An Automatic Traffic Sign Recognizer for Intelligent Vehicles

Maru M. Velas, Gerardo E. Domingo II, Enrico I. Pangan, Erickson T. Miranda
Computer Vision & Machine Intelligence Group
Department of Computer Science
University of the Philippines
Diliman, Q.C.
idnet@updiliman.com

ABSTRACT

In this paper we describe a method for accurately locating and recognizing traffic road signs in real time from a stream of images taken by a video camera mounted on a vehicle. For the computer system to classify the colors of pixels in an image from the stream of images, color segmentation is implemented using the Self Organizing Map (SOM). Using this, pixels are sorted into eight color categories. The output of the SOM is then fed to the Region-Growing Algorithm (RGA). RGA grows regions of pixels labeled as red until all potential traffic signs are bounded. These potential traffic signs are then fed into our Backpropagation Neural Network (BPNN), which outputs the classification of the sign. The primary goal of the researchers is to lay the foundations of an Intelligent Transportation System (ITS) for automated driving. This traffic sign recognizer module is expected to play an important part in ITS.

Keywords

Self Organizing Map (SOM), Region-Growing Algorithm (RGA), Backpropagation Neural Network (BPNN), Intelligent Transport System (ITS).

1. INTRODUCTION

Vision-based Intelligent Transportation System (ITS) deals with the development of autonomous driver-assistance modules using computer vision. This includes automating one or more driving tasks such as: road following (automatic movement along a given path, which includes lane and obstacle detection), platooning (an automatic vehicle following a manually driven vehicle), vehicle overtaking, automatic parking, collision avoidance and driver-status monitoring(Broggi, 1998).

The primary goal of such systems includes road safety and driver efficiency. ITS can use machine vision to detect lane markings, vehicles, pedestrians, road signs, traffic conditions, traffic accidents, and even driver drowsiness. By exploiting ITS technologies, road vehicle systems can be made safer, more efficient, and more environment-friendly (Masaki, 1998).

It is a fact that a lot of road accidents happen nowadays due to the ever-increasing number of vehicles and careless drivers. Some of these are partly caused by unrecognized traffic signs, which drivers sometimes ignore or may have failed to notice.

We present an automatic traffic sign recognizer which locates and recognizes traffic road signs in urban highways. It incorporates color image processing and neural networks to achieve the goal of acquiring pertinent road information.

As a stand-alone system, it informs the driver of the information contained in the traffic sign encountered. This is particularly crucial in avoiding road accidents, with drivers who may have become negligent due to driving fatigue, preoccupation with other matters or human inattention.

The main focus of the research is the development of a software for locating and recognizing traffic signs in a roadway scenery. For this, processing speed is essential because driving happens in real-time and decisions will have to be made in fractions of a second. One of the major difficulties posed by this problem is the presence of extremely varied weather and lighting conditions in the street scenery, affecting the appearance of traffic signs. Skewing of traffic signs posed additional difficulty to traffic sign recognizers.

In this research, we examine the promise of combining Self-Organized Map (SOM), Region-Growing Algorithm (RGA) and the Back-Propagation Neural Network (BPNN) for locating and recognizing Philippine traffic road signs.

It is important to stress that our system was trained for the most commonly found traffic signs in Philippine urban roads and highways. Additional training is expected to give similar performance for traffic signs found in other countries.

2. METHODOLOGY

The program uses Self-Organized Map (SOM) for color segmentation, Region-Growing Algorithm (RGA) for object location and Backpropagation Neural Network (BPNN) for object recognition.

The first activity in the implementation of the research is the gathering of video images of different street environments. Image size frames fall under the size of 320x160. These images are initially digitized into a Quicktime format and are lately converted into an AVI format.

Sample videos of traffic signs and other objects found in the

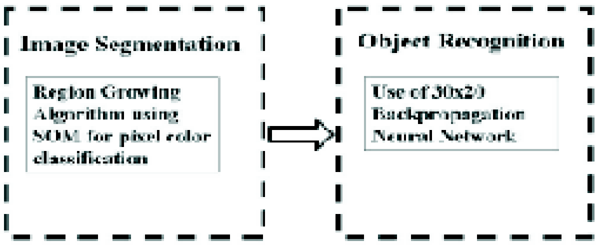

Figure 1: The Methodology of the Application

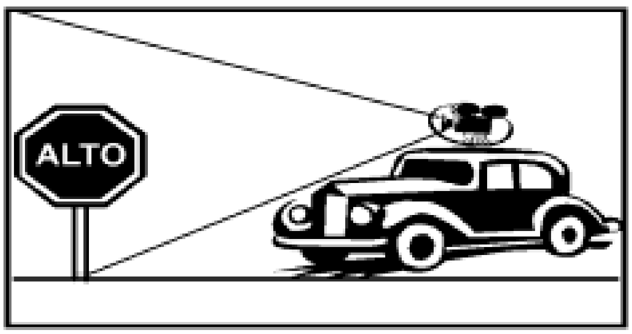

Figure 2: Setup for the Taking of the Stream of Images

street environment were gathered and used in the training of an 8-node SOM.

Figure 1 illustrates the two major parts of the methodology: Image Segmentation and Object Recognition.

The trained SOM distinguishes traffic sign and non-traffic sign pixels. RGA analyzes video file frames and bounds potential traffic signs in rectangles.

The frames from all the video samples are then analyzed using RGA, and the bound potential traffic signs are stored. These traffic signs are tagged as nonsign or as one of the different traffic signs. Afterward, these are used as training and test samples for a neural network that recognizes traffic signs (the object recognition module).

The different subsystems – SOM, RGA and BPNN – are altogether integrated into the Vision-based Driving Assistance System. This system analyzes street environment scenery frame by frame by bounding potential traffic signs.

Figure 2 shows how traffic road signs can be captured by a video camera mounted on top of a vehicle.

3. SYSTEM ARCHITECTURE

We used the Microsoft Developer Studio's Visual C++ as the programming language simulator running under the Windows 98 Operating System. This simulator uses its Windows-

Figure 3: Sample Video File Before the Linear Scanning Process

Figure 4: Sample Video File After the Linear Scanning Process

based user interface.

The experiments are made using three Pentium PCs (a Pentium 133, a Pentium II and a Pentium III) for the completion of the Vision-Based Driving Assistant project. The Pentium III is used for the Back-Propagation Neural Network training and for the actual demonstration run of the program. The Pentium 133 and IIs are used for the design of the application.

4. PIXEL CLASSIFICATION: SELF ORGANIZING MAP

The Self Organizing Map (SOM) is used for color segmentation. Among other methods like the Linear Vector Quantization (LVQ) and the Backpropagation Neural Network (BPNN), the group's devised SOM uses a relatively simple function and is therefore a fast algorithm for color segmentation. Linearly decreasing learning rates and a shrinking neighborhood function are used during the training of SOM. The Red, Green and Blue (RGB) pixel values corresponding to the RGB channels of the video are used for training of the SOM. Each RGB is in their normalized forms.

Figures 3 and 4 illustrate how a No Parking video file can

Figure 5: Another Sample Video File Before the Linear Scanning Process

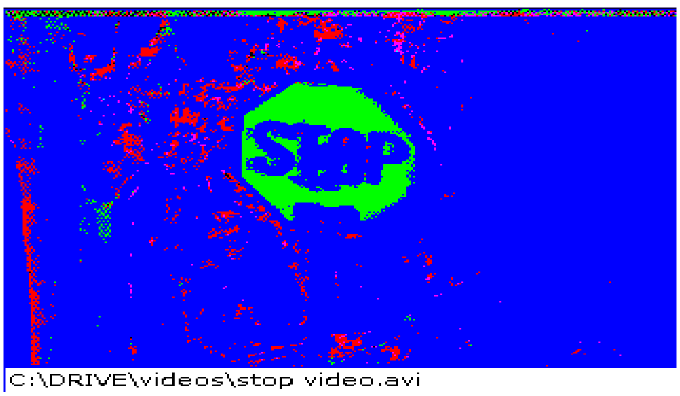

Figure 6: Another Sample Video File After the Linear Scanning Process

be image segmented using SOM.

The SOM outputs eight classifications as represented by the color labels Black, Blue, Green, Cyan, Red, Magenta, Yellow and White. These representations facilitated the visualization of the outputs of the SOM.

The SOM is trained with 1000 randomly-selected pixel samples. Training was stopped when SOM was able to perform color segmentation correctly on the images in the training set. The SOM is now ready for the next process, the Region-Growing Algorithm.

The image is scanned from right-to-left and top-to-bottom fashion submitting the pixels to the trained SOM. The output of the process is a colored bitmap which represents a pixel classified as a red potential traffic sign pixel with the color cyan, and black for pixels which have been placed in a different classification. The output of this process helps monitor the performance of the SOM through visual inspection.

Figures 5 and 6 illustrate how a Stop video file can be image segmented using SOM.

5. OBJECT LOCATION: REGION-GROWING ALGORITHM

RGA implements the object location subsystem. It grows only those regions corresponding to potential traffic signs. The detected regions are then placed within bounding boxes and then passed on the BPNN for classification.

The algorithm checks whether pixels on a video frame are potential pixels of traffic signs in discrete intervals. When it recognizes a pixel as a potential traffic sign pixel, it marks its representation in the image map and checks the surrounding pixels one by one, for whether they are already marked as sign or non-sign. Then, for those which are unmarked, the algorithm checks for whether they are traffic sign pixels or not. The x coordinates of the pixels that are nearest from left, and are farthest to the right and the y coordinates of the pixels that are nearest from top, and are farthest to the bottom are stored. These coordinates represent the potential sign's upper-leftmost and bottom-rightmost corners of the grown region bounded in a rectangle. This process repeats until all the pixels in the whole video frame are analyzed and grown.

All potential signs are stored. Each sign is cross-checked with in a top to bottom manner. If a sign pair passes the set measure of nearness with respect to the sign corners, they will be combined into one sign–the boundary rectangles will be merged, and only one potential sign object will be kept. This represents the new larger object. Potential signs that are too small are discarded and signs that overlap are merged.

The RGA has a main looping block of code that checks for potential traffic sign pixels found. This loop calls external functions. One function returns the RGB values of a pixel on the current open video frame, given its coordinates. SOM returns the color classification of a pixel given its RGB values. The loop calls another function that classifies a pixel as a potential traffic sign.

RGA functions to combine potential signs if they are near to each other or if they are on top of each other. It also eliminates signs that are too small.

Given the coordinates of a potential traffic sign pixel and its color, a function is built to grow region of potential traffic sign pixels around the pixel in question. It puts all traffic sign pixels in a stack and analyzes each one's surrounding pixels. Afterwards, it gets the corners of the grown region, and stores them in an object. It combines this object with the another objects in the potential signs list, if it passes a measure of nearness with that object.

6. OBJECT RECOGNITION: BACKPROPAGATION NEURAL NETWORK

The three-layer fully-connected feedforward architecture of a Backpropagation Neural Network (BPNN) is implemented to achieve object recognition part. This method is chosen based from the good results of its training in the Face Recognition Machine Problem implemented in UNIX. The modified BPNN (implemented in Visual C++) is trained for various number of epochs, different learning rate and momentum

Figure 7: Sample Collection of Binary Sign Files used in BPNN

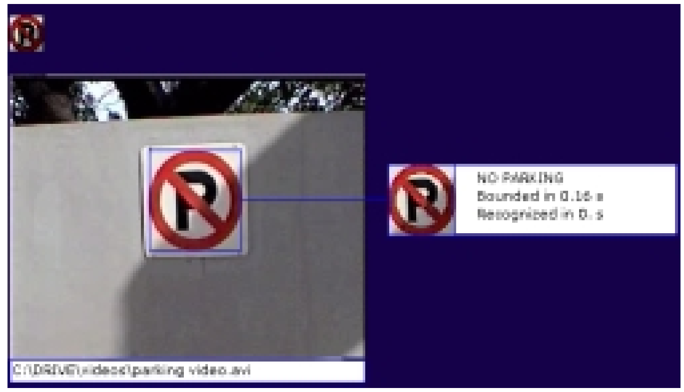

Figure 8: A Detected and Recognized No Parking Sign

parameters to achieve recognition accuracy

The 30x20 neural network accepts input of binary image files seen in figure 7. These binary files are carefully selected to fit the required number training and validation sets to correctly classify and recognize the different types of traffic signs enumerated in table below. The training set consists of 308 images. There are 140 and 148 images in the validation sets.

These signs are named correspondingly to organize the image sets for training and validation. This naming convention easily keeps track of signs detected and recognized from the output image lists of the BPNN. BPNN outputs two sets of image lists for every set. The first list consists of all images that the BPNN is able to correctly recognize as traffic signs. The other list consists of the images it failed to recognize as traffic signs.

Symbol	Meaning
AL	Ascending Line
DL	Descending Line
LS	Lower-half of Stop
MT	Merging Traffic
NE	No Entry
NH	No Horns
NP	No Parking
NS	Non-Sign
OW	One way
SS	Stop Sign
TA	Tow Away Zone
US	Upper-half of Stop
ZN	Zone Line

Table 1: Traffic Sign Naming Convention and their Meanings (as used in the training and validation sets)

7. EXPERIMENTS AND EXPERIMENTAL RESULTS

Figure 9: A Detected and Recognized Stop Sign

An analysis of all the frames of the sample AVI files with the object location module showed that the median of the time intervals taken by the module to analyze a video frame is 0.16 seconds. 0.11 and 0.22 seconds are the next most frequent time intervals. 0.27 seconds is the ceiling time interval for the different frames (based on the speed of a Pentium III processor).

Figures 8 and 9 illustrates a No Parking and Stop Sign Detection and Recognition using the results of SOM, RGA and BPNN.

Visual inspection of the bitmap outputs showed that, on the average, red pixels of traffic signs are classified correctly. However, red pixels of signs that are very far from the mounted camera are not classified as potential traffic sign pixels. Incorrect classification of pixels is also reflected through thinner cyan regions in the bitmap outputs than expected from the corresponding traffic signs in the video frame inputs. Some stray pixels in the background, particularly specks from trees were classified as red potential traffic sign pixels. In several frames, it is noticeable that only pixels from a few disjoint parts of true traffic signs are classified as red potential traffic sign pixels. RGA compensates imperfection in the pixel classification.

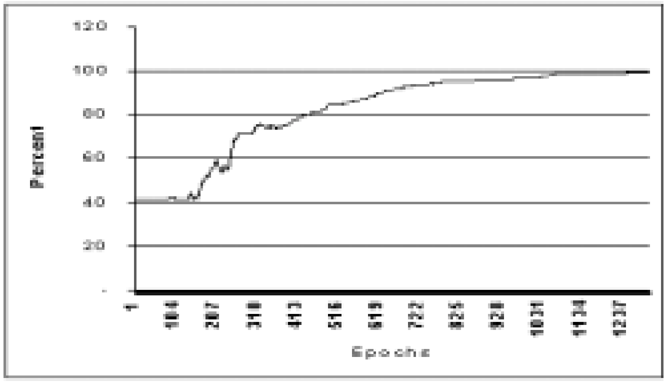

Figure 10: Performance of the BPNN on the Training Set

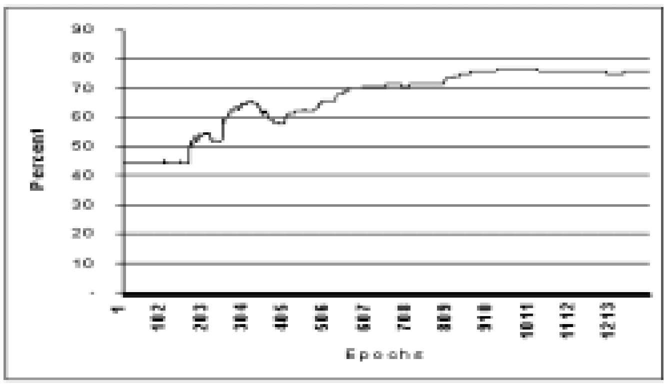

Figure 11: Performance of the BPNN on the Validation Set

Growing/binding of potential traffic signs using the Region Growing implementation is done for frames of the different AVI file samples. To indicate the position and boundary of the potential traffic signs, rectangles are drawn around these objects on the bitmap representations of the video frames.

Visual inspection showed successful binding of potential traffic signs in most frames. In some frames, even though the true traffic signs are bound correctly, some regions that do not represent true traffic signs are also bound as candidate traffic signs. Several frames showed fragmentation of traffic signs. Disjoint pieces of the true traffic signs in these frames are bound separately.

Different sizes of the Backpropagation Neural Network have been tested, and a 30x20 neural network is chosen for the Vision-based Driving Assistant application.

The graphs presented here outline its performance. Figure 10 shows the performance of the neural network in the training set. Figures 11 and 12 show its performance for the validation sets.

The y-axes for graphs represent the percentage of the image samples classified correctly by the neural network. The x-axes represent the number of epochs for which the neural network has so far been trained. (An epoch is an interval wherein all of the training and validation set samples have been analyzed by the neural network.)

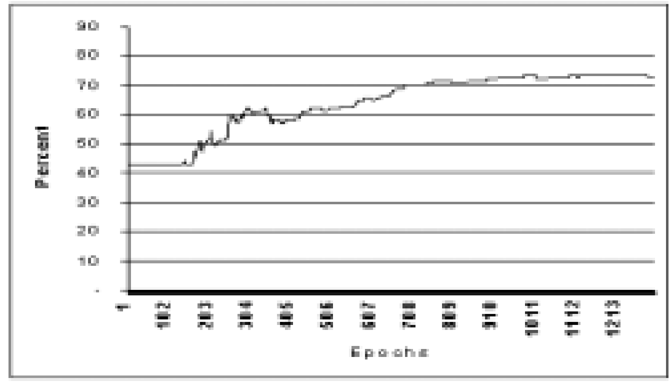

Figure 12: Performance of the BPNN on the other Validation Set

The trained performance of the network is 99.35% recognition rate for the training set and 75.71% and 72.97% for the validation sets, respectively. The learning rate and momentum for training of the neural network is periodically reduced from a starting value of 0.6 because of observed instability in the recognition rate for constant values of learning rate and momentum as training progressed. Even with this adjustment of variables, fluctuations with regards to recognition rate have not been avoided. The final value for both learning rate and momentum is 0.7.

learning rate	momentum	epochs
0.3	0.3	1 to 160
1.5	1.5	161 to 240
1.0	1.0	241 to 350
0.7	0.7	351 to 440
0.8	0.8	441 to 470
0.6	0.6	471 to 740
0.7	0.7	741 to 1280

Table 2: Learning Rates and Momenta for Corresponding Epochs

Many hardware constraints especially on image capture prevented the research experiments from being more accurate simulations of the target environment that is driving. We therefore recommend the following hardware and hardware set-up improvements.

a. The use of better image-capture set-ups A video camera with a wide angle lens that approximates the angle of vision of a human driver, or the use of two cameras mounted on opposite sides of the vehicle will provide a view of two sides of the street, and capture a view of the traffic signs even if they are already near the vehicle.

b. The incorporation of the TV Video in/out facility. The

TV Video in/out will allow the simultaneous capturing and analysis of street environment images. With this, the Vision-based Driving application will be tested while mounted on a real v ehicle and cruising on a real highw ay.

c. The use of LV Q for pixel classification is also a recommended alternative for color training. With this method, training of the representative vectors can be controlled because the expected classifications of the pixels are already stated. The use of more representative (codebook) vectors for the different classifications, instead of just one, may also improv e the classification of pixels.

8. CONCLUSION

A major strength of this research and its implementation, is its quic k processing time. This is essential for a dynamic en vironment such as driving. Space requirements are modest at 900 Kb. Extension of the project to include specifications for blue traffic signs, and other traffic signs is easily done because of the robustness of the methods. Neural netw orks, a primary component of the application, fails gracefully, and is able to express very complicated functions.

9. ACKNOWLEDGEMENTS

We thank our advisers, Prof. Pros C. Naval Jr. and Mr. Arnold M. Putong for the stimulating discussions and helpful insights in the development of this researc h project and paper.

10. REFERENCES

[1] A. Broggi, "Vision-Based Driving Assistance," *IEEE Intelligent Systems*, 1998.

[2] A. Dela Escalera et. al., "Road Traffic Sign Detection Detection and Classification," *IEEE*, 0278-0046/97, June 1997.

[3] D. Ghica et. al., "Recognition of Traffic Signs by Artificial Neural Netw ork," *IEEE*.

[4] U. Frank e et. al., "Autonomous Driving Goes Do wn town," *IEEE Intelligent Systems*, 1998.

[5] D.S. Kang et. al., "An Invariant Traffic Sign Recognition System Based on Sequential Color Processing and Geometrical Transformation," *IEEE*, 0-8186-6250-6/94, June 1994,88-93.

[6] D. Kellmeyer and H. Zwahlen, "Detection of Highw ay Warning Signs in Natural Video Images using Color Image Processing and Neural Netw orks" *IEEE*.

[7] J. Kim et. al., "A Design and Characteristic Analysis of the Motion Base for Vehicle Driving Simulator," in *IEEE International Workshop on Robot ad Human Communication*, 0-7803-4076-0, 1997.

[8] R. Lichauco, "The Philippine Road Rules and Regulations Handbook," *The Green Ads Gr oup* 1993.

[9] W. Lu, "Recognition of Traffic Signs Using a Multilay er Neural Netw ork," *IEEE*.

[10] M. Marsella et. al.,"Neural Techniques for Image Segmentation," *IEEE*, 0-8186-8548-4, 1998.

[11] I. Masaki, "Machine-Vision Systems for Intelligen t T ransportation Systems," *IEEE Intelligent Systems*, 1094-7167/98, 1998.

[12] U. Regensburger and V. Graefe, "Visual Recognition of Obstacles on Roads," *IEEE*.