

Street Detection with Asymmetric Haar Features

Geovany A. Ramirez and Olac Fuentes

Computer Science Department, University of Texas at El Paso
garamirez@miners.utep.edu, ofuentes@utep.edu

Abstract. We present a system for object detection applied to street detection in satellite images. Our system is based on asymmetric Haar features. Asymmetric Haar features provide a rich feature space, which allows to build classifiers that are accurate and much simpler than those obtained with other features. The extremely large parameter space of potential features is explored using a genetic algorithm. Our system uses specialized detectors in different street orientations that are built using AdaBoost and the C4.5 rule induction algorithm. Experimental results show that Asymmetric Haar features are better than basic Haar features for street detection.

Keywords: Object Detection, Asymmetric Haar Features, Machine Learning, Street Detection.

1 Introduction

Object detection has received a great deal of attention in the last few years. Most recent approaches to this problem pose it as a binary classification problem, where one needs to classify every window in an image as belonging to the class of interest (i.e. streets), or not. Recent research contributions have focused mostly on two main aspects of the problem: feature engineering and classifier design. Feature engineering consists of designing and selecting classes of features that can improve the performance of the classifiers that are based on them. Work on classifier design consists of adapting existing classification algorithms to the detection and recognition problems, or of designing special-purpose algorithms that are targeted to these problems.

In a series of papers, Viola and co-workers advocated an approach to object recognition based on Haar features, which are equivalent to the difference of the sum of the intensity levels of two contiguous equal-sized rectangular image regions (see Figure 1a). They presented an algorithm for computing these features in constant time, which makes them suitable for real-time performance [13]. Using Haar features, a cascade of classifiers based on the Adaboost algorithm was constructed, yielding accurate detection, albeit at the expense of long training times. Successful applications of this methodology were presented in face detection [13], image retrieval [11], and pedestrian detection [14]. There are some extensions to the original set of Haar features proposed by Viola and

Jones such as the work of Lienhart and Maydt, which introduced rotated Haar features as shown in Figure 1b [6]. Another extension was presented by Ramirez and Fuentes, which proposed asymmetric Haar features, which can have regions with either different width or height, but not both [10]. This results in a more expressive feature space, which, as they showed, allows to build classifiers that are much simpler than those obtained with the standard features.

Road extraction is a problem where the goal is to segment roads in satellite/aerial images. It is typically performed in three steps: image pre-processing, road detection, and road following [3]. Road detection consists of determining what regions of the images are likely to be roads. The goal of road following is to build a road network by connecting all the road segments found using road detection. Street extraction is a subproblem of road extraction, where the goal is to extract the roads in satellite/aerial images in urban areas. In this paper we focus on the subproblem of automatic street detection in satellite images. We use the extensions to the the original set of Haar features proposed by Ramirez and Fuentes [10]. We present experimental results that show that our method can attain better results than the original set of Haar features proposed by Viola and Jones [13] for street detection.

2 Related Work

Street detection, has received some attention in recent years. Vosselman and Knecht detect roads matching the average gray value of a road model with pixels of some manually selected segments of the image [15]. With the matches they obtain some parameters that describe the shape and position of the road. The parameters are used in combination with the Kalman filter to predict the position of the complete road.

Doucette *et al.* presents an algorithm for automatic building of road maps in multi-spectral images [2]. The algorithm is inspired by Kohonen's self-organizing map neural network algorithm. First a human-supervised segmentation of roads is performed, then the segmented image is given as input to a K-medians algorithm. The points that represent the road are initialized in a uniform grid. After that they use a minimum spanning tree algorithm to derive the road topology.

Tupin *et al.* presents a road detection method for synthetic aperture radar images [12]. They used a line detector that is based on thresholding. After that, they construct a graph using the interconnected segments. To identify a road, they minimize an energy function associated with the graph.

Yan and Zhao presents a method based on segmentation [16]. The segmentation is performed using region growing based on three rules: gray comparability, width consistency, and continuity. After segmenting, they measure some features such as parallelism and connectivity and also fuse information from different image channels to obtain more precise road detection. Péteri and Ranchin presents a road extractor based on a multi-resolution analysis [8]. They used the Canny

edge detector and a wavelet transform to construct a multi-resolution image representation. Then they use double active contours called double snakes. They fit a street minimizing an energy function that represents continuity, curvature and parallelism. A drawback is that the search has to be initialized near the road to ensure convergence.

Christophe and Inglada presents a road detector based on color and geometric properties of roads [1]. The detector needs user interaction to define the color of the road to be detected. The detector's output is a set of vectors that can be integrated to GIS software. Guo *et al.* presents a method for road extraction on a large set of aerial images [4]. The method is based on detection of road footprints. A footprint describes the geometric characteristics of a possible road segment. To build a road network the authors used a growing algorithm to connect all the road segments.

3 Haar Features

Haar features are based on Haar wavelets, which are functions that consist of a brief positive impulse followed of a brief negative impulse. In image processing, a Haar feature is the difference of the sums intensities of all pixels in two or more contiguous regions. Papageorgiou *et al.* were the first to use Haar features for face detection [7]. They used three types of Haar features of size 2×2 and 4×4 pixels, for a total of 1,734 different features in a 19×19 face image. Viola and Jones proposed a basic set of four types of Haar features that are shown in Figure 1a [13]. The value of Haar feature is given by the sum of intensities of the pixels in the light region minus the sum of intensities in the dark region. Using all possible sizes, they generate around 180,000 features for a 24×24 pixel image. Lienhart and Maydt presented an extension to the basic set with rotated Haar features as shown in Figure 1b [6]. Using a straightforward implementation, the time required to perform the sum of pixels increases linearly with the number of pixels. Viola and Jones proposed to use the integral image as preprocessing to compute the sum of regions of any size in constant time [13]. Each element of the integral image contains the sum of pixels in the original image that are above and to the left of that pixel; using this idea allows to compute a two-region Haar feature using only six memory access and a three-region Haar feature with only eight.

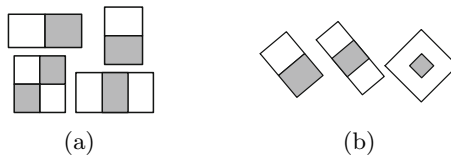


Fig. 1. (a) Haar features introduced by Viola and Jones [13]. (b) Some Haar features variations proposed by Lienhart and Maydt [6].

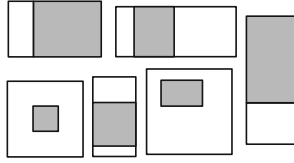


Fig. 2. Asymmetric Haar features used

4 Proposed Algorithm

In contrast with basic Haar features, the features we presented in [10] can have regions with different width or height, but not both (see Fig. 2). By allowing asymmetry, the number of possible configurations for Haar features grows exponentially and is an overcomplete set. For the 6 Haar features shown in Figure 2, there are around 200 million possible configurations for a 24×24 image. Using all the possible configurations is unfeasible, therefore, to deal with this limitation we use a Genetic Algorithm to select a subset of features.

We use specialized detectors to detect streets with different orientations. We build a specialized detector for every 30 degrees in orientation to cover the 360 degree in-plane rotation of each possible street orientation. Each specialized detector consists of a cascade of strong classifiers created with the AdaBoost algorithm used by Viola and Jones [13]. We use a weak classifier based on the C4.5 rule induction algorithm [9] that is associated with only one Haar feature. We only need to train detectors for 0 and 30 degrees; the remaining ten detectors can be created by combinations of 90-degree rotations and inversions of original Haar features as shown in Figure 3. Since a street is symmetric horizontal and vertically, we require only 6 specialized detectors to cover the 360 degree in-plane rotation.

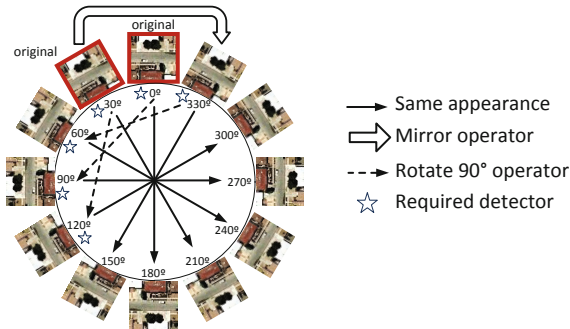


Fig. 3. Specialized detector for streets

4.1 Selecting Haar Features with a Genetic Algorithm

Training a specialized detector using all the possible configurations of the asymmetric region Haar features would be impractical. For instance, using an initial training set of 4,000 examples and all the possible Haar configurations, we would need about 4 months on a 2.5GHz Intel Xeon computer for only one specialized detector. Therefore, to reduce the number of possible Haar features, we use a Genetic Algorithm (GA) to select the width and the height of the regions associated with a feature. In the GA, one individual is a weak classifier (WC) that contains only one Haar feature and is trained with the C4.5 algorithm. The fitness of each individual corresponds to the classification accuracy on an initial training set. We compute a WC for each place on the image and for each type of feature, for a total of 2,431 Haar features. We use a decimal representation that avoids the creation of invalid individuals by crossover. For mutation, we generate an uniformly distributed random value inside the allowed range. We use two point crossover with a deterministic crossover model presented in [5]. This model consists of combining the best and the worst individuals in the population, then the second best with the second worst, and so on. With this crossover model, it is possible to perform a larger exploration on the search space. The best result was obtained with a 10% mutation rate. Using the GA to select a subset of Haar features, we can reduce training time to 6 hours on our 2.5GHz Intel Xeon computer; this corresponds to a 99.8% reduction in time.

4.2 Training a Specialized Detector

A specialized detector is a cascade of strong classifiers (SC). We can create a cascade of SC using the same algorithm presented in [10]. After obtaining a subset of WC with GA, new examples are added to the training set and then the AdaBoost algorithm is used to create a new SC. AdaBoost ends when the minimum detection rate and the maximum false positive rate per layer in the cascade are attained. If the target false positive rate is achieved, the algorithm ends. Otherwise all negative examples correctly classified are eliminated and the training set is balanced adding negative examples using a bootstrapping technique. With the training set updated, all the WCs are retrained and then used in AdaBoost.

5 Experimental Results

We use satellite images taken from Google Earth. We use 20 high resolution images from urban places and manually crop square regions that contain streets. A total of 800 regions were cut and normalized in rotation to 0 degrees and a size of 24×24 pixels. For each region we add 4 variations in angle inside the range of ± 15 degrees for a total of 4,000 positive images. To generate negative samples, we replaced the streets in the 20 original images with other parts of the image that do not contain streets. Then we randomly selected 4,000 regions

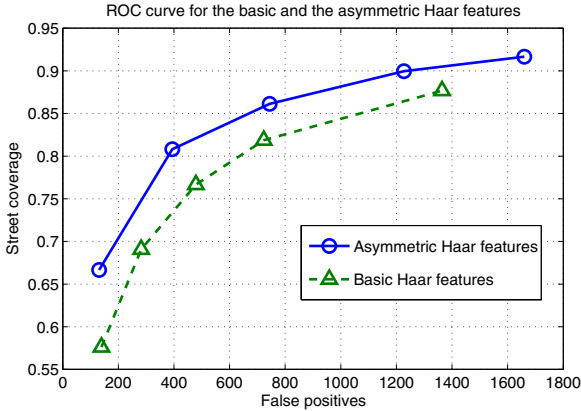


Fig. 4. Comparison between the basic and the asymmetric Haar features

as negative samples for the training set. The training set contains a total of 8,000 examples and can be transformed for any rotation angle in order to create other specialized detectors. In addition, we use 160 scenery images with the streets replaced with image regions without streets to add negative examples using bootstrapping. We trained 2 specialized detectors, one in 0 degrees and other in 30 degrees, the other 4 detectors were created rotating 90 degrees of inverting the Haar features of the original detectors as shown in Fig. 3. For comparison, we also trained specialized detectors using the implementation of the original set of Haar features proposed by Viola and Jones [13] provided by OpenCV. To detect streets in any rotation angle, we combined the output of the 6 specialized detectors. For testing we created a set of 40 satellite images at a resolution of 1280×900 pixels. For each image we manually segmented the street regions to create the ground truth data. The performance of the full detectors was measured as the percentage of the streets that were correctly detected by the detectors and we called it *street coverage*. We found that on average, our detector based on asymmetric Haar features has 35% fewer false positives than the detector based on basic Haar features when they have similar levels of street coverage as we can see in the ROC curve showed in Fig. 4.

Table 1. Comparison of Haar features sets

Haar features set	Number of layers	Total number of Haar features	Training time
Basic	16	0 degrees: 970, 30 degrees: 891	88 hours
Asymmetric	5	0 degrees: 452, 30 degrees: 436	6 hours

Table 1 shows a comparison of the Haar features sets used in our experiments. As we can see, we only need about 7% of the time for training a detector using our proposed approach in comparison with the exhaustive approach used by OpenCV. Also, our detector based on asymmetric Haar features only needs around half of the features to archive a better performance for street detection.

A lower number of features results in a faster detection. Figure 5 shows some results of our street detector.



Fig. 5. Some results for street detection. (a) and (c) Input images. (b) and (d) After use our street detector.

6 Conclusions and Future Work

In this paper we have presented a street detection system based on asymmetric Haar features. According to our experiments, these new features can represent object appearance more accurately than basic Haar features. We used a genetic algorithm to reduce the training time; this allowed us to exploit the expressive advantage of asymmetric features while requiring only 0.02% of the time that would be required using the full feature set. Our detector has on average 11% better performance than the detector based on basic Haar features. Future work will include testing our system in other object detection problems such as car detection in aerial images. In addition, we will perform experiments using other classification algorithms such as FloatBoost and Real AdaBoost.

References

1. Christophe, E., Inglada, J.: Robust road extraction for high resolution satellite images. In: IEEE International Conference on Image Processing, ICIP 2007, October 16-19, vol. 5, pp. V-437–V-440 (2007)
2. Doucette, P., Agouris, P., Stefanidis, A., Musavi, M.: Self-organised clustering for road extraction in classified imagery. *ISPRS Journal of photogrammetry and Remote Sensing* 55(5-6) (2001)
3. Gruen, A., Li, H.: Road extraction from aerial and satellite images by dynamic programming. *ISPRS Journal of Photogrammetry and Remote Sensing* 50(4), 11–20 (1995)
4. Guo, X., Dean, D., Denman, S., Fookes, C., Sridharan, S.: Evaluating automatic road detection across a large aerial imagery collection. In: 2011 International Conference on Digital Image Computing Techniques and Applications (DICTA), pp. 140–145 (December 2011)
5. Kuri-Morales, A.F.: Efficient compression from non-ergodic sources with genetic algorithms. In: Fourth Mexican International Conference on Computer Science, pp. 324–329 (2003)
6. Lienhart, R., Maydt, J.: An extended set of Haar-like features for rapid object detection. In: International Conference on Image Processing, vol. 1, pp. I-900–I-903 (2002)
7. Papageorgiou, C.P., Oren, M., Poggio, T.: A general framework for object detection. In: ICCV 1998: Proceedings of the Sixth International Conference on Computer Vision, p. 555. IEEE Computer Society, Washington, DC (1998)
8. Péteri, R., Ranchin, T.: Multiresolution snakes for urban road extraction from ikonos and quickbird images. In: 23rd EARSeL Symposium Remote Sensing in Transition (2003)
9. Quinlan, J.R.: C4.5: programs for machine learning. Morgan Kaufmann Publishers Inc., San Francisco (1993)
10. Ramirez, G.A., Fuentes, O.: Multi-pose face detection with asymmetric haar features. In: IEEE Workshop on Applications of Computer Vision, WACV 2008, pp. 1–6 (January 2008)
11. Tieu, K., Viola, P.: Boosting image retrieval. *International Journal of Computer Vision* 56(1-2), 17–36 (2004)
12. Tupin, F., Houshmand, B., Datcu, M.: Road detection in dense urban areas using sar imagery and the usefulness of multiple views. *IEEE Trans. Geosci. and Remote Sensing* 40(11) (2002)
13. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proceedings of 2001 IEEE International Conference on Computer Vision and Pattern Recognition, pp. 511–518 (2001)
14. Viola, P., Jones, M., Snow, D.: Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision* 63(2), 153–161 (2005)
15. Vosselman, G., Knecht, J.: Road tracing by profile matching and kalman filtering. In: Automatic Extraction of Man-Made Objects from Aerial and Space Images (1995)
16. Yan, D., Zhao, Z.: Road detection from quickbird fused image using ihs transform and morphology. In: IEEE International Geoscience and Remote Sensing Symposium (2003)