Extraction of geographical entities from aerial images

Manuele Bicego, Silvio Dalfini, and Vittorio Murino
Dipartimento di Informatica, Università di Verona
Ca' Vignal 2, Strada Le Grazie 15, 37134 Verona - ITALY
e-mail: {bicego@sci.univr.it, vittorio.murino@univr.it}

Abstract—**In this paper a complete system for extracting geographical entities from aerial images is proposed. The method treats separately the extraction of green areas, forestry and buildings from the extraction of roads. The former issue is addressed with a color segmentation followed by a region growing approach. The latter issue is addressed by a probabilistic contour tracking approach, able to follow roads in highly urbanized areas. Experimental evaluation on real highly complex images showed that the proposed method is able to extract most part of the geographical entities present in the image, even in highly urbanized areas.**

I. INTRODUCTION

The extraction and classification of geographical entities from aerial images has grown in importance in the last years, due to its wide applicability in different research areas. From the others, the GIS (*Geographic Information System*) topic is one of the most interesting, for both practical and theoretical motivations. In this context, annotation and classification of aerial images are used for different objectives, as creating vectorial representations, validating existing data, finding disagreements between older representations and current situations (to discover abusive buildings, for example), or moreover performing some kind of content-based retrieval on raster data [1]. Moreover, the identification and classification of the symbols of interest from raster data is the basic step towards the integration between raster and vector data.

In the literature, many approaches have been proposed in the last years: for a review of the automatic object extraction techniques recently proposed see [2]; for another interesting survey on roads extraction, see [3].

In this paper a new complete extraction scheme is proposed, able to extract all the important entities from an aerial image. The method treats separately the extraction of fields, forestry and buildings from the extraction of roads. The former issue is addressed with a quite conservative color segmentation (performed in the HSI color space), followed by a region growing approach (with a hysteresis thresholding mechanism). The latter issue is addressed by a probabilistic contour tracking approach [4], able to follow roads in highly urbanized areas. The high density of roofs and the agglomerate road networks, together with the occlusions due to cars and shadows, makes the extraction of these kinds of entities very difficult.

Experimental evaluation on real highly complex images showed that the proposed method is able to extract most part of the geographical entities present in the image, even if the images were gathered from a not perfectly orthogonal view (presence of several shadows).

The rest of the paper is organized as follows: in Section II the method for the extraction of fields, forestry and buildings are described, while the approach for detecting roads is detailed in Section III. In Section IV some results are presented, while in Section V conclusions are drawn.

II. FIELDS, FORESTRY AND BUILDINGS EXTRACTION

The extraction of fields, forestry, and buildings is based on the observation that these entities share similar topological characteristics. Fields and forestry are very broad regions with regular shape and quite homogeneous color. The color of the buildings is quite homogeneous in both sides of the roof, even if of different colors because of shadows: moreover, the homogeneity is weakened by the different color of roof tiles. Buildings, if individually examined, have small dimensions and regular shapes; if considered as an agglomerate (as in urban areas) they have large dimensions and irregular shapes. These characteristics indicate that color based segmentation could be an appropriate approach for extracting these entities.

The method proposed in this paper combines a color segmentation process with a hysteresis threshold region growing approach, able to refine the coarse color segmentation. Before starting, the image is firstly transformed from the RGB (i.e., the color additive format composed by Red Green Blu channels) to the HSI (Hue, Saturation, Intensity) color system. The HSI system has been preferred since it allows for a separation between chrominance (color and saturation) from intensity, so that it is possible to discriminate on the basis of color information only. More in detail, the extraction process is divided into two steps: in the first step (color segmentation), the image is segmented by thresholding the H and the S values, using a really conservative threshold: only "sure" pixels are extracted. From this initial segmentation a growing algorithm is applied, where another threshold is used; this second threshold is less conservative than the first. The recursive rule applied extracts a pixel only if it is near to an already labelled pixel ("sure" or a neighbor of a "sure"). In this way the extracted regions are composed by "good" pixels and by pixels near to them, having some slightly different characteristics due to noise or shadows, but still valid.

III. ROAD EXTRACTION

In this paper, the road extraction problem has been faced using a contour tracking approach, as proposed in [4], [5]: starting from an initial point, the road contour is followed in the image, with a space-tracking process. In particular, in this paper a modified version of the JetStream algorithm [4] has been proposed. JetStream is the space version of the famous CONDENSATION algorithm [6], a Bayesian approach to tracking recently introduced. This Monte Carlo technique, based on sequential importance sampling/resampling, provides a statistical framework for propagating sample-based approximations of posterior distributions, with almost no restriction on the ingredients of the model. This kind of approach outperforms Kalman approaches [5], because of its ability to perform multi-hypothesis tracking, which is a necessary condition for dealing with noisy situations. Algorithms based on the Kalman filter have also the disadvantage of being unimodal, i.e., only one contour can be followed at each time step.

The main idea under the JetStream algorithm is to approximate the posterior probability, representing all the information of the model obtainable from the image, as a set of samples $S^t = \{s_1^t ... s_N^t\}$, or particles, each featured with a weight $\{\pi_1^t ... \pi_N^t\}$. Each particle represents a path in the image (the road), while the weight is the "reliability" of the path (the probability). In particular,

$$s_i^t = (x_{i,0}, m_{i,0}), (x_{i,1}, m_{i,1}), ..., (x_{i,t}, m_{i,0})$$

where $x_{i,t}$ represents the middle point of a road large m_i pixels. In other words, m_i is the distance between the two side points $x_{i,t}^+$ and $x_{i,t}^-$. The point $x_{i,0}$ is the first road point (given), and $x_{i,1}, ..., x_{i,t}$ represent the evolution of the road (*i.e.* the subsequent middle points). The distance between two consecutive points represents the "space step", and denotes the "speed" with which the algorithm follows the road in the image.

Supposing that N samples are maintained during all evolution, at each step t the following operations are performed:

1) *Sampling.* N particles are sampled from the set S^{t-1}, based on the weights $\{\pi_i^{t-1}\}$: the higher the weight π_i^{t-1}, the larger the probability of the sample s_i^{t-1} to survive (to be extracted). Let us denote this set as $\tilde{S}^{t-1}$
2) *Prediction.* In this step a dynamic (prediction) is applied to the selected set of particles $\tilde{S}^{t-1}$, in order to predict the next position in the road and obtain the set S^t: this dynamics encodes the a priori knowledge on the possible evolution of the road contour, defined as

$$s_i^t = \tilde{s}_i^{t-1}, (x_{i,t}, m_{i,t})$$

with $x_{i,t} = x_{i,t-1} + R(\theta_i)(x_{i,t} - x_{i,t-1})$. $R(\theta_i)$ denotes a rotation of an angle θ_i. The angle θ_i is drawn from two different distributions, depending from the value taken by the corner function $c(x_{i,t})$, defined as

$$c(x_{i,t}) = \begin{cases} 1 & \text{if there is a corner in } x_{i,t}^+ \text{ or in } x_{i,t}^- \\ 0 & \text{otherwise} \end{cases}$$

The angle θ_i is then drawn from:

- an uniform distribution in $(-\frac{\pi}{2}, \frac{\pi}{2})$, if $c(x_{i,t}) = 1$
- a Gaussian distribution $\mathcal{N}(0, \sigma_\theta^2)$, otherwise.

The evolution of the width of the road m_i is driven by

$$m_{i,t} = m_{i,t-1} + v_{i,t} \tag{1}$$

with $v_{i,t}$ drawn from a Gaussian of zero mean and fixed small variance.
3) *Weighting.* These predictions are then validated using information from the image (likelihood), obtaining the new weights. The likelihood is computed by using image gradient information (both direction and magnitude), taking into account also the corners. More in detail, the likelihood ℓ of a sample s_i^t is determined by the following formula

$$\ell(s_i^t) = \frac{p_{\text{on}}(x_{i,t}^+) p_{\text{on}}(x_{i,t}^-)}{p_{\text{off}}(x_{i,t}^+) p_{\text{off}}(x_{i,t}^-)} \tag{2}$$

where p_{on} and p_{off} are the probabilities of the "on contours" and the "off contours", respectively. These quantities are computed for both the points $x_{i,t}^+$ and $x_{i,t}^-$, with the use of the gradient ∇I and the corner function c, and are defined as (removing the i-th index and the +/- apex for readability):

$$p_{\text{on}}(x_t) \propto \frac{c(x_t)}{\pi} + (1 - c(x_t)) \mathcal{N}\left(\psi(x_t); 0, \frac{\sigma_\psi^2}{|\nabla I(x_t)|}\right) \tag{3}$$

and

$$p_{\text{off}}(x_t) \propto \exp - \frac{\nabla I(x_t)}{\lambda} \tag{4}$$

where $\psi(x_t)$ is the angle between the gradient normal $\nabla I(x_t)^\perp$ and the segment (x_{t-1}, x_t), λ is the average gradient norm over the image, and σ_ψ^2 has experimentally been fixed to 1.36.

At each step, the extracted road is represented by the most probable sample.

A. The strategy

In this paper a modification of the JetStream algorithm is proposed: more in detail, we change the weighting step (step 3) in a twofold manner. First, we propose to augment the information used in the likelihood computation (i.e., gradient) by considering also the color information (roads are typically homogeneous inside). Color information are obtained by segmenting the image in the HSV (Hue Saturation Value) color space. Also in this case this space has been chosen as it permits to separate the chroma information from the luminosity, resulting in a more effective segmentation. This results in a change of the definition in eq. (2):

$$\ell(s_i^t) = \frac{p_{\text{on}}(x_{i,t}^+) p_{\text{on}}(x_{i,t}^-) U(x_{i,t}^+, x_{i,t}^-)}{p_{\text{off}}(x_{i,t}^+) p_{\text{off}}(x_{i,t}^-)} \tag{5}$$

where $U(x_{i,t}^+, x_{i,t}^-)$ measures the color uniformity of the pixels between $x_{i,t}^+$ and $x_{i,t}^-$.

The second change to the weighting step proposes to add an inertial term to the likelihood computation procedure, that allows a particle to survive along its direction for few additional steps, even if not supported by high likelihood. This permits to go beyond small occlusions like isolated trees or cars. This is obtained by defining the new likelihood $\hat{\ell}(x_{i,t})$ as:

$$\hat{\ell}(s_i^t) = \alpha\ell(s_i^t) + (1-\alpha)\ell(s_i^{t-1}) \quad (6)$$

where α is the parameter driving the decay of the process memory. In this way also older likelihood is taken into consideration. When $\ell(x_{i,t})$ is zero (sure non-contour point), we force the $R(\theta)$ function of the prediction step (step 2) to 0. By means of these two modifications, a particle can continue to survive (along its last good run direction) for few more iterations, even if not supported by a high likelihood, allowing to overcome from small occlusions.

These modification allows to recover streets also in a highly urbanized areas, where shadows and occlusions make the process more difficult.

IV. Results

The proposed approach has been tested on several real images[1], with respect to the extraction of fields, forestry and buildings and to the extraction of roads. Regarding the former class, two examples of the extraction are proposed in Fig. 1 and Fig. 2: buildings are red, groups of tree are green, and isolated tree are light green. From Fig. 1 one could notice that segmentation is quite accurate: in particular both agglomerate and isolated buildings are properly segmented, as well as agglomerate and isolated trees. From Fig. 2 it is important to note that the fields are correctly segmented; moreover, also the small buildings situated in the center of the image are detected, even if the color of the roof is very similar to the color of the field. In this case, the first step of the color segmentation algorithm, *i.e.* the first conservative segmentation, permits to succeed in this discrimination.

Results from the roads extraction process are proposed in Fig. 3 (a) and (b): one could notice that here the task is quite difficult, due to shadows and car occlusions. Nevertheless the proposed approach is able to correctly extract all relevant roads in the images.

(a)

(b)

Fig. 1. Buildings and woods extraction, first example: (a) original image; (b) segmented image.

V. Conclusions

This paper proposed a complete approach to the extraction of interesting entities from aerial images. Buildings, woods and fields were detected with color segmentation followed by a hysteresis region growing process on the HSI color space, while roads were extracted with a modified probabilistic approach for contour tracking. The proposed approach has been tested on real complex images, presenting encouraging extraction results.

[1] All images are courtesy of CO.GE.ME. Informatica - Rovato (Brescia).

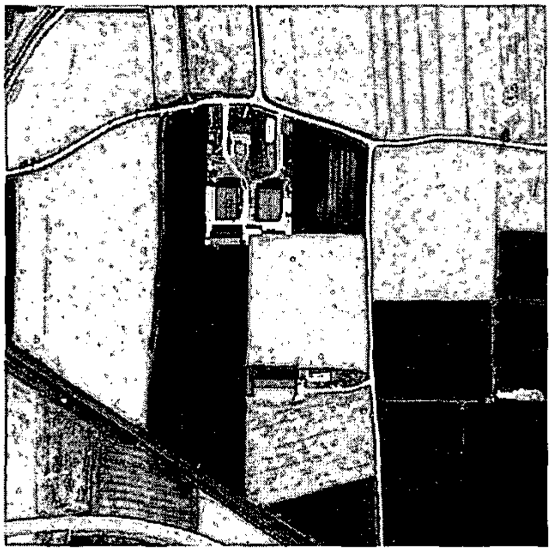
References

[1] H. Sammet and A. Soffer, "MARCO: Map retrieval by content," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 783–798, 1996.

[2] H. Mayer, "Automatic object extraction from aerial imagery: A survey focusing on buildings," *Computer Vision and Image Understanding*, vol. 74, pp. 138–149, 1999.

[3] M.-F. Auclair-Fortier, D. Ziou, C. Armenakis, and S. Wang, "Survey of Work on Road Extraction in Aerial and Satellite Images," Département de mathématiques et d'informatique, Université de Sherbrooke, Tech. Rep. 247, 2000.

[4] P. Pérez, A. Blake, and M. Gangnet, "Jetstream: Probabilistic contour


(a)

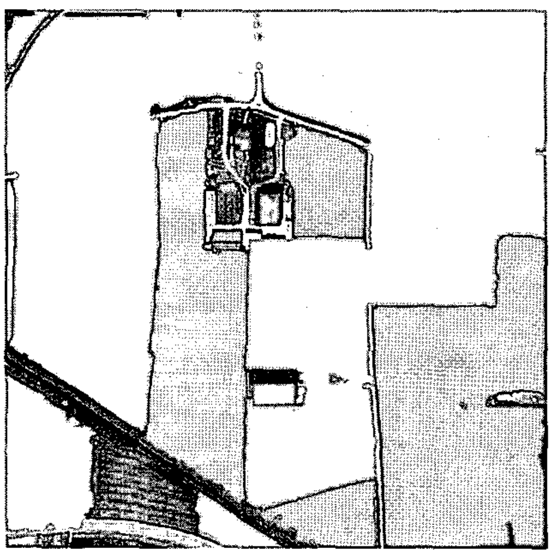

(b)

Fig. 2. Buildings and woods extraction, second example: (a) original image; (b) segmented image.

extraction with particles," *International Conference on Computer Vision*, 2001.

[5] G. Vosselman and J. de Knecht, "Road tracking by prole matching and kalman filtering," *Birkhäuser Verlag*, pp. 265–274, 1995.

[6] A. Blake and M. Isard, "Condensation - conditional density propagation for visual tracking," *International Journal Computer Vision*, vol. 29(1), pp. 5–28, 1998.


(a)

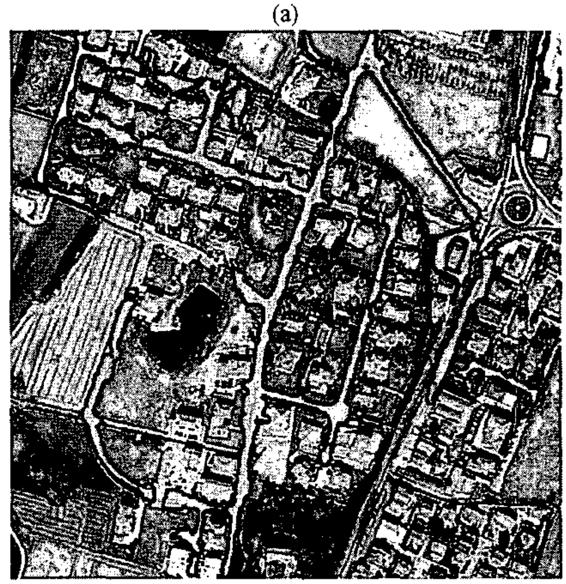

(b)

Fig. 3. Roads extraction.