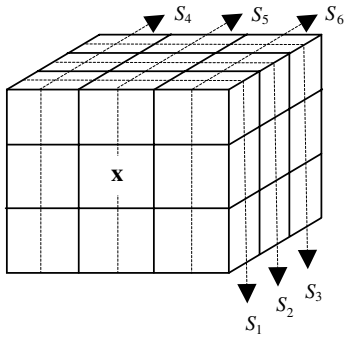

Fig. 2. The texture cube.

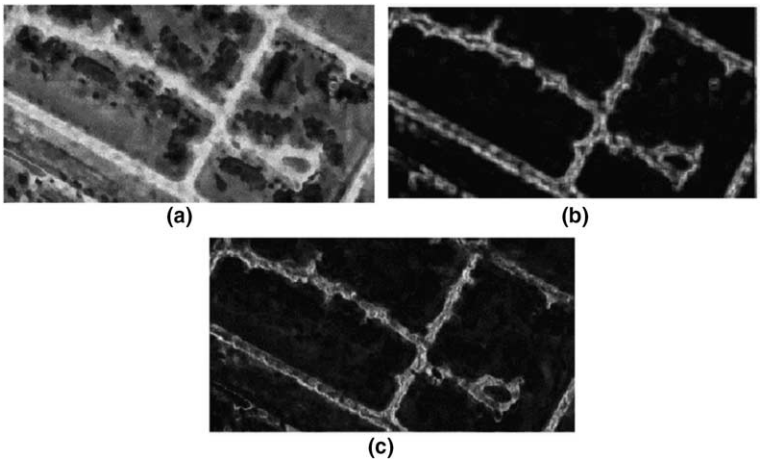

Fig. 3. (a) First order layer; (b) 1.5 order layer; (c) second order layer.

cube will be called *the cube texture*. Firstly, the three sections formed for the bands are considered, that is, each of the three sections S_k, $k = 1,2,3$. They will correspond to each of the bands in the HSI decomposition of the image, where the pixel $\boldsymbol{x}$ is in the central position of the cube. The cube texture is formed by 27 cells, and since it has been consider 256 digital levels, each of the cells have a number between 0 and 255. This was done to avoid a threshold and to retain the maximum amount of information possible.

As a second step, three new sections from the cube texture will be considered. These sections are traced transversely to the firsts, as shown in Fig. 2, and they are designed whit $k = 4,5,6$. No more sections were considered in order to avoid the undesirable redundancy, which could distort the results. Therefore there would be six sections altogether.

$$S_k = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix}; \begin{cases} k = 1,2,\ldots,6 \\ t_{ij} = 0,1,\ldots,255 \quad \forall i,j = 1,2,3 \end{cases} \tag{3}$$

For each section S_k, four cooccurrence matrices are built by empiric directional criteria. Therefore there are 24 matrices altogether. In fact, these matrices are not built directly for computing time reasons, since each matrix would need to have 65,536 elements (256×256). Besides, given that each section S_k is only of order three, the cooccurrence matrices will have a high number of zeros. To avoid the building of these matrices we propose to substitute them for cooccurrence distributions, which are promptly defined.

Let us consider section S_k of pixel $\boldsymbol{x}$ of Eq. (3). The Cooccurrence Distribution (CD) in the level of grey t_{ij} according to our directional criterion 1 is defined by considering the following succession of pairs:

$$\begin{aligned} \mathrm{CD}_{k[1]} = \{&(t_{11},t_{12}),(t_{12},t_{13}),(t_{13},t_{21}),(t_{21},t_{22}),\\ &(t_{22},t_{23}),(t_{23},t_{31}),(t_{31},t_{32}),(t_{32},t_{33})\} \end{aligned}$$

Note that the relations between the pixel grey levels belonging to section edges, in this distribution, are also considered. In this way we try to interweave ad hoc the pixel information.

According to our suggestion, the corresponding CD for directional criteria 2, 3 and 4, are respectively:

$$\begin{aligned} \mathrm{CD}_{k[2]} = \{&(t_{31},t_{21}),(t_{21},t_{32}),(t_{32},t_{11}),(t_{11},t_{22}),\\ &(t_{22},t_{33}),(t_{33},t_{12}),(t_{12},t_{23}),(t_{23},t_{13})\} \end{aligned}$$

$$\begin{aligned} \mathrm{CD}_{k[3]} = \{&(t_{11},t_{21}),(t_{21},t_{31}),(t_{31},t_{12}),(t_{12},t_{22}),\\ &(t_{22},t_{32}),(t_{32},t_{13}),(t_{13},t_{23}),(t_{23},t_{33})\} \end{aligned}$$

$$\begin{aligned} \mathrm{CD}_{k[4]} = \{&(t_{11},t_{12}),(t_{12},t_{21}),(t_{21},t_{13}),(t_{13},t_{22}),\\ &(t_{22},t_{31}),(t_{31},t_{23}),(t_{23},t_{32}),(t_{32},t_{33})\} \end{aligned}$$

Note that all distributions are always formed by eight elements, normally with frequency equal to one, that is to say, without repetitions. In this way the 65,536 values of the cooccurrence matrix have been substituted by the pairs of cooccurrence which actually happen on the levels of grey.

The features given in (Haralick, 1979) have been calculated from the 24 matrices constructed in above. These features are: correlation, energy, entropy, maximum probability, contrast, and inverse difference moment, which have been chosen ad hoc in order to have an adequate amount of data for texture analysis.

However, to avoid the problem of the sparsely in the matrices, as mentioned before, the concept of probability has been modified. This modification is needed because a same value of any feature produces, how we'll see immediately, the singularity of the covariance matrix corresponding to the considered pixel. This is done as follows:

Let the bidimensional distribution be:

$$\mathrm{CD}_{k[n]} = \{(x_1,y_1),(x_2,y_2),\ldots,(x_8,y_8)\} \quad n = 1,2,3,4 \tag{4}$$

we define empirically the variable s of element (x_i, y_i) as:

$$s_i = \frac{\sqrt{x_i y_i}}{\sum_{j=1}^{8} \sqrt{x_j y_j}} \tag{5}$$

This equation was chosen ad hoc. It minimizes the sparsely matrix problem, which is presented when using 256 grey levels, as happen when the classical probability concept is applied. Subsequently, the values for the elected Haralick features are the following:

$$\begin{aligned}
H_1 &= \frac{1}{8} \sum_i \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \text{ (correlation)} \\
H_2 &= \sum_i s_i^2 \text{ (energy)} \\
H_3 &= \sum_i s_i L s_i \text{ (enthropy)} \\
H_4 &= \mathrm{Max}(s_i) \text{ (maximum } s\text{)} \\
H_5 &= \sum_i |\, x_i - y_i \,|\, s_i \text{ (contrast)} \\
H_6 &= \sum_i \frac{s_i}{|\, x_i - y_i \,|} \text{ (inverse difference moment)}
\end{aligned} \tag{6}$$

Since four CD in six sections have been considered, a distribution of 24 vectors with dimension six (one for each Haralick feature) for each $\boldsymbol{x}$ was obtained. This distribution $(\boldsymbol{m}_x, \Sigma_x)$ was then compared to the one corresponding to the training set $(\boldsymbol{m}_t, \Sigma_t)$ through the Bhattacharyya distance (2); it culminates with the calculation of the respective value $\exp(-b)$.

If the 6-dimensional distribution of pixel $\boldsymbol{x}$ have a component equal in all their vectors, the six covariances corresponding to such variable are all nulls. Consequently, the covariance matrix becomes singular, and the Bhattacharyya distance cannot be applied. Now we can see why the concept of the new variable s has been defined.

Fig. 3c shows the layer that is obtained after applying the texture cube technique to the image in Fig. 1.

4.1.4. The RGB and HSI systems

It appears that as the order of statistics increase the closer we are to the psychological concept of color detection. Some authors are using psychological concepts in automatic segmentation in order to improve some existing methods. Such is the work done by Mirmehdi and Petrou (2000).

In the first order statistics, or first layer, the RGB system has been applied. It appears that when using HSI the results are not as good as with RGB. In the intermediate layer some results surge regardless of which system is applied, either RGB or HSI. In our method HSI has been applied. In the last layer, the results using HSI are significantly better than applying RGB, that is why the former system has been applied. After many tryouts, the following empirical yields have been obtained.

In consequence, we have estimated that HSI system is more appropriate for texture analysis as the statistic order is increased. In fact, the RGB system is only a mathematical structure based on primary colors addition. Even though the HSI is a linear transformation for each pixel, it is not linear for the entirely image. The HSI system was proposed by A.R. Smith in 1978 system, with the idea to show the intuitive characteristics of color. Subsequently, we conclude that HSI system has psychological aspects and it looks like that is more adequate for human color interpretation than RGB.

For transforming RGB into HSI the following algorithm has been used:

$$\delta = \mathrm{Max}(r,g,b) - \mathrm{Min}(r,g,b)$$

$$h = \begin{cases} 60^\circ \dfrac{g-b}{\delta} & \text{if } \mathrm{Max}(r,g,b) = r \\ 60^\circ \left(2 + \dfrac{b-r}{\delta}\right) & \text{if } \mathrm{Max}(r,g,b) = g \\ 60^\circ \left(4 + \dfrac{r-g}{\delta}\right) & \text{if } \mathrm{Max}(r,g,b) = b \end{cases}$$

$$s = \frac{\delta}{\mathrm{Max}(r,g,b)}; \quad i = \mathrm{Max}(r,g,b)$$

4.2. Fusion of the information layers with the theory of the evidence

The results obtained, with the Mahalanobis distance in the first layer and the Battacharyya distance in the second and third layers from the above, gave values that will be used for fusing of these three techniques. This means that these val-

ues can be considered pieces of evidence for the recognition of the studied texture (Shafer, 1976).

We proceed by considering only two classes in the Dempster–Shafer theory of evidence. Either a pixel belongs to the studied texture (to be detected) ω, or it belongs to the background ϖ. There is also an uncertainty θ inherent in the theory of evidence. All this constitutes the frame of discernment Θ in our case.

$$\Theta = \{\omega, \varpi, \theta\}$$

For each pixel three values of evidence, μ_i ($i = 1,2,3$), corresponding to each order statistic could be obtained

$$[\mu_i(\omega), \mu_i(\varpi), \mu_i(\theta)] \quad \text{where}$$
$$\mu_i(\omega) + \mu_i(\varpi) + \mu_i(\theta) = 1 \quad \forall i = 1,2,3. \tag{7}$$

We will see that using the theory of evidence the three groups of values are fused to yield only one group for each pixel.

4.2.1. The first order statistic

The Mahalanobis distances d (1) between pixel $\boldsymbol{x}$ and the set of training pixels are calculated. Then the maximum $d_{\max}$ and minimum $d_{\min}$ values are obtained for normalising, and computing its complement to one (d'), as the following equation shows:

$$d' = 1 - \frac{d - d_{\min}}{d_{\max} - d_{\min}} \tag{8}$$

The standard deviation for the all d' values obtained according to (8) is taken as the uncertainty $\mu_1(\theta) = \sigma$, because standard deviation is a statistical measure of dispersion which represents the accuracy of mean value in the respective distribution. Thus, in order to verify the condition of adding equal to one, Eq. (7), the values $\mu_1(\omega)$ are obtained as follows:

$$\mu_1(\omega) = d'(1 - \sigma) \tag{9}$$

Using these results the new evidence masses are obtained. These are the definitive values of evidence of each pixel of being close to the feature represented by the training set. Therefore the values for not belonging to the training set would be given by:

$$\begin{aligned}\mu_1(\varpi) &= 1 - \mu_1(\omega) - \mu_1(\theta) \\ &= (1 - d')(1 - \sigma)\end{aligned} \tag{10}$$

4.2.2. The first and a half order statistics

The Bhattacharyya distances b (2) between the training set ($\boldsymbol{m}_t$, Σ_t) and each pixel ($\boldsymbol{m}_x$, Σ_x) are computed. Then the evidence mass ω will be obtained by Eq. (11), inspired in the *Jeffries Matusita* distance (Richards and Jia, 1999) between a pair of probability distributions:

$$d' = \frac{1}{e^b} \tag{11}$$

where b is the Bhattacharyya distance.

Once all d' values for the image have been calculated, the maximum and minimum are obtained following similar steps as in above for the previous order. The standard deviation for the distances values was assigned to the uncertainty $\mu_2(\theta) = \sigma$, and the evidence masses underwent similar transformation as given by (9) and (10).

4.2.3. The second order statistics

The second order statistics is calculated very similarly to the 1.5 order. Again the Bhattacharyya distance is used, and the evidence mass will be obtained by Eqs. (9) and (10).

	$\mu_j(\omega)$	$\mu_j(\varpi)$	$\mu_j(\theta)$
$\mu_i(\omega)$	ω		ω
$\mu_i(\varpi)$		ϖ	ϖ
$\mu_i(\theta)$	ω	ϖ	θ

4.2.4. The orthogonal sum

With the values $\mu_i(\omega)$, $\mu_i\varpi$ and $\mu_i(\theta)$ for the three orders, $i = 1,2,3$, the Dempster's rule or orthogonal sum of the theory of evidence is ap-

plied according to Richards and Jia (1999) and Shafer (1976). For understanding this concept we are going to consider the frame of discernment given above. The orthogonal sum of two triads of evidence masses for a same pixel, which have been obtained from different sources, is performed by constructing a unit square. This square is partitioned vertically in proportion to the mass distribution from one source (μ_i) and horizontally in proportion to the mass distribution from the other source (μ_j). In order to compute the orthogonal sum $\mu_i \oplus \mu_j$ the rectangle areas where ω appears are summed, because only them contribute to give evidence for the hypothesis ω. A similar process is followed for ϖ and the uncertainty θ.

This intuitive idea with unit square for the orthogonal sum $\mu_i \oplus \mu_j$ is defined algebraically as follows:

$$(\mu_i \oplus \mu_j)(\omega) = \frac{\mu_i(\omega)\mu_j(\omega) + \mu_i(\omega)\mu_j(\theta) + \mu_i(\theta)\mu_j(\omega)}{k}$$

$$(\mu_i \oplus \mu_j)(\varpi) = \frac{\mu_i(\varpi)\mu_j(\varpi) + \mu_i(\varpi)\mu_j(\theta) + \mu_i(\theta)\mu_j(\varpi)}{k}$$

$$(\mu_i \oplus \mu_j)(\theta) = \frac{\mu_i(\theta)\mu_j(\theta)}{k}$$

$$k = 1 - [\mu_i(\omega)\mu_j(\varpi) + \mu_i(\varpi)\mu_j(\theta)]$$

Fig. 4a shows the layer that is obtained after applying the orthogonal sum of the theory of evidence to the images in Fig. 3. Since the orthogonal sum is associative (Bauer, 1997) a new set of masses is obtained by fusing the information obtained from the three order statistics:

$$\mu_f(\omega), \mu_f(\varpi), \mu_f(\theta) \quad \text{where } \mu_f = \mu_1 \oplus \mu_2 \oplus \mu_3.$$

4.2.5. Plausibility layer

In order to secure the fusion result a binary image is obtained that shows which pixels belong to the texture and which do not. The concept of *plausibility* from the theory of evidence is applied. In our case, the plausibility $\Pi(x)$ for the pixel $\boldsymbol{x}$ is the sum of the masses of the evidences for labelling the pixel as belonging to the training set plus the uncertainty. In other words:

$$\Pi(x) = \mu_f(\omega) + \mu_f(\theta) = 1 - \mu_f(\varpi) \tag{12}$$

Here it is necessary to define a threshold in order to proceed with the segmentation. As mentioned and observed above, this is the second and more relevant parameter of our algorithm, the other was the window size for studying the texture. Fig. 4b represents the obtained binary image, where this threshold has been tuned manually. Also, in the results shown in this paper, this threshold has been tuned in the same way. It is

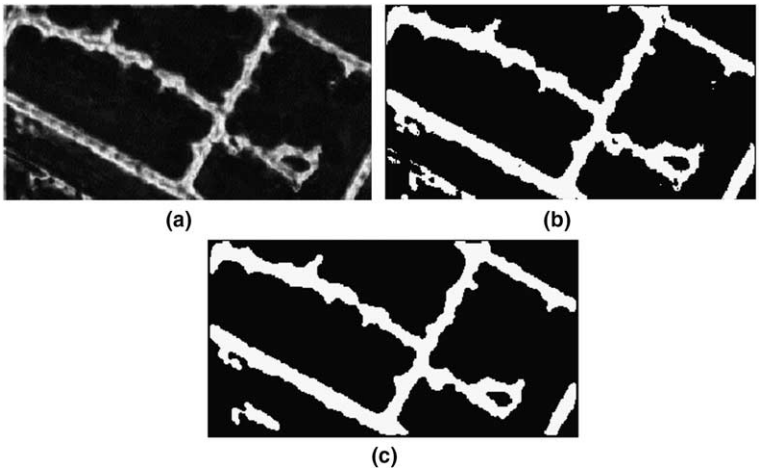

Fig. 4. (a) Fusion layer; (b) plausibility layer; (c) clean layer.

important to notice that frequently the parameter final value obtained manually is closed to the mean value of the plausibility distribution.

4.2.6. Clean process

The plausibility image is improved by an iterative clean process based on computing the numbers of black and white pixels which are included in a 3×3 window open around each pixel. According to these numbers the pixels are or not changed. The process stop when no pixel is changed. Then the final result of the segmentation is achieved. Fig. 4c shows this clean layer.

5. Third module: Geometrical and topological vectorization process (mid-level)

Starting from the segmented image, now our system seeks the geometrical and topological definition of extracted road network. For this goal an algorithm based on skeleton extraction and graph theory is applied. This algorithm comprises the following steps.

5.1. Edge generalization

In order to achieve the edge smoothness a generalization process is applied. This process is necessary because in the skeleton extraction phase only pixels which are allocate between parallel edges are used.

5.2. Main skeleton extraction

Using the method proposed in (Borgefors, 1988; Sanniti di Baja, 1994) the skeleton of the road network is automatically achieved. Fig. 5(a) shows the corresponding result.

5.3. Secondary skeleton extraction

Starting from the plausibility image, a new method for extracting the skeleton is applied. This method provides a rude result which will serve to establish the principal roads of the road network. The system is an approach presented in (Mena, 2002) which is based on the previous work proposed in (Doucette et al., 2001). It seeks to obtain the skeleton and the basic topology of the road network in the binary image by means of the progressive degeneration of Delaunay triangulation. For it, the *K*-means technique in order to stabilize the initial nodes group is used. Next the Voronoi diagram and the respective Delaunay triangulation is built. The rectification of this triangulation is made with a reiterative process which substitutes in each step the triangles for segments. When the Delaunay triangulation has not yet any triangle, the skeleton is obtained. In Fig. 5(b) the result of secondary skeleton extraction is shown.

The main advantage of this method of skeleton extraction is that it offers the main roads of the road network. This main roads could be used in the high level of knowledge in order to validate the cartographic elements obtained for the system in its last step.

5.4. Graph generation

Starting from the main skeleton a graph of relations is constructed. For it, the *sac* concept is introduced, which is defined as any set of adjacent pixels comprised between one or two nodes. The

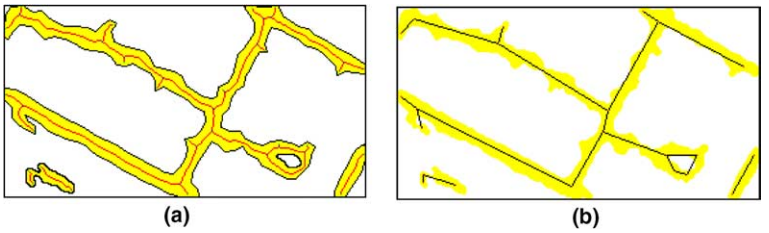

Fig. 5. (a) Main skeleton extraction; (b) secondary skeleton extraction.

sac and node concepts permit to generate cartographic objects which are considered as successions of sacs and adjacent nodes. An order relation and other of equivalence for object construction are enunciated. Likewise a semantic net according to Förstner and Pluemer (1997) is used in order to combine sacs and nodes efficiently and obtaining a minimum of curvature in the object drawing.

In order to improve the extracted vectorial structure which is stored in the graph, this phase of our system accepts additional information coming from GIS, as well as the cyclic reiteration using a neural network (Ripley, 1996). This possibility includes pruning and study of topological relations of our road network with other objects, which are achieved through models, maps, databases, etc....

5.5. Geometrical adjustment

With the entity object as work element, the algorithm studies and provides the geometrical adjustment of all objects. For it, a polynomial adjustment technique based on McGlone (1998) is applied. A polynomial whose degree coincides with the number of nodes is applied for the robust adjustment of opened objects. In the sharp objects case the adjustment is achieved by means of irregular polygons which are adapted progressively to the object. This polygonal adjustment is a reiterative process similar to the snakes technique (Kass et al., 1987), and it finalizes when the achieved accuracy is equal to image resolution.

5.6. Topological adjustment

After the geometrical adjustment, connections between different roads have been changed. In fact, the initial nodes which defined these connections have been excluded of their corresponding objects. For solving this problem a method of morphologic adjustment is applied according to Serra (1982). Now our system seeks those points which belong at the same time at two or more objects. These points are the new positions of nodes. Likewise, the initial and final points of each object are morphologically modified. Fig. 10 shows graphically the automatic result obtained.

Numerically, this result is provided by means of a file where all important characteristics like polynomial coefficients and topological information are available. Fig. 6 shows a part of the file corresponding with the graphic elements represented in Fig. 7.

6. Fourth module: Evaluation and validation of results (high level of knowledge)

6.1. Other obtained results

Variables	Fig. 8	Fig. 9	Fig. 10	Fig. 11	Fig. 12	Means
Resolution	3	2	1	2	2	2 m
Completeness	86	66	86	75	91	81%
Correctness	96	65	93	85	94	87%
Quality	83	48	81	67	87	73%
RMS	1.7	1.2	0.6	1.2	1.2	1.2 m
Redundancy	0.01	0.01	0.02	0.01	0.01	0.01
Gaps number	3	0	0	2	0	1
Gaps/km	1	0	0	1	0	0.4
Mean gap length	73	0	0	63	0	27 m

Below other results are shown. In each figure only the original image, the training area and the automatic vectorization are presented. The numerical results are omitted in order to abbreviate this paper.

```
- OBJECT n. 1
- Polynomial degree.....:          3
- Adjustment type.......:          0
- Connection number.....:          1
- Initial node..........:    13-   1
- End node..............:     2-  18
- Coefficients..........:
              1.30122896152407E+001
              2.73937205356257E-001
             -9.21918700043995E-002
              2.34193500349711E-003
- Connections...........:
        Object  14   Node   12-   6
**********************************
- OBJECT n. 2
- Polynomial degree.....:          3
- Adjustment type.......:          1
- Connection number.....:          1
- Initial node..........:     1-  69
- End node..............:   142- 144
- Coefficients..........:
              6.86616370808245E+001
              5.32838676585783E-001
              5.69911559248673E-004
             -4.29060043244976E-006
- Connections...........:
        Object  13   Node  136- 141
**********************************
- OBJECT n. 3
- Polynomial degree.....:          2
- Adjustment type.......:          0
- Connection number.....:          0
- Initial node..........:    20- 128
- End node..............:    22- 138
- Coefficients..........:
              5.38475630378092E+002
             -8.00898338075373E+000
              3.09144946732652E-002
**********************************
..................................
```

```
- OBJECT n. 8
- Polynomial degree.....:          5
- Adjustment type.......:          1
- Connection number.....:          3
- Initial node..........:   192-   4
- End node..............:   279-  49
- Coefficients..........:
              6.74863202403080E+003
             -1.62869065818306E+002
              1.55259462841762E+000
             -7.33226132036197E-003
              1.71903645964111E-005
             -1.59924605212772E-008
- Connections...........:
        Object  10   Node  206-  10
        Object  11   Node  264-  43
        Object  13   Node  202-   8
 **********************************
 ..................................

- OBJECT n. 13
- Polynomial degree.....:          6
- Adjustment type.......:          0
- Connection number.....:          6
- Initial node..........:   202-   8
- End node..............:   136- 141
- Coefficients..........:
              2.01599555626917E+002
              8.27992026824099E-002
             -8.28828830325273E-003
             -1.68477491732478E-004
              4.81493078944628E-006
             -3.88693468621847E-008
              1.04736408686212E-010
- Connections...........:
        Object   2   Node  136- 141
        Object   7   Node  151- 110
        Object   8   Node  202-   8
        Object   9   Node  195-  28
        Object  14   Node  165-  82
        Object  15   Node  158-  97
**********************************
..................................
```

```
- OBJECT n. 14
- Polynomial degree.....:          4
- Adjustment type.......:          1
- Connection number.....:          4
- Initial node..........:    12-   6
- End node..............:   165-  82
- Coefficients..........:
             -3.49507513628284E+000
              9.63658571515880E-001
             -1.52437910641889E-002
              1.37599905402965E-004
             -3.73853221543839E-007
- Connections...........:
        Object   1   Node   12-   6
        Object   5   Node   77-  30
        Object   6   Node  123-  55
        Object  13   Node  165-  82
**********************************
- OBJECT n. 15
- Polynomial degree.....:          2
- Adjustment type.......:          1
- Connection number.....:          2
- Initial node..........:   158-  97
- End node..............:   204- 115
- Coefficients..........:
              9.97241996674620E+001
             -3.18496258066717E-001
              1.93811540543648E-003
- Connections...........:
        Object  13   Node  158-  97
        Object  16   Node  204- 115
**********************************
- OBJECT n. 16
- Polygonal edges.......:          8
- Accuracy..............:          1
- Node..................:   204- 115
- Intermediate vertex...:
                           208- 122
                           221- 130
                           224- 130
                           233- 124
                           233- 113
                           213- 109
                           209- 110
- Connection............:
        Object  15   Node  204- 115
**********************************
```

Fig. 6. Numerical elements.

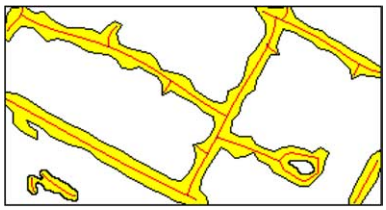

Fig. 7. Graphics elements.

6.2. Evaluation

According to the method developed in Wiedemann et al. (1998) for evaluating automatic road extraction systems, and by comparison between automatic results and manual results, the follow-

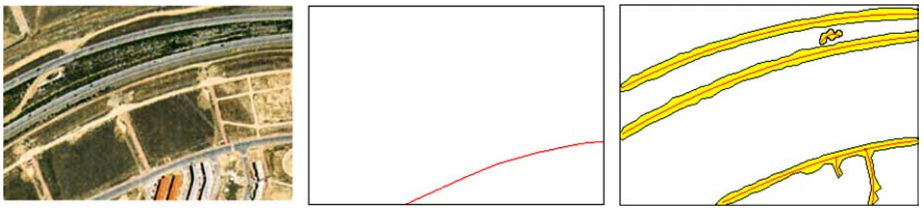

Fig. 8. Original image, training set, and graphic elements.

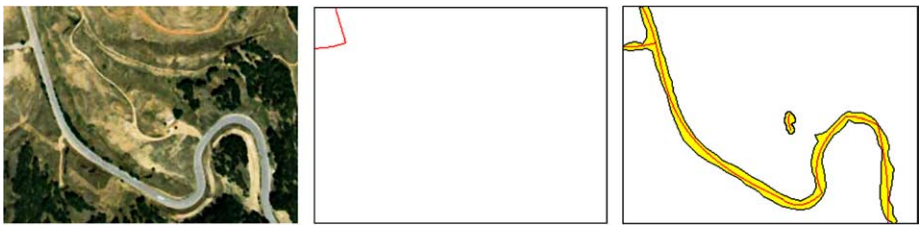

Fig. 9. Original image, training set, and graphic elements.

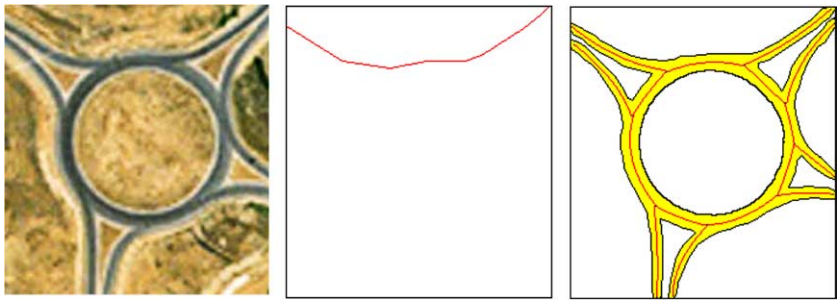

Fig. 10. Original image, training set, and graphic elements.

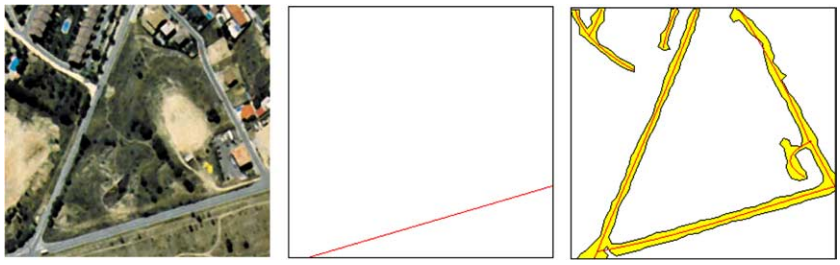

Fig. 11. Original image, training set, and graphic elements.

ing data have been obtained for our system (Figs. 8, 9, 11, and 12).

Likewise, an empirical evaluation process developed for us independently of this, has provided a quality percentage around 70%. Comparing this value whit the corresponding value obtained through the Wiedemann method, an high efficiency and quality in around seven of each ten cases can be inferred for our automatic road extraction system.

6.3. Result validation and GIS update

According to evaluation results, the validation task is often reduced. Therefore using our system for road extraction and GIS update, the manual work of operators is highly decreased. However, the conversion between image coordinates and geode-sic coordinates, including altitude data, can not be forgotten. Since we have started from a rectified and referenced image, the z coordinate can be introduced by means of a Digital Terrain Elevation Data (DTED) when it is available. If original image is rectified but it is not referenced,

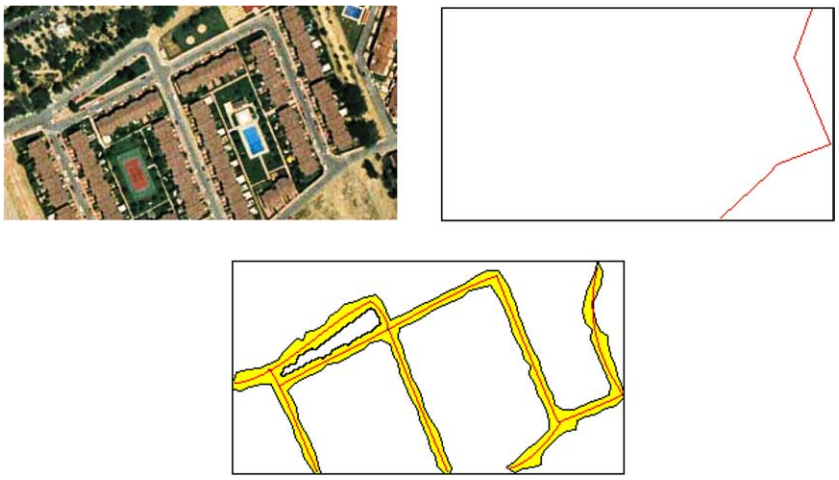

Fig. 12. Original image, training set, and graphic elements.

then a process of mathematical transform between image and terrain is needed. For it, any mathematical model of many available models (Grafarend, 2002) can be applied.

Likewise, other additional works can be needed depending of the format type which is used in the GIS database for linear object storage.

7. Conclusions

In this paper a solid system for automatic road extraction has been presented. This system includes the geometrical and topological definition of the graphic elements, and it seeks decreasing the work which is manually realized for the operators in their tasks for GIS update.

Our system constitutes a proposal of high efficiency, since admitting imperfections in the results, only GIS information and the RGB bands of a satellite or aerial color image are required for starting the automatic process.

In the segmentation process, one of the major advantages is that it needs only a few parameters (size of the window and plausibility threshold), and most of the information is retained until the very end. Nevertheless, the main feature of the process is the use of three different statistics levels for carrying out the texture progressive analysis. These levels provide complementary results which are merged by means of evidence theory. The final result shows a superior quality than the obtained by segmentation methods based only on one statistic level like maximum likelihood. This fact is because the errors obtained by one segmentation level are frequently corrected by the results of another one.

About the mentioned thresholds, the most important is the window size in the segmentation levels of order 1 and 1.5. Such size should be chosen depending on the kind of object to detect. For instance, in high resolution images road extraction and other small width lineal features demand the use of small windows (3×3 or 5×5), while shape objects extraction such vegetation or farming areas, allow the use of bigger windows. On the other hand, the window size in the segmentation level of order 2 is conditioned to the number of bands available for the image, since in this order the analysis is carried out according to the texture cube. Nevertheless, nice results are achieved when a 3×3 window is used in level 1 and 5×5 window in 1.5 for any kind of feature extraction in a three bands color image.

The plausibility threshold, that is our second parameter, can be automatically fixed giving it the arithmetic mean of all plausibility values computed, incremented in the standard deviation value. In this case the computation of the plausibility image is provided automatically, since the

quality improving was not too important when the computed threshold is manually changed.

A high stability of the process is achieved when the mentioned values for window size and plausibility threshold are accepted, and therefore the method here presented can be considered fully automatic. Besides, it can be stated as the main conclusion, that the introduction of information of high order of statistics is satisfactory for segmenting images in most of the cases.

On the other hand, the general structure of our algorithm permits the simple modification in order to integrate it with each particular environment. Likewise, the system can be improved simply introducing additional layers in the segmentation process, or refining manually the plausibility image before the vectorization. Besides, the evaluation results confirm a minimum task of manual validation. Therefore an excellent semi-automatic method for road extraction could be obtained if manual intervention along the process is assumed. Likewise, our system could be integrated with a knowledge based system and an object oriented database in order to increase the knowledge level in the validation step. However this task is very difficult and it requires future research works.

Finally, we have proposed an automatic road extraction system which is a small contribution in solving the multiple problems that GIS updating from digital imagery has associated.

References

Agouris, P., Gyftakis, S., Stefanidis, A., 1998. Using a fuzzy supervisor for object extraction within an integrated geospatial environment. International Archives of Photogrammetry and Remote Sensing 32 (Part III/1), 191–195.

Amini, J., Saradjian, M.R., Blais, J.A.R., Lucas, C., Azizi, A., 2002. Automatic road side extraction from large scale imagemaps. International Journal of Applied Earth Observation and Geo-information 4, 95–107.

Andersen, H.E., Reutebuch, S.E., Schreuder, G.F., 2002. Bayesian object recognition for the analysis of complex forest scenes in airborne laser scanner data. ISPRS. Photogrammetric Computer Vision, Graz, Austria, September 9–13 pp. A-73.

Baltsavias, E.P., Gruen, A., Gool, L.V. (Eds.), 2001. Automatic Extraction of Man-Made Objects from Aerial and Space Images (III). A.A. Balkema Publishers, UK.

Bauer, M., 1997. Approximation algorithms and decision making in the dempster—shafer theory of evidence. An empirical study. International Journal of Approximate Reasoning 17 (2–3), 217–237.

Baumgartner, A., Hinz, S., Wiedemann, C., 2002. Efficient methods and interfaces for road tracking. In: Proceedings of the ISPRS-Commision III Symposium "Photogrammetric Computer Vision" (PCV′02), Graz 2002; International Archives of Photogrammetry and Remote Sensing 34 (Part 3B), pp. 28–31.

Bhattacharya, A.K., Srivastava, P.K., Bhagat, A., 2001. A modified texture filtering technique for satellite images. In: 22nd Asian Conference on Remote Sensing, 5–9 November, Singapore.

Belongie, S., Carson, C., Greenspanand, H., Malik, J., 1998. Color and texture based image segmentation using EM and its application to content based image retrieval. In: Sixth International Conference on Computer Vision (IEEE Cat. No. 98 CH36271). Narosa Publishing House, New Delhi, India, pp. 675–682.

Bonnefon, R., Dhérété, P., Desachy, J., 2002. Geographic information system updating using remote sensing images. Pattern Recognition Letters 23 (9), 1073–1083.

Borgefors, G., 1988. Hierarchical chamfer matching: a parametric edge matching algorithm. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10.

Bückner, J., Müller, S., Pahl, M., Stahlhut, O., 2002. Semantic Interpretation of Remote Sensing Data. ISPRS. Photogrammetric Computer Vision, Graz, Austria, September 9–13, pp. A-62 ff.

Burks, T.F., Shearer, S.A., Payne, F.A., 2000a. Classification of weed species using color texture features and discriminant analysis. Transactions of the ASAE 43 (2), 441–448.

Burks, T.F., Shearer, S.A., Gates, R.S., Donohue, K.D., 2000b. Backpropagation neural network design and evaluation for classifying weed species using color image texture. Transaction of the ASAE 43 (4), 1029–1037.

Chang, C.C., Wang, L.L., 1996. Color texture segmentation for clothing in a computer-aided fashion design system. Image and Vision Computing. 14 (9), 685–702, October.

Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J., 2001. Color image segmentation: Advances and prospects. Pattern Recognition 34, 2259–2281.

Conci, A., Proenca, C.B., 1997. A box-counting approach to color segmentation. In: Proceedings of International Conference on Image Processing, IEEE Computer Society, Los Alamitos, California, pp. 228–230.

Dell'aqua, F., Gamba, P., 2001. Detection of urban structures in SAR images by robust fuzzy clustering algorithms: The example of street tracking. IEEE Transactions on Geoscience and Remote Sensing 39 (10), 2287–2297.

Dial, G., Gibson, L., Poulsen, R., 2001. IKONOS Satellite Imagery and its Use in Automated Road Extraction. Baltsavias, E.P., Gruen, A., Van Gool, L. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images (III), Zurich.

Doucette, P., Agouris, P., Stefanidis, A., Musavi, M., 2001. Self-organised clustering for road extraction in classified imagery. ISPRS Journal of Photogrammetry and Remote Sensing 55, 347–358.

Eidenbenz, C., Käser, C., Baltsavias, E., 2000. Atomi-Automated reconstruction of topographic objects from aerial images using vectorized map information. In: International Archives of Photogrammetry and Remote Sensing, vol. XXIII, Amsterdam.

Fabre, S., Briottet, X., Appriou, A., 2001. Impact of Contextual Information Integration on Pixel Fusion. IEEE Transactions on Geoscience and Remote Sensing 40 (9), 1997–2010.

Förstner, W., Pluemer, L. (Eds.), 1997. Semantic Modeling for the Acquisition of Topographic Information from Images and Maps. Birkhaeuser Verlag, Basel.

Fukunaga, K., 1990. Statistical Pattern Recognition, second ed. Academic Press.

Gagalowicz, A., 1987. Texture modelling applications. The Visual Computer, Springer Verlag, vol. 3, pp. 186–200.

Gerke, M., 2002. Scene Analysis in Urban Areas Using a Knowledge Based Interpretation System. ISPRS. In: Photogrammetric Computer Vision, Graz, Austria, September 9–13, pp. B-63.

Grafarend, E.W., 2002. Geodesy: The Challenge of the Third Millennium, first ed. Springer-Verlag.

Haralick, R.M., 1979. Statistical and Structural Approaches to Texture. Proceedings of IEEE 67 (5), 786–804.

Heijden, F. Van Der, 1995. Image Based Measurement Systems. John Wiley & Sons.

Heipke, C., Pakzad, K., Straub, B.M., 2000. Image analysis for GIS data acquisition. Photogrammetric Record 16 (96), 963–985.

Hellwich, O., Wiedemann, C., 2000. Object extraction from high resolution multisensor image data. Chair for Photogrammetry and Remote Sensing, Fusion of Earth Data, Technical University Munich, Arcisstr. 21, D-80290 Munich, Germany, Sophia Antipolis, France, January 26–28.

Hinz, S., Baumgartner, A., Ebner, H., 2001. Modeling contextual knowledge for controlling road extraction in urban areas. In: IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas. Rome, 2001.

Jayanta, M., 2002. MRF clustering for segmentation of color images. Pattern Recognition Letters 23 (8), 917–929.

Jeon, B.K., Jang, J., Hong, K., 2000. Map based road detection in spaceborne synthetic aperture radar images based on curvilinear structure extraction. In: Lessard, R.A. (Ed.), Optical Engineering vol. 39 (09). pp. 2413–2421.

Kass, M., Witkin, A., Terzopoulos, D., 1987. Snakes: Active contour models. International Journal of Computer Vision, 321–331.

Kohlas, J., 1995. Mathematical Foundations of Evidence Theory. In: Coletti, G., Dubois, D., Scozzafava, R. (Eds.), Mathematical Models for Handling Partial Knowledge in Artificial Intelligence. Plenum Press, UK, pp. 31–64.

Kohlas, J., Besnard, P., 1995a. An algebraic study of argumentation systems and evidence theory. Technical Report 95-13. Institute of Informatics, University of Fribourg.

Kohlas, J., Besnard, P., 1995b. Evidence theory based on general consequence relations. International Journal of Foundations of Computer Science 6 (2), 119–135.

Kohlas, J., 1997. Allocation of arguments and evidence theory. Theoretical Computer Science 171, 221–246.

Krishnamoorthi, R., Bhattacharyya, P., 1997. On unsupervised segmentation of color texture images. In: Proceedings of the Fourth International Conference on High-Performance Computing, IEEE Computer Society Press, Los Alamitos, California, pp. 500–504.

Mao, J., Jain, A.K., 1992. Texture classification and segmentation using multiresolution simultaneous autoregressive models. Pattern Recognition 25 (2), 173–188.

Markov, N.G., Napryushkin, A.A., 2000. Use of remote sensing data at thematic mapping in GIS. Third AGILE Conference on Geographic Information Science. Helsinki/Espoo, Finland.

Mayer, H., 1999. Automatic object extraction from aerial imagery—a survey focusing on buildings. Computer Vision and Image Understanding 2 (74), 138–149.

McGlone, C., 1998. Block adjustment of linear pushbroom imagery with geometric constraints. In: International Archives of Photogrammetry, Cambridge, UK 32-2.

McKeown, D.M., McMahill, J., Cadwell, D., 1999a. The use of spatial context awareness in feature simplification. In: Geocomputation, CD ROM, Greenwich, UK, pp. 25–28.

McKeown, D.M., Cochran, S.D., Ford, S.J., McGlone, J.C., Shufelt, J.A., Yocum, D.A., 1999b. Fusion of HYDICE hyperspectral data with panchromatic imagery for cartographic feature extraction. IEEE Transactions on Geoscience and Remote Sensing 37 (3), 1261–1277.

Mena, J.B., 2002. Vectorización automática de una imagen binaria mediante *K*-means y degeneración de la triangulación de Delaunay. Revista de la Asociación Española de Teledetección (17), 21–29.

Mena, J.B., 2003. State of the art on automatic road extraction for GIS update: a novel classification. Pattern Recognition Letters 24 (16), 3037–3058.

Mirmehdi, M., Petrou, M., 2000. Segmentation of color texture. IEEE Transactions Pattern Analysis and Machine Intelligence 22 (2), 142–158.

Moghaddamzadeh, A., Bourbakis, N., 1997. A Fuzzy region growing approach for segmentation of color images. Pattern Recognition 30 (6), 867–881.

Ohlhof, T., Emge, T., Reinhardt, W., Leukert, K., Heipke, C., Pakzad, K., 2000. Generation and Update of VMAP Data using Satellite and Airborne Imagery. IntArchPhRS (33) B4/2, pp. 762–768.

Park, S.H.O., Yun, D., Uk Lee, S., 1998. Color image segmentation based on 3D clustering: Morphological approach. Pattern Recognition 31 (8), 1061–1076.

Paschos, G., Valavanis, F.P., 1999. A color texture based visual monitoring system for automated surveillance. IEEE Transactions on Systems, Man and Cybernetics 29 (2), 298–307.

Peddle, D.R., Ferguson, D.T., 2002. Optimization of multisource data analysis: an example using evidential reasoning for GIS data classification. Computers and Geosciencies 28, 45–52.

Richards, J.A., Jia, Xiuping, 1999. Remote Sensing Digital Image Analysis. Springer.

Rellier, G., Descombes, X., Zerubia, J., 2002. Local registration and deformation of a road cartographic database on a SPOT satellite image. Pattern Recognition 35, 2213–2221.

Ripley, B.D., 1996. Pattern recognition and neural networks. Cambridge University Press.

Ruskoné, R., Airault, S., 1997. Toward an automatic extraction of the road network by local interpretation of the scene. In: 46th Photogrammetric Week, Stuttgart.

Sanniti di Baja, G., 1994. Well-Shaped, stable and reversible skeletons from the (3,4)-distance transform. Journal of Visual Communication and Image Representation 5 (1), 107–115.

Serra, J., 1982. Image Analysis and Mathematical Morphology, vol. 1. Theoretical Advances, vol. 2. Academic Press, London.

Shafer, G., 1976. A Mathematical Theory of Evidence. Princeton University Press, Princeton, NJ.

Skarbek, W., Koschan, A., 1994. Color Image Segmentation. Tech report, Technical University, Berlin.

Song, L., Ruikang, Y., Saarinen, I., Gabbouj, M., 1996. Use of fractals and median type filters in color texture segmentation. In: Proceedings of the IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World, ISCAS 96 IEEE, New York, NY, USA, pp. 108–111.

Sowmya, A., Trinder, J., 2000. Modelling and representation issues in automated feature extraction from aerial and satellite images. Journal of Photogrammetry and Remote Sensing 55, 34–47.

Wallace, S.J., Hatcher, M.J., Priestnall, G., Morton, R.D., 2001. Research into a framework for automatic linear feature identification and extraction. In: Baltsavias, E.P., Gruen, A., Van Gool, L. (Eds.), Automatic Extraction of Man Made Objects from Aerial and Space Images (III), Zurich.

Weszka, J., Dyer, C., Rosenfeld, A., 1976. A comparative study of texture measures for terrain classification. IEEE Transactions on System, Man, and Cybernetics SMC-6 (4), 269–285.

Wiedemann, C., Heipke, C., Mayer, H., Jamet, O., 1998. Empirical evaluation of automatically extracted road axes. In: Kevin, J., Bowyer, P., Phillips, J. (Eds.), Empirical Evaluation Methods in Computer Vision. IEEE Computer Society Press, pp. 172–187, Also in: 9th Australasian Remote Sensing and Photogrammetry Conference, The University of New South Wales, Sydney, Paper No. 239 (CD).

Willrich, F., 2002. Quality control and updating of road data by gis driven road extraction from imagery. In: Proceedings of the International Symposium on Geospatial Theory, Processing and Applications, and ISPRS IV symposium, Ottawa, Canada, p. 7.

Yang, J.F., Hao, S.-S., Chung, P.-C., 2002. Color image segmentation using fuzzy *C*-means and eigenspace projections. Signal Processing 82 (3), 461–472.

Zhang, C., Baltsavias, E., 2002a. Edge matching and 3D road reconstruction using knowledge based methods. Institute of Geodesy and Photogrammetry. ETH Hönggerberg, CH-8093 Zurich, Switzerland.

Zhang, C., Baltsavias, E., 2002b. Improving cartographic road databases by image analysis. ISPRS. Photogrammetric Computer Vision, September 9–13, pp. A-400. Graz, Austria.

Zhao, H., Kumagai, J., Nakagawa, M., Shibasaki, R., 2002. Semiautomatic road extraction from high resolution satellite image. In: ISPRS Photogrammetric Computer Vision, September 9–13, Graz, Austria, p. A-406.

Zhong, Y., Jain, A.K., 2000. Object location using color texture and shape. Pattern Recognition 33 (4), 671–684.

Zugaj, D., Lattuati, V., 1998. A new approach of color images segmentation based on fusing region and edge segmentations outputs. Pattern Recognition 31 (2), 105–113.