Submitted to the IEEE Computer Society International Conference on Computer Vision and Pattern Recognition, June 2005.

Road Sign Detection and Recognition

Michael Shneier
National Institute of Standards and Technology
100 Bureau Drive, Stop 8230
Gaithersburg, MD 20899

Abstract

Road sign detection is important to a robotic vehicle that automatically drives on roads. In this paper, road signs are detected by means of rules that restrict color and shape and require signs to appear only in limited regions in an image. They are then recognized using a template matching method and tracked through a sequence of images. The method is fast and can easily be modified to include new classes of signs.

Introduction

As with any vehicle, an autonomous vehicle driving on public roads must obey the rules of the road. Many of these rules are conveyed through the use of road signs, so an autonomous vehicle must be able to detect and recognize signs and change its behavior accordingly. This paper describes a fast method for locating and recognizing road signs in a sequence of images. In the United States, the Manual on Uniform Traffic Control Devices (http://mutcd.fhwa.dot.gov/pdfs/2003r1/pdf-index.htm) defines the appearance of all signs and road markings. They divide the signs into classes, each with distinctive colors and shapes. For example, warning signs are orange and predominantly diamond shaped, while regulatory signs are usually red with white writing or white with black writing. We make use of both color and shape in locating signs in image sequences.

There has been a surge in recent years of papers describing road sign detection methods. Many of them segment the signs using color and shape. For example, Piccioli et al. [1] use color and a priori information to limit the possible locations of signs in the image. They then extract edges and look for circular or triangular regions before applying a cross-correlation technique for recognizing the signs. In [2], a redness measure is used to locate stop, yield, and "do not enter" signs. This step is followed by edge detection and shape analysis to identify the sign. Escalera et al., [3], also start with color matching, which they follow with corner detection in which they look for corners in specific relationships that correspond to triangular, rectangular, or circular signs. Classification makes use of a neural network. In their approach to detecting stop signs, Yuille and his colleagues [4] correct for the color of the ambient illumination, locate the boundaries of the signs and map the sign into a frontoparallel position before reading the sign.

Huang and Hsu [5] use shape and color in a wide angle view to locate signs as circular or triangular shapes. They then control the camera to point at candidate sign locations for a closer view which is used to identify the sign based on matching pursuit filters. Another paper that describes actively controlling the camera for sign detection is [6]. Again, the researchers start with color and shape (from edges). They predict the location of the sign and point the camera for a closer view. Signs are recognized and their contents are read by template matching.

A decision tree method is used in [7] to detect and recognize signs without using color. Detection is based on shape using local orientations of image edges and hierarchical templates. The results are sent to a decision tree which either labels the regions by sign type or rejects them. A method to detect speed limit signs is given in [8]. It is based on first using color to locate candidate signs, followed by a multi-resolution application of templates that look for circular regions. Finally, the numbers are read to recognize the signs.

A very different approach is taken by Fleischer et al., [9], who use a model-based, top down approach. Predictions are made of locations in which signs may appear and shape, size, and color are used to specify edge models for the signs in the 3D world. Signs that are found are tracked through subsequent images using a Kalman filter.

Shaposhnikov et al., [10], make use of color segmentation using the CIECAM97 color appearance model. They then use histograms of oriented edge elements to determine the shape of the sign followed by location of the center of the sign. Signs are described by a set of descriptors which are matched with stored models to recognize the signs. Hsien and Chen [11] quantize colors in HSV space and locate candidate signs by projecting regions that may be signs onto the horizontal and vertical axes. Extracted sign regions are matched with templates using a Markov model, and the match with highest rank is regarded as the recognition result.

An impressive study focusing on sign detection in cluttered environments is that of Fang et al. [12]. Neural networks are used to locate regions that may be the center of signs. Both color (hue) and shape features are used. Candidate signs are tracked through time using a Kalman filter and signs are verified by a set of rules concerning the colors and shapes of the regions.

The approach taken in this paper follows the general trend in previous work. Candidate regions are identified by color and pruned using shape criteria. Recognition uses template matching, and signs are tracked over time. The details of the approach are different, and the method has the advantages of being fast and easily modified to recognize new classes of signs.

Detecting the Signs

The program accepts a video stream taken from a camera on the roof of our test vehicle. Because the images are interlaced and the vehicle is moving, there is significant blur between fields in the images. To minimize this, we subsample the images

(a) (b)

Figure 1. (a) shows a frame from the video with a detected sign. (b) shows the binary map for warning signs, with the shape of the sign clearly visible. The other blobs in the image are rejected as signs because they are either too small, are of the wrong shape, or are in a region where signs are not expected.

and use only every second line and column. Signs are detected in a multistage process, starting with segmentation based on color. Because of variation in the ambient illumination, it is not possible to search for regions with specific red, green, and blue values. To get around this problem, we use ratios of RGB colors tailored to different classes of signs (warning, regulatory, informational, etc.). Several other authors suggested using the HSV color space instead of RGB (e.g., [11]),[1]), but we found that ratios of RGB colors worked better and did not require color space conversion. For warning signs, we require that for each pixel the red (r), green (g) and blue (b) values conform to the following.

$$r / g > \alpha_{warn} \;\&\; r / b > \beta_{warn} \;\&\; g / b > \gamma_{warn}$$

For predominantly red signs (stop, yield, no entry), we require

$$r / g > \alpha_{red} \;\&\; r / b > \beta_{red} \;\&\; g / b > \gamma_{red}$$

where α, β, and γ are constants. These constants are determined by sampling the red, green, and blue values of images of typical signs. The precise values of the constants are not critical for sign detection. For additional classes of sign, such as guiding signs on highways (green with white letters) we would define a similar set of constraints, making it very easy to increase the range of signs recognized by the algorithm.

The constraints are applied pixel by pixel to the image and result in a binary image with 1's where pixels are candidates for belonging to a sign. The different classes of signs can all be combined into a single binary image, or each class can have its own binary image. In the latter case, more control is possible in the processing that follows this step. In this work, all the classes were combined into a single binary image.

Following the creation of the binary images, a morphological erosion is performed to get rid of single pixels and two or three dilations are done to join parts of signs that may have become separated due to the presence of writing or ideograms on the signs. Figure 1 shows a scene taken from a video camera and the corresponding binary image for a warning sign (orange). The sign outline is clearly visible in the binary image. We will use the sign outline as a mask in the sign recognition stage.

The next step is to find connected components in the binary images and identify blobs that are likely to be signs. Properties of each component are computed, including the centroid, area, and bounding box. They are used in a set of rules that accept or reject each blob as a sign candidate. The rules require the area of the blob to be greater than a minimum and less than a maximum, the height to width ratio to be in a specified range, and the centroid to be in a restricted part of the image where signs can be expected to appear. The ratio of the area of the blob to the area of the bounding box is restricted to prevent blobs that are too thin from being accepted. Blobs that conform to the rules are considered to be candidate signs and are tracked from image to image. If a blob is seen in five successive frames, it is confirmed as a candidate and goes on to the recognition phase of the algorithm.

Recognition is achieved by template matching. A preprocessing step is first applied to each candidate sign. It masks out the background surrounding the sign which would otherwise interfere with the template matching. We make use of the results of the sign detection phase that already constructed a mask for the sign (Figure 1). Using this mask results in good segmentation of the sign region from the background (Figure 2). The masked candidate signs are scaled to a standard size (48x48 pixels) and are compared with stored signs of the same size. The stored sign templates are taken from video sequences similar to those being recognized. Because there is a lot of variation in the signs, several stored templates may be needed for each canonical sign.

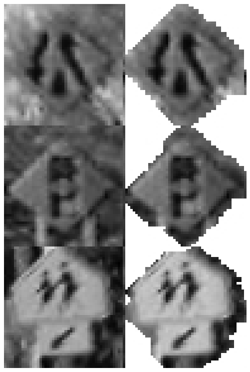

Figure 2 Candidate signs before and after background masking.

Experiments

The algorithm was tested on video sequences recorded with a camera mounted on the roof of a vehicle driving on suburban roads at normal driving speeds. The results reported here are for a total of 23,637 frames containing 92 warning and stop signs. Figure 3 shows examples of recognized signs. Table 1 shows the results for three individual runs and the combined totals, while Table 2 shows the corresponding percentages. As can be seen, the sign detection phase misses very few signs, but at the cost of substantial false detections. In the first data set, this is due to a single vehicle that was in front of the test vehicle in most of the frames. The sign detection algorithm was sensitive to the brake lights of the vehicle, and almost all the false detections are related to this feature. The second data set was acquired in early spring, when the forsythia bushes were in flower. The color of the flowers is very close to that of warning signs, and the bulk of false detections were due to this coincidental color match. All of the false detections are rejected by the recognition phase. The false recognitions in the tables resulted when a correctly-detected sign was incorrectly recognized. The penalty for the extra detections is that the recognition process has to be run on each of them. This slows down the processing, although the algorithm still runs at over 20 frames per second on a 1.6 MHz Intel Pentium Mobile. This is fast enough to interact with the control system in real time.

Figure 3 Examples of recognized road signs

Run	No. Frames	No. Signs	No. Detected	No. Recognized	False Detections	False Recognitions
1	10, 924	39	33	29	19	1
2	9, 439	45	43	38	34	5
3	3, 274	8	5	5	0	0
1, 2, 3	23, 637	92	81	72	53	6

Table 1 Performance of Sign Detection and Recognition

Run	No. Frames	No. Signs	Percent Detected	Percent Recognized	False Detections	False Recognitions
1	10, 924	39	96 %	84 %	76 %	11 %
2	9, 439	45	95.4 %	86 %	75 %	7 %
3	3, 274	8	63 %	63 %	0 %	0 %
1, 2, 3	23, 637	92	88 %	78 %	58 %	6.5 %

Table 2 Percentage results for Sign Detection and Recognition

Conclusions and Future Work

Road signs are deliberately designed and positioned to make it easy for humans to detect them and recognize what they mean. It is still a challenge to enable a computer to perform as well as humans, especially over the full range of possible signs. This paper has covered only two types of sign: warning signs and a subset of regulatory signs. While the sign detection method is robust and accurate, sign recognition suffers because the extracted sign regions are small and often blurred (Figure 4). Actively pointing a camera at the locations of detected signs and zooming in would enable more accurate sign recognition, as in [5] and [6].

While it is straightforward to extend the current method to recognize new signs that have colored backgrounds, the method does not work well for white signs with black writing. This is because many regions in the images may have uniform gray values (e.g., the sky, road surface, and buildings) and, depending on the lighting, the sign may appear lighter or darker than these regions. This means that the detection process either fails or is overwhelmed by non-sign regions. A more structural approach is needed for these signs, perhaps looking for pairs of edges required to match a rectangular shape.

A road sign detection and recognition algorithm has been described that makes use of color and shape to detect road signs in video imagery. Regions that are candidate signs are further processed using a template-matching approach to identify the signs. Information about signs is determined in close to real time, which means that it can be provided to an autonomous robotic vehicle quickly enough to enable it to modify its behavior according to the information on the signs.

Acknowledgements: This work was partly funded by the Army Research Laboratory and by the DARPA MARS program. Their support is gratefully recognized.

Figure 4 Typical detected sign appearance.

Bibliography

[1] Piccioli, G., De Micheli, E., Parodi, P., and Campani, M., "A Robust Method for Road Sign Detection and Recognition," *Image and Vision Computing*, vol. 14, no. 3, pp. 209-223, 1996.

[2] Estevez, L. and Kehtarnavaz, N. A., Real-Time Histographic Approach to Road Sign Recognition. Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, 95-100. 1996. San Antonio, TX.

[3] de la Escalera, A., Moreno, L. E., Salichs, M. A., and Armingol, J. M., "Road Traffic Sign Detection and Classification," *IEEE Transactions Industrial Electronics*, vol. 44, no. 6, pp. 848-859, 1997.

[4] Yuille, A. L., Snow, D., and Nitzberg, M. Signfinder: Using Color to Detect, Localize and Identify Informational Signs. International Conference on Computer Vision (ICCV98), 628-633. 1998. Bombay, India.

[5] Huang, Chung-Lin and Hsu, Shih-Hung. Road Sign Interpretation using Matching Pursuit Method. IEEE International Conference on Pattern Recognition (ICPR2000), 1329-1333. 2000. Barcelona, Spain.

[6] Miura, Jun, Tsuyoshi, Kanda, and Shirai, Yoshiaki. An Active Vision System for Real-Time Traffic Sign Recognition. IEEE Conference on Intelligent Transportation Systems, 52-57. 2000. Dearborn, MI.

[7] Paclik, Pavel and Novovicova, Jana. Road Sign Classification Without Color Information. Proceedings of the 6th Conference of Advanced School of Imaging and Computing. 2000. Lommel, Belgium.

[8] Sekanina, Lukas and Torresen, Jim. Detection of Norwegian Speed Limit Signs. Proceedings of the 16th European Simulation Multiconference (ESM-2002), 337-340. 2002. Darmstadt, Germany.

[9] Fleischer, K, Nagel, H.-H., and Rath, T. M. 3D-Model-Based-Vision for Innercity Driving Scenes. IEEE Intelligent Vehicles Symposium (IV'2002), 477-482. 2002. Versailles, France, IEEE.

[10] Shaposhnikov, Lubov N., Podladchikova, Alexander V., Golovan, Natalia, Shevtsova, A., Hong, Kunbin, and Gao, Xiaohong. Road Sign Recognition by Single Positioning of Space-Variant Sensor Window. Proc. 15th International Conference on Vision Interface, 213-217, 2002. Calgary, Canada.

[11] Hsien, Jung-Chieh and Chen, Shu-Yuan. Road Sign Detection and Recognition Using Markov Model. 14th Workshop on OOTA, 529-536, 2003.

[12] Fang, C.-Y., Chen, S.-W., and Fuh, C.-S., "Road-Sign Detection and Tracking," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 5, pp. 1329-1341, 2003.