Extraction of Main Urban Roads from High Resolution Satellite Images by Machine Learning

Conference Paper *in* Lecture Notes in Computer Science · January 2006

DOI: 10.1007/11612032_25 · Source: DBLP

Xianqing Tai
Chinese Academy of Sciences

Extraction of Main Urban Roads from High Resolution Satellite Images by Machine Learning

Yanqing Wang[1,2], Yuan Tian[2], Xianqing Tai[2], and Lixia Shu[1]

[1] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing, 100080, P.R. China
{yqwang, lxshu}@nlpr.ia.ac.cn
[2] Integrated Information System Research Center, 100080, P.R. China
{yuan.tian, xianqing.tai}@ia.ac.cn

Abstract. This paper focuses on automatic road extraction in urban areas from high resolution satellite images. We propose a new approach based on machine learning. First, many features reflecting road characteristics are extracted, which consist of the ratio of bright regions, the direction consistency of edges and local binary patterns. Then these features are input into a learning container, and AdaBoost is adopted to train classifiers and select most effective features. Finally, roads are detected with a sliding window by using the learning results and validated by combining the road connectivity. Experimental results on real Quickbird images demonstrate the effectiveness and robustness of the proposed method.

Keywords: AdaBoost, Local Binary Pattern, Machine Learning, Road Extraction.

1 Introduction

Road extraction in urban areas has been an important task for generating geographic information systems (GIS). Especially in recent years, the rapid development of urban areas makes it urgent to provide up-to-date road maps. The timely road information is very useful for the decision-makers in urban planning, traffic management and car navigation fields, etc.

Nowadays, we are experiencing an explosion in the amount of satellite image data, which provides us abundant data and also brings challenges to the road extraction task at the same time. The conventional road extraction methods by manual are time consuming and tedious, and cannot meet the increasing requirement for such tremendous data. Therefore, it has drawn considerable attention of many researchers on how to develop automatic road extraction systems. And much work has been done for this task. However, automatic extraction of urban roads from high resolution remote sensing imagery is still a challenging problem in digital photogrammetry and computer vision. The main reason is that the diverse road surfaces and the complex surrounding environments such as trees, vehicles and shadows induced by high buildings make the urban roads take on different textures and gray levels in images.

P.J. Narayanan et al. (Eds.): ACCV 2006, LNCS 3851, pp. 236–245, 2006.

1.1 Related Work

In the past decades, a large number of papers have been published for automatic road extraction. However, most of them focus on extracting roads in rural or open areas. By contrast, the efforts made for urban road extraction are relatively few [1], [2], [3], [4], [5], [6]. These methods can be roughly divided into two categories: heuristic-based methods and Bayesian-based methods.

Heuristic-based approaches usually model roads in a semantic way and group the extracted road components with the "hypothesis and test" paradigm. Hinz et al [1] used road substructures such as markings and lanes to extract road segments and further linked them into a global road network. In their later work [2] road details from multiple sources were integrated and then roads were found by iteratively grouping. The approach presented by Price [3] modelled the road network as a regular grid. He assumed that the roads crossed at the specific angles and the road width was approximately constant. After three initial seed points on the grid were given by manual, the grid propagation, verification and refinement process was performed based on edge and contextual information. McKeown et al described a multi-level cooperative methods for road tracking by assuming that there existed some specific patterns or textures for road surface [4]. They used a texture-correlation-based tracker and an edge linker to obtain the road candidates.

By contrast, Bayesian approaches generally build stochastic process models for road data and find roads by probability methods. For example, Barzohar and Cooper [5] established a geometric-stochastic model for road image generation and used maximum a posteriori probability to estimate road boundaries.

As can be seen from these methods, road details such as markings and structural information of road surface, are valuable cues for urban road detection from high resolution images. Furthermore, it is advisable to combine the local and global properties to extract road network. However, the two types of methods described above are limited because of the difficulties for building comprehensive road models covering all possible situations. As is known, the diversity of roads, the complexity of surrounding environments, the variation of illumination, the appearance of cars and trees and different view angles of sensors, make it very difficult to built a general road model.

1.2 Overview of the Proposed Method

In order to deal with the difficulties for building comprehensive road models and to make full use of the characteristics of urban roads, we propose an automatic approach based on machine learning. It can be divided into three steps. First, a series of features reflecting road characteristics are extracted. They include the ratio of bright lines on the road surface, the directional consistency of road markings and local binary patterns (LBP). These features are then input into a learning container, and AdaBoost is adopted to train classifiers and select distinct features. Finally, on the basis of the learning results roads are detected with a sliding window and further validated by combing the road connectivity.

The road extraction process is performed based on the essential features of urban roads which can be achieved by learning from a great amount of training examples with diverse appearances.

The remainder of this paper is organized as follows. Section 2 and Section 3 describe the features and the machine learning process based on AdaBoost, respectively. The experimental results and discussions are given in Section 4. Finally, Section 5 concludes the paper.

2 Features

There are many valuable indications about urban roads in high resolution images, for example, road markings are bright and parallel lines; road markings cover only a part of road surfaces; there exist some patterns or textures for urban roads.

Obviously, these assumptions are reasonable because the majority of main urban roads satisfy these conditions, especially those in built-up areas. To make the best use of these characteristics, it is important to extract features that are robust to illumination variations, building shadows and disturbances by cars or trees.

Here, three kinds of features are extracted, namely, the coverage ratio, the direction consistency of road markings and LBP-based features for road textures. There are mainly three reasons for choosing LBP as the road texture descriptor. First, the adopted LBP-based features are invariant to orientations and the monotonic transform of gray levels. Secondly, it can perform multi-resolution analysis by combining different neighborhoods for LBP. Furthermore, LBP is theoretically simple and easy to implement.

2.1 Coverage Ratio of Bright Lines

This feature is used to describe the distribution of road markings. First, the image is segmented to obtain bright lines. This can be accomplished by ridge detection based on the methods presented by Steger [12]. If bright regions are denoted as foreground with 1, and dark regions as background with 0, then the coverage ratio of the bright lines can be computed by the formula:

$$ratio = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} I(i,j)}{M \times N}$$

where

$$I(i,j) = \begin{cases} 1, \text{foreground;} \\ 0, \text{background.} \end{cases}$$

Here, M and N are the number of rows and columns of the image, respectively.

As is known, road markings cover only a part of road surface, therefore, the ratio feature can be used as one of the indicators for urban roads.

2.2 Direction Consistency

Most of road markings are parallel, so the direction consistency feature is considered to make use of their direction information. First, we obtain edges with Canny edge detector, and then Hough transform is carried out to further get their direction information. The results are shown in Figs. 1 and 2. The first two rows in these figures are roads and their edge features, and the last two ones are non-roads and their results. It can be seen that the directions of road markings are obviously consistent. Fig. 2 (b) is obtained by accumulating the votes of straight lines at different directions. The horizontal axis denotes the angles ranging from 0 to 179, and the vertical axis is the number of occurring times. Furthermore, the standard deviation can be computed directly from this figure. In Fig. 2, one can see that when the directions of edges are similar, the accumulation values converge on a small range of direction angles; otherwise,

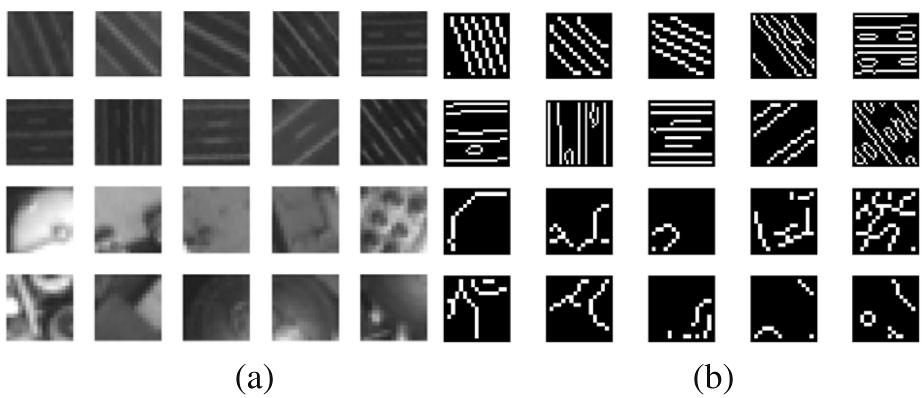

Fig. 1. (a) Original sub-images. (b) Detected edges.

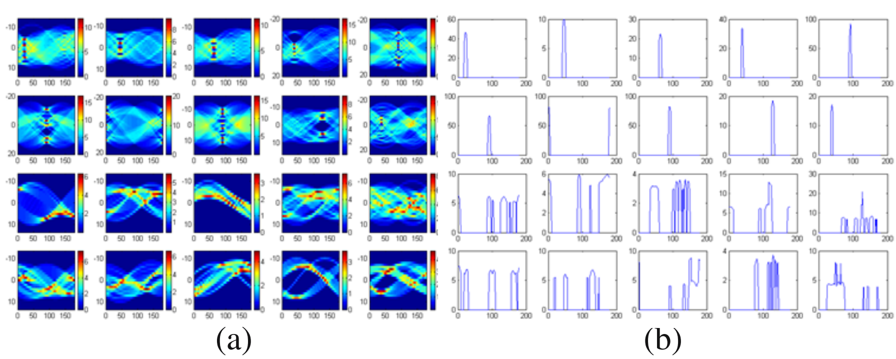

Fig. 2. (a) Hough transform. (b) Accumulation of the threshold result according to different directions.

they are spread on different angles. It shows that the direction consistency is a good feature for urban road extraction.

2.3 Features Based on LBP

Local Binary Pattern (LBP) [7], [8] is a descriptor for local texture. The original LBP operator labels each pixel in an image by comparing the gray values in a circularly symmetric neighborhood with that of the center pixel and then transforming the binary pattern into an integer. An example for LBP formation is shown in Fig. 3. The operator is denoted as $LBP_{P,R}$ for a neighborhood of P pixels that are symmetrically located on a circle of radius R. It can produce 2^P different binary patterns by the P pixels in the neighbor set. LBP is insensitive to the monotonic intensity transformation, because it is only dependent on the relative order of the gray levels. From its formation process, one can see that

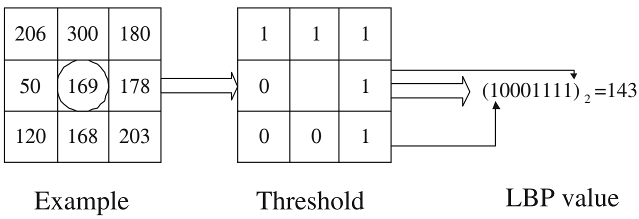

Fig. 3. An example for $LBP_{8,1}$ formation

$LBP_{P,R}$ is sensitive to image rotation, and cannot provide good discrimination for small size images because of the divergence of pattern occurrences. Therefore, as a variant of $LBP_{P,R}$, $LBP^{riu2}_{P,R}$, is introduced in [7]. This improved operator is rotation invariant and deals with the frequently occurring patterns and the less occurring ones in a different way. The procedure for $LBP^{riu2}_{P,R}$ is detailed as follows:

$$LBP^{riu2}_{P,R} = \begin{cases} \sum_{p=0}^{P-1} s(p), & \text{if } U(LBP_{P,R}) \le 2; \\ \text{P+1}, & \text{otherwise.} \end{cases}$$

where

$$U(LBP_{P,R}) = |s(p-1) - s(0)| + \sum_{p=1}^{P-1} |s(p) - s(p-1)|,$$

and s(x)= 0 or 1.

$LBP^{riu2}_{P,R}$ has only P+2 different output values, less than 2^P ones, which makes it more convergent for texture discrimination.

Choosing different P and R, we can get operators of different spatial resolutions. Consequently, multi-resolution analysis can be realized by combining multiple operators with different (P, R) pairs.

Here, by simultaneously using $LBP_{8,1}^{riu2}$, $LBP_{16,2}^{riu2}$ and $LBP_{24,3}^{riu2}$ to extract road texture features, we are able to get 54 (10+18+26)-bin histogram of LBP-based features. Together with the two features described above, i.e., the coverage ratio of bright lines and the direction consistency, 56 features are obtained for the learning process.

3 Learning Based on AdaBoost

Given the road and non-road samples and their features, the current important task is to choose a suitable learning algorithm. Consideration of the redundancy of the 56-dimensional features, AdaBoost, which can serve as both a classifier trainer and a feature selector, is used in this study.

AdaBoost is an adaptive learning algorithm that aims to build a strong classifier by linearly combining a set of weak learners [10], [11]. It works by updating the weights of training samples dynamically according to the training error. Freund and Schapire have proved that the training error of the strong classifier decreases exponentially with the number of iterations. Furthermore, AdaBoost achieves a good generalization performance because it manages to maximize the margin between positive and negative examples.

In order to select important features, the weak learner of AdaBoost can be constrained to rely on a single feature. Consequently, AdaBoost obtains a strong classifier and effective features simultaneously during its learning process.

Motivated by the work of Viola et. al [11] we introduce the pyramid idea into the learning process to improve the system efficiency. Similar to a pyramid running from the top to bottom, at the first layer of the learning process a simple classifier with less features is trained to obtain a high detection rate, while propagating to the next layers more accurate and complex classifiers with more features are built to remove those false positives. The sub-images rejected by earlier classifiers will not be evaluated by subsequent classifiers. So the highly distinguishable but complex classifiers are only required to examine those potential regions, which can reduce the system spending effectively. As for the number of pyramid layers and the number of features, they are determined by considering the detection rate and false positive rate. The learning algorithm for each layer of the pyramid AdaBoost can be summarized as follows:

– Given training examples (x_i, y_i), $i = 1, 2, ..., n$, where x_i is a 56-D feature vector, and y_i= 0, 1 for road (positive) and non-road (negative) samples respectively.
– Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for y_i=0, 1 respectively, where m and l are the number of positive and negative samples respectively.
– t=0; While (FAR and AR are not satisfied):
  1. Update t: $t := t + 1$;
  2. Normalize the weights:

$$w_{t,i} := \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}.$$

3. Train a classifier h_j for each feature j, the error related to w_t is $\epsilon_j = \Sigma_i w_i|(h_j(x_i) - y_i)|$.
   Choose the classifier $h_t = h_{argmin_j \epsilon_j}$.
4. $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$, $\alpha_t = log\frac{1}{\beta_t}$, $e_i = |(h_t(x_i) - y_i)|$.
5. Update the weights:
$$w_{t+1,i} = w_{t,i}\beta_t^{1-e_i}.$$

– Obtain the number of weak classifiers: T= t.
– Output the strong classifier of this layer:
$$h(x) = \begin{cases} 1, \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2}\sum_{t=1}^{T} \alpha_t; \\ 0, \text{otherwise.} \end{cases}$$

From the above, it can be seen that for a certain layer features and weak classifiers will be continuously added until both the FAR and AR requirements are met. Then these weak classifiers are combined to get a strong classifier for this layer. The number of layers can also be obtained when FAR and AR both reach the expected value. Once the learning process is completed, all of the classifiers for different layers and the selected features obtained by AdaBoost are regarded as a paradigm. For a given image, urban roads are then detected according to this paradigm.

4 Experiments

In order to demonstrate the performance of the proposed method, Quickbird imagery, whose resolution is 0.61 m/pixel, is used in this study.

4.1 Training Data

We collected 833 positive samples from Quickbird images manually. These sub-images consisted of various road directions and appearances, and their sizes were also different, ranging from 14×14 to 30×30 pixels. In order to enlarge the training data set, we rotated the road samples by $2°$, $4°$, $6°$, $8°$, $10°$, respectively. Therefore, we obtained 4998 road samples in all for training. Some of the road examples are shown in Fig. 4(a). The negative samples came from 320 gray-level images which were manually examined and found no roads in them. Here, 5000 negative sub-images are used for the learning process, which are obtained randomly from these gray-level images.

4.2 Experimental Results

The features extracted from the training set are input into the pyramid AdaBoost procedure for learning. Three learning curves are obtained in this process, as shown in Fig. 4(b). The horizontal axis in Fig. 4(b) denotes the learning layers, and the vertical axis denotes false accept rate (FAR), accept rate (AR) and correct classification rate (CCR), respectively. The curves reflect the fluctuating trends of these indices according to different layers. As can be seen, FAR and AR

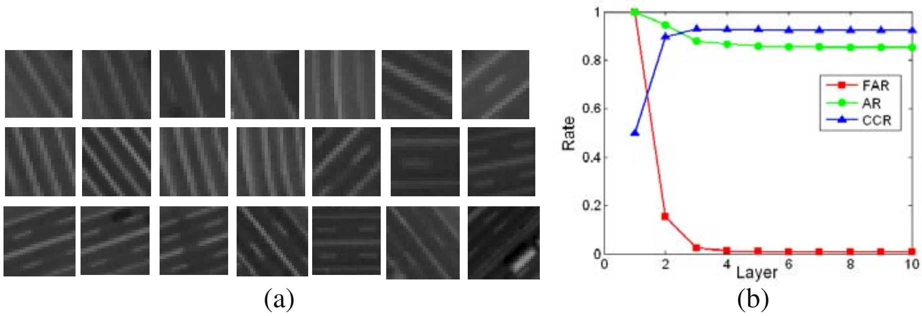

Fig. 4. (a)Road examples and (b)FAR, AR and CCR curves

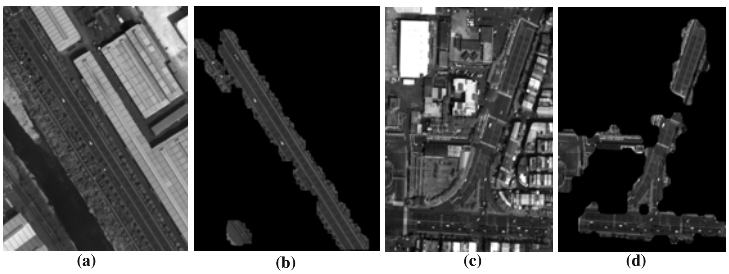

Fig. 5. Experiment 1: (a) original image1 and (b) road extraction result of image1. Experiment 2: (c) original image2 and (d) road extraction result of image2.

curves decrease with layers while CCR increases with layers, and the three curves almost keep unchanged after four layers. Therefore, it is sufficient to choose the first four layers for our AdaBoost. The corresponding number of selected features is 1, 1, 4, 6, respectively. As is expected, the coverage ratio and the direction consistency features are selected successfully.

When given an image, urban roads are to be detected with a sliding window. First, feature extraction is performed for the window of interest according to the learning results, which is followed by the multi-layer classification. Only when

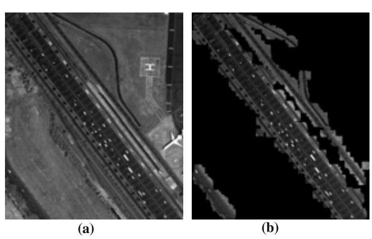

Fig. 6. Experiment 3: (a) original image3 and (b) road extraction result of image3.

the window is considered as a road candidate by the earlier classifiers will it be examined by the next classifiers. Road candidates can be found by altering the size of the sliding window. Decisions are made by the majority-vote method for these different windows, and the largest window of the voting for ones are kept. Finally, the global connectivity of roads is integrated to obtain the final results.

Figs. 5, and 6 are three examples of our road extraction results on real Quickbird Pan imagery. Most of the roads are detected correctly though there are many disturbances, such as, buildings, trees and cars in the image.

4.3 Analysis and Discussions

In Fig. 5(a), there are buildings, trees and a few cars around or on the roads. Roads and buildings have a high similarity in intensity and local edges. Moreover, cars on roads have a negative effect on detection results. However, our method works very well on this image, as shown in Fig. 5(b), the majority of main urban roads has been extracted correctly. Other experiments with more complex surrounding environments and dense traffic flow in Figs. 5(c)(d)and 6 also give fairly good results. All of them demonstrate the effectiveness and robustness of the proposed method.

The qualitative evaluation results are shown in Table 1. Here, we select two evaluation measures provided by [13], namely, completeness and correctness. The reference road maps are obtained by manual, and the evaluation is based on the length of extraction results and reference roads. One can see that both the completeness and correctness are fairly good. The correctness for Experiment 1 and 3 are 100% and both of their completeness exceed 92%. Even for the complex environments in Experiment 2, the completeness and correctness are 89.3% and 84.6%, respectively.

One reason for the robustness of the proposed method is likely because that the vehicles on roads are aligned with the road direction and also appear as bright lines, which resemble road markings. Therefore, vehicles, road markings and road surfaces can be considered as an organic element. The robustness also benefits from the fusion of many features, namely, the coverage ratio of bright lines, the direction consistency and $LBPs$. Integrating these features is favorable to distinguish urban roads from buildings. Although the results are encouraging, there are some places to be improved further. For example, the road boundary is not very precise. Our future work will focus on integrating the road boundary detection into the current system. Additionally, some efforts are needed to promote the proposed method in an operational road detection system.

Table 1. External evaluation of the extraction results

Quality measures	Experiment 1	Experiment 2	Experiment 3
Completeness	92. 6%	89. 3%	94.9%
Correctness	100%	84. 6%	100%

5 Conclusion

In this paper, we present a new approach for main road extraction in urban areas from high resolution satellite images. This method is distinguished from previous work by two highlights. One is that a large number of robust features reflecting the structural and texture properties of urban roads are extracted. The other is to adopt the AdaBoost-based learning algorithm. AdaBoost can not only train the classifiers but select most effective features as well. The experimental results on real Quickbird imagery demonstrate that it is an effective way to detect urban roads by learning from many intrinsic road features.

References

1. S. Hinz, A. Baumgartner, C. Steger, H. Mayer, W. Eckstein, H. Ebner and B. Radig, "Road Extraction in Rural and Urban Areas," *Semantic Modeling for the Acquistion of Topographic Information from Images and Maps*, pp. 7-27, 1999.
2. S. Hinz, "Automatic Road Extraction in Urban Scenes and Beyond," *ISPRS*, Vol. 35, pp. 349-354, July 2004.
3. K. Price, "Road Grid Extraction and Verification," *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 3-2W5, pp. 101-106, 1999.
4. D. M. McKeown and J. L. Denlinger, "Cooperative Methods for Road Tracking in Aerial Imagery," *CVPR*, pp. 662-672, 1988.
5. Meir Barzohar and David B. Cooper, "Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation," *IEEE Trans. PAMI*, Vol. 18, No. 7, pp. 707-721, July 1996.
6. Geman, D., B.Jedynak, "An active testing model for tracking roads in satellite images," *IEEE Trans. PAMI*, Vol.18, No.1, pp.1-14, January 1996.
7. T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 24, No. 7, pp. 971-987, July 2002.
8. A. Hadid, M. Pietikainen and T. Ahonen, "A Discriminative Feature Space for Detecting and Recognizing Faces," *CVPR*, Vol. 2, No. 2, pp. 797-804, 2004.
9. Xiangyun Hu and C.Vincent Tao, "Automatic Main Road Extraction from High Resolution Satellite Imagery," *ISPRS*, XXXIV, August 2002.
10. Y. Freund and R. E. Schapire, "A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting," *European Conference on Computational Learning Theory*, Springer-Verlag, pp. 23-37, March 1995.
11. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *CVPR*, Vol. 1, pp. 511-518, 2001.
12. C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. PAMI*, Vol. 20, No. 2, pp. 113-125, 1998.
13. C. Wiedemann, "Automatic Evaluation of Road Networks," *ISPRS Archives*, Vol. XXXIV, Part 3/W8, Munich, pp. 17-19, September 2003.