VISUAL INVENTORY OF ROAD SIGN NO-BLOCKING OF PASSAGEWAY

Luca Lombardi and Roberto Marmo and Andrea Toccalini

University of Pavia
Department of Informatica and Sistemistica
via Ferrata 1, 27100 Pavia, Italy http://vision.unipv.it
email: roberto.marmo@unipv.it

KEY WORDS: Classification, identification, image, inventory, mapping, recognition

ABSTRACT:

This paper describes a method to detect and identify the Italian road sign for no-blocking of passageway. The approach detects the sign in the image by color processing and multi-layer perceptron neural network. The approach locates the region of interest within the image, using color segmentation, then the signal of restricted no stopping is identified using shape and color information, and finally the text label is recognized with a state diagram and test on the shape of the words. It is important to examine also the text label passo-carrabile as an additional test and to differentiate this specific sign among all other prohibition signs. The obtained results show the feasibility of the system.

1 INTRODUCTION

Italian road sign no-blocking of passageway, in Italian language also called passo-carrabile is composed by different sizes, but the standard model (Fig. 1) is 24 cm wide and 43 cm long, with a restricted no stopping sign of 10 cm radius. The plate is painted only with red, blue, white and black colors. This specific sign gives the order to keep clear a transit zone, it forbids parking in a lateral area of the street located into town, so other objects with similar colors can make more difficult the recognition.

It is required to pay the fee to the municipality in order to install the sign on the exit from private area to street, to avoid cars parked on the private area that can make impossible to access from private area to street. In case of fee paid to municipality the Police removes the car that obstructs the passageway.

The row on the bottom contains the code of permission from municipality and the year which corresponds the fee.

Figure 1: The road sign passo-carrabile.

Inventory and localization are performed by GIS based on georeference and visual content information within the sign plate. It is possible to install this software as module in mobile mapping vehicle, measuring vehicle in moving traffic carries that make possible mobile mapping data acquisition. Therefore, a mobile mapping vehicle can locate each sign and related GIS information regards km and meter at the point along the street where is the plate, in order to check if the private area paid the fee by searching on the database if a record occurs at the same point and with the same code of permission.

The image processing techniques (Forsyth and Ponce, 2003) allow to extract from digital image a set of numerical features, expressed as coded characteristics of the selected object, and used to differentiate one class of objects from another. In the following sections we explain the details of these techniques applied to this specific sign.

2 SIGN LOCALIZATION

Our vision system is based on a single digital camera mounted in the right hand of the vehicle, bitmap color images of frame size 640x480 pixel, RGB 24 bit pixel, frame rate of 5 frames per second on a standard PC. Due to high quality of images, the median filter and a neighborhood of 3x3 pixels are sufficient to remove small distortions without reducing the sharpness of the image.

The Italian Highway Code requires to install the plate in a way that it is possible to note the sign when a driver stop the car to park along the street, so the plate is well visible and the plate is located always in front of the camera.

After a sub-sampling for de-interlace (Fang et al., 2004) in order to reduce the noise based on step of 2 rows, the image size is 320x240 pixel.

It is important to delimit a small area of interest, minimizing the possibility to get noise from undesired objects. Due to location of the sign, for every image we consider as region of interest the area of 40x200 (Fig. 2).

In some case the plate are located behind bars of gates (Fig. 3) and we can not analyze the entire plate as unique shape. Another problem regards the high number of visual data in image and to obtain a fast computation, that can make possible to increase the velocity of mapping van, it is required to select the relevant data.

This approach make possible a sort of sampling of the image: we localize a set of subwindow of size 11x11 pixel and the distance between each subwindow is 4 (Fig. 4).

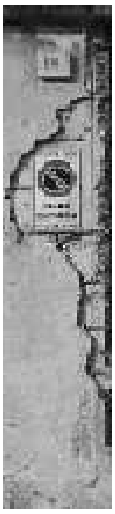

Figure 2: A plate and the region of interest.

Figure 3: The sign passo-carrabile behind a gate.

Figure 4: Zoom of subwindows on the plate.

For each subwindow we extract the RGB color components as Red, Green, Blu in this way:

1. if $R > 150$ and $R - B > 40$ and $R - G > 40$ we have a red pixel;
2. if $R > 150$ and $B - R > 40$ and $B - G > 40$ we have a blu pixel;
3. if $(R + G + B)/3 > 150$ we have a white pixel;
4. if $(R + G + B)/3 < 50$ we have a black pixel;
5. in other case the color pixel is not relevant and it is classified as no useful color.

These thresholds have been estimated experimentally. For each subwindow we obtain 121 values related to color for each pixel and we obtain 5 values related to number of pixel for each color. A feature is a numerical characteristic and these values compose the vector of features related to sign.

Now it is relevant to classify the features in order to decide if the sign occurs in the image. Statistical pattern recognition study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make sound and reasonable decisions about the categories of the patterns (Jain et al., 2000).

Due to difficult to classify the pattern using 121 features, we try to use the 5 features from color information but is difficult to well separate features of sign from features of no sign based on statistical classifier. Therefore, we use the neural networks to classify the number of pixel for each color.

Neural networks are models to express knowledge using a connectionist paradigm inspired by the working mechanisms of the human brain (Bishop, 1998, Jain et al., 2000). The neural networks classify a pattern of numerical features after a training on a specific data set by taking into account all extracted features, or a part of them. Each output of the neural network is the probability to classify the features as input belonging to a specific class. We assign the input to the class associated with the highest value to the output from the network.

The neural network used is a multi-layer perceptron composed by two-layer feed-forward network. The neurons are grouped into input, output and hidden (i.e. those units that are neither input nor output) layers. Each neuron of a given layer is connected to all neurons of the next one. The input layer has five neurons corresponding to the analyzed color, the hidden layer has four neurons, and the output layer has two neurons corresponding to identified or not identified subwindow of sign. Supervised training must be used whereby the network learns from a training set consisting of features input and the desired class outputs. The hidden units use the hyperbolic tangent as activation function and the output units use the softmax activation function, so it is possible to interpret the output values as probabilities. We use the cross-entropy error function and the quasi-Newton minimization error function. These parameters are established by varying the number of neurons in the hidden layer: the chosen architecture reaches a small training error.

To create the neurale network we consider images with sign painted on different weather condition and images without sign but with percentage of colors that are similar to percentage of color on the sign. The training set is composed of 240 images, 120 for each class of sign and no sign. The validation set is composed of 80 pixel, 40 for each class of sign and no sign. We use the early stopping method to control the effectiveness of the training step. The test set is composed of 80 images, 40 for each class of sign and no sign. As test error we have only 1 case of subwindow classified as sign instead of no sign, due to high percentage of red color on the subwindow. Another criterion is to calculate the difference between the values of the two output, we always obtain a difference of at least 0.2.

In this way we recognize the sign by classification of color information from subwindows instead from the entire shape of sign, a sort of local classification versus global classification.

At this step we obtain a black/white image where the white pixel corresponds to subwindows classified as part of sign and black pixel corresponds to subwindow classifieds as part of no sign (Fig. 5). It is possible to note that the white results it is very similar to shape of sign, but it is composed by parts that are not connected as unique shape. We use the morphological operator dilate, method of processing digital images on the basis of shape (Forsyth and Ponce, 2003), with structural element 3x3 to join all the white pixels as unique shape (Fig. 6). A labelling algorithm (Forsyth and Ponce, 2003) locates and counts the shapes composed of white pixels. The sign is found in case of 1 white area composed by number of pixel in [40,600] and eccentricity less than 0.6. The eccentricty of an object is a numerical feature related to the ellipse that has the same second-moments as the region that contain the white object. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1 that are degenerate cases, an ellipse whose eccentricity is 0 is actually a circle, while an ellipse whose eccentricity is 1 is a line segment. The values were determined by analysis of a wide variety of images.

We acquire the image for next analysis in case of 3 images with positive response on sign localization.

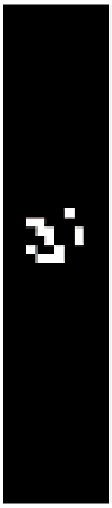

Figure 5: Result of analysis on color of subwindos related to Fig. 2.

A video sequence af about 20 km has been acquired on road using a digital camera located in front of car at velocity 50 km/h and good weather condition. We use 200 images as test set composed by wide variety of images. No error occurs, a relevant factor is the radius: the minimum radius measures only 11 pixel.

As first step on the whole image, this approach selects the area that will be analyzed more accurately in order to decide if discard the image or continue with more in-depth analysis.

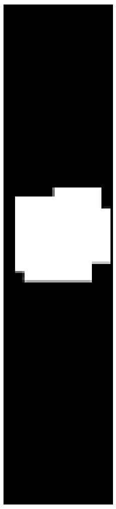

Figure 6: Result of morphological operator related to Fig. 5.

The approach is based on three step. First, the program performs a rapid analysis of the whole image, and selects the area that will be analyzed more accurately, according to the distribution of the sign colors. Next step analyzes the selected region, in case of negative response the approach turns back to the first level to search a new region of interest. If the previous constrains are satisfied, we use the Hough Transform to find red and blue circles, which correspond to circle sof the prohibition signal. In case of circles are detected and they have a common center, we use specific measurements to estimate the size of the sign. If results are admittable, the approach continues with more detailed test, so we apply a pattern matching of the prohibition signal with a model, and the analysis of the text label passo-carrabile with of specific test.

3 COLOR SEGMENTATION

At this step we use the entire image of 640x480 pixels. Working with the RGB color space is high sensibility to lighting conditions, so we study relations between color components (Escalera and Moreno, 1997, Zadeh et al., 1998, Bnallal and Meunier, 2003, Sekanina and Torresen, 2002). Converting the RGB space to HSV (Hue, Saturation, Value) allows to gain a better control on chromatic variations of each hue (it is a separated component), but the computational cost is prohibitive, because the space transformation is nonlinear. HSV space is very tipic (Escalera et al., 2003, Vitabile et al., 2001, Gavrila, 1999, Liu and Ran, 2001, Vitabile et al., 2002), but some studies shows that Hue component changes considerably according to distance, brightness and sign age.

Therefore, we decided to use the RGB model to reduce execution time, also because we analyze only two colors (red and blue) and these are represented by two separately channels in RGB space. White color can be easily obtained putting together the three components.

Three color image histograms (Fig. 7) are analyzed to find a region with a sufficient number of colored pixels as blue, red, white on the same row and also on the same column (Piccioli et al., 1994).

4 ANALYSIS OF REGION OF INTEREST

The algorithm for circles detection works on binary images, that is the blue and red masks of the inspected area (Fig. 8).

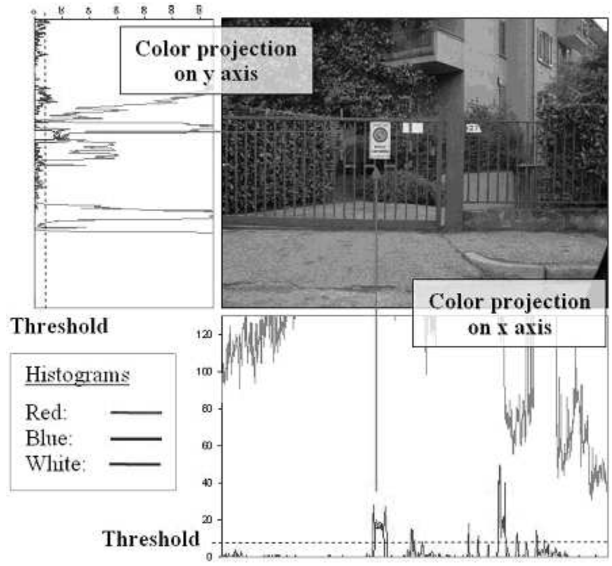

Figure 7: Extraction of region of interest from the image using histograms of colors red, blue and white.

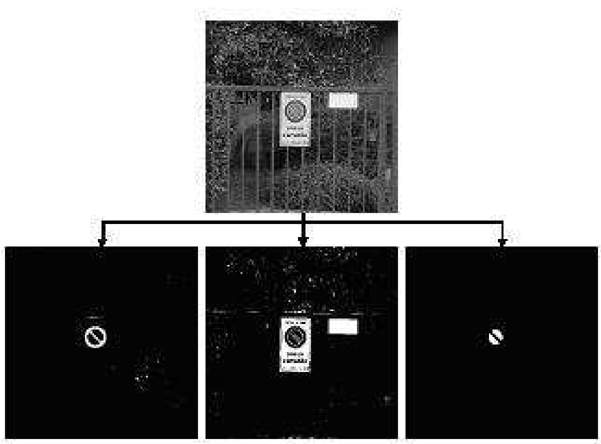

Figure 8: Color masks corresponding to red, white and blue.

We refine the colors segmentation, adjuste the mean brightness of the region. A gaussian filter (mean value among adjacent pixels) removes noise and spare pixels. The edges of the two binary images are extracted by isotropic operator, which consists in the convolution of the image with 3x3 masks.

Hough transform (Forsyth and Ponce, 2003) then is performed to locate red and blue circles. We do not know the radius of these circles, so we search in a three-dimensional parameter space to find the three values (x_c, y_c, r) corresponding to circumference. We apply the Hough method either on red and blue circles; it allows to detect the centre of the signal with good result. Using these results we want to find exactly position and size of signal painted on the sign panel.

Starting from the centre found by Hough, we execute four analysis (right, left, up and down directions) and we store coordinates of color changes, from blue to red and then to white. So we can identify the dimensions of the three different regions within the signal, from these values we can easily detect the radius of the blue and the red rings and obtain the radius of the signal (Fig. 9).

At this point we know exactly the position and the dimension of the signal, so we can apply a pattern matching with a sample template (Fig. 9).

Every colored pixel is altered to its pure tonality (for example: R=255, G=0, B=0 for every red pixel). We obtain good results even when road sign is dirty or affected by bad lighting.

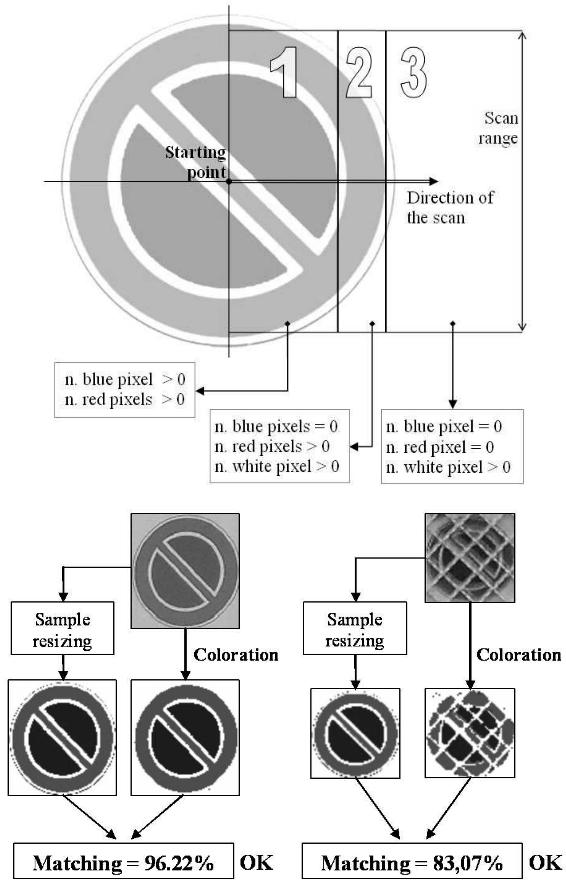

Figure 9: Extraction of sign dimensions: on upper horizontal scan to right direction, on lower examples of pattern matching in case of entire sign (left) and in case of sign behind a gate (right).

Finally, we apply a further control restricted to no stopping at any time sign, to identify and exclude this case that is very similar to restricted no stopping sign. The main difference is a double red strip instead of a single one. These signs are able to pass the pattern maching test with positive result. The approach consists in counting the number of blue pixels on the diagonal placed at +45 degree. When the signal has a double red strip we obtain that this number is equal to zero, in the other case the number of blue pixels always be major then zero.

5 TEXT EXTRACTION

The text label passo-carrabile is located in a specific region, under the signal of no-waiting on the bottom of the sign. It is possible to identify characters using an O.C.R. method, but in this case we realized that there is no need to recognize each letter because the label is fixed a priori by Italian Highway Code.

The image is binaryzed according to a threshold value, to simplify next operations like characters extraction and testing. We use dynamic parameters to perform these adjustments to the image: they are related to the mean brightness of the region analyzed.

To identify the exact dimensions and positions of the two words we execute a vertical scanning of the binary image, analyzing the black projection on the y axis (Miura et al., 2000). Working with a finite state machine we can detect both position and height of the two words (which are always one above the other), as we examine the percentage of black pixels located on each row. Sometime the sign is placed beyond a gate of the house and we obtain a noise caused by gate. We can find out this error performing the previous analysis, and we achieve good results even if the sign is corrupt in this way.

We can identify the error ϵ using informations from the first derivative of y projection, by the incremental ratio: $y'(i) = y(i+1) - y(i)$ where y vector contains, for each element i, the count of black pixels situated on the image row i; vector y' represents its derivative, i.e. the variation of black pixels between two adjacent rows. Examining the vector y', we can note positive peaks corresponding to upper boundaries of words and negative peaks on words ends. Our technique consists in detecting these peaks (so locating where the words are) and extracts the minimum value $y(i)$ among all rows i, outside these words, where $y'(i) = 0$. Briefly, we calculate the fixed error as the minimum number of black pixels located on white rows; we also assume that this noise vertically involves all the image (Fig. 10). So, we obtain two images which fits exactly the two words we have to examine.

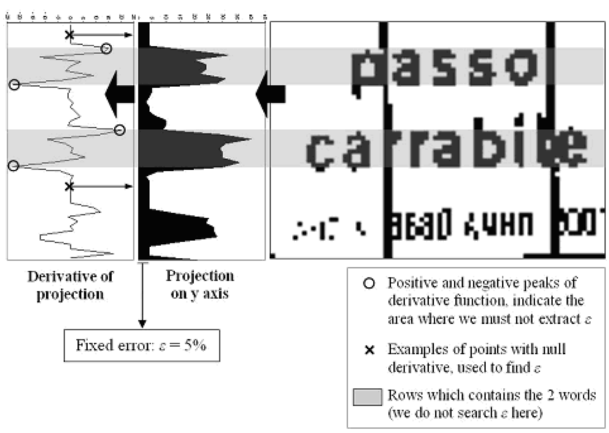

Figure 10: Extraction of a fixed noise, using the derivative of projection on y axis.

Using horizontal scans on the two words separately we can extract the number of characters, their dimensions and width analyzing the count of black pixels on each column of the image (Fig. 11). We use again a state diagram to describe the behavior of a system. State diagrams describe all of the possible states of an object as events occur. Each diagram usually represents objects of a single class and track the different states of its objects through the system. We use the transition from states 'character' and 'space' that is defined by the percentage of black pixels on each column.

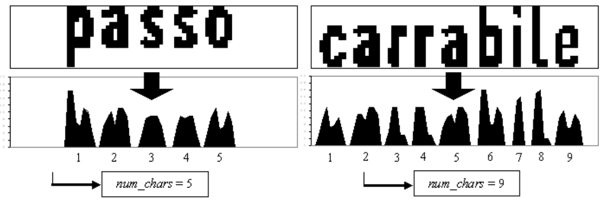

Figure 11: Projection of black pixels on x axis.

We perform some specific test that allows only a little tolerance of dissimilarity from the model of a standard sign. The identification of the label is verified successfully in case of sufficient number of tests have positive results, so we can also remove a high number of false positives (thanks to the low tolerance of each test). Briefly, this method allows to recognize the label passo-carrabile if it has enough common features with a sample image and using this tests:

1. analogy of heights of the 2 words;
2. word 'carrabile' larger then word 'passo';
3. width of the 2 words proportionate;
4. all characters of 'passo' with adequate width;
5. all characters of 'carrabile' with adequate width;
6. count of characters from 'passo' equal to 5;
7. count of characters from 'carrabile' equal to 9.

Each test has a different weight, relative to its importance for the recognition; if we still have an uncertain result after these 7 tests, we apply other 6 verifications:

8. width of word 'passo' proportional to width of the sign;
9. height of word 'passo' proportional to height of the sign;
10. matching between the 2 characters 's' from the word 'passo';
11. matching between the 2 characters 'r' from the word 'carrabile';
12. matching between the word 'passo' and a sample image;
13. matching between the word 'carrabile' and a sample image;

6 EXPERIMENTAL RESULTS

Results produced by the developed system are satisfying: we recognize correctly the signs of passo-carrabile under different lightness conditions and also with high noise. Test set is composed of 420 images and 13 frame per second. The application runs on a computer with a Intel Pentium 4 processor at 3GHz and 512MB of RAM DDR memory at 400MHz. The mean execution time is 74 ms.

Text can be correctly detected until a distance of 11.5 meters, characters are 2 pixels wide and 3 pixels high and the minimum sign radius measures only 11 pixel.

Accuracy is equal to 98.6%. As we can see, the program works very good with false positives: thanks to the low tolerance of final detect test on the label passo-carrabile we do not have any false positive. The approach is robust against noise and occlusions of the signal, as we recognize the road sign even in very prohibitive circumstances (Fig. 12, 13). The few failures are caused by too heavy noise on the writing or excessive faded colors.

The testing has been performed acquiring images at different distances from the sign, from 1 meter to 10 meters. Sometimes the road sign is placed far from the acquisition source, so we need high image resolution for a correct text recognition.

Figure 12: Examples of right results.

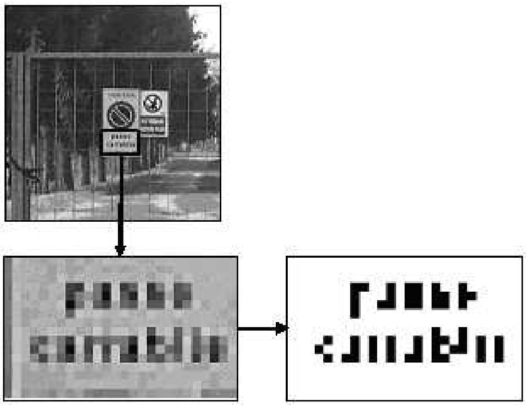

Figure 13: Example of right detection of plate and label with deformed shape of characters. Note the other prohibition plate close to sign and the distorted shape of characters

7 CONCLUSIONS AND FUTURE WORK

In this paper we have proposed an approach based on image processing, color processing and neural networks for the detection of Italian road sign no-blocking of passageway in colour images distinguishing between other similar prohibition signs.

Experimental results show the effectiveness of the approach implemented even in presence of perspective deformation, different conditions of light, partially hidden signs behind gates. This percentage is satisfactory for this preliminary application of the proposed methodology in mobile mapping van.

The codes have been written in C language as part of a package, which can be used and extended for future applications. Future works include the integration with a GIS system.

REFERENCES

Bishop, C., 1998. Neural networks for pattern recognition. Clarendon Press, Oxford.

Bnallal, M. and Meunier, J., 2003. Real-time color segmentation of road signs. IEEE Canadian Conference on Electrical and Computer Engineering, Montral, Qubec, Canada pp. 1823–1826.

Escalera, A. and Moreno, S., 1997. Road traffic sign detection and classification. IEEE Transactions on Industrial Electronics 44(3/1), pp. 848–859.

Escalera, A., Armingol, J. and Mata, M., 2003. Traffic sign recognition and analysis for intelligent vehicles. Image and Vision Computing 11, pp. 247–258.

Fang, C., Fuh, C., Yen, P., Cherng, S. and Chen, S., 2004. An automatic road sign recognition system based on a computational model of human recognition processing. Computer Vision adn Image Understanding 96, pp. 237–268.

Forsyth, D. and Ponce, J., 2003. Computer vision : a modern approach. Prentice Hall, Upper Saddle River.

Gavrila, D. M., 1999. Traffic sign recognition revisited. Proceedings of the 21st DAGM Symposium fr Mustererkennung, Bonn, Germany pp. 86–93.

Jain, A. K., Duin, P. and Mao, J., 2000. Statistical pattern recognition: a review. IEEE Transactions on pattern analysis and machine intelligence 22, pp. 4–37.

Liu, H. X. and Ran, B., 2001. Vision-based stop sign detection and recognition system for intelligent vehicles. Transportation Research Record 1748, pp. 161–166.

Miura, J., Kanda, T. and Shirail, Y., 2000. An active vision system for real-time traffic sign recognition. Proceedings of IEEE Conference on Intelligent Transportation System pp. 52–57.

Piccioli, G., DeMicheli, E. and Campani, M., 1994. A robust method for road sign detection and recognition. Proceedings of Third European Conference on Computer Vision pp. 495–500.

Sekanina, L. and Torresen, J., 2002. Detection of Norwegian speed limit signs. Proceedings of 16th European Simulation Multiconference, Darmstadt, Germany pp. 337–340.

Vitabile, S., Gentile, A., Dammone, G. and Sorbello, F., 2002. MLP neural network implementation on a SIMD architecture. Lecture Notes in Computer Science 2486, pp. 99–108.

Vitabile, S., Pilato, G., Pollaccia, G. and Corbello, F., 2001. Road Signs Recognition Using a Dynamic Pixel Aggregation Technique in the HSV Color Space. Proceedings of 11 International Conference on Image Analysis and Processing, Palermo, Italy pp. 572–577.

Zadeh, M., Kasvand, T. and Suen, C. Y., 1998. Localization and recognition of traffic signs for automated vehicle control systems. Conference on Intelligent Transportation Systems, Pittsburgh, PA pp. 272–282.