A NEURAL MACHINE VISION MODEL FOR ROAD DETECTION IN AUTONOMOUS NAVIGATION

Victor-Emil NEAGOE[1], Cristian-Tudor TUDORAN[2]

În această lucrare este prezentată o abordare originală privind soluționarea problemei identificării direcției de deplasare a unui robot vehicular autonom prin utilizarea clasificatorului cu module neurale autoorganizabile concurente (MONACO) - Concurrent Self-Organizing Maps (CSOM), model introdus de primul autor. Pentru comparație, sunt evaluate de asemenea și performanțele obținute cu ajutorul altor trei clasificatori uzuali. Direcția ce urmează a fi identificată este cuantizată în trei clase: înainte, stânga și dreapta. Sunt prezentate rezultatele obținute prin simularea pe calculator a modelului propus, precum și implementarea practică pe un robot mobil (mașină-jucărie telecomandată) a acestui model.

This paper presents an original approach for visual identification of road direction in autonomous vehicle navigation using a neural network classifier called Concurrent Self-Organizing Maps (CSOM), previously introduced by first author. For comparison, we also evaluate the performances of three other usual classifiers. The path to be identified has been quantized in three output directions: straight ahead, left and right. We present the experimental results obtained by computer simulation (using a data set of 210 road images from the CMU VASC Image Database), as well as a real-time neural path follower based on CSOM model, implemented on a mobile robot (car toy).

Keywords: road following, autonomous vehicle navigation, neural network classifier, Concurrent Self-Organizing Maps, Hough transform

1. Introduction

After many years of extensive research in the autonomous navigation field, building a robust driverless vehicle is still a challenge [1], mainly because of the very high variability of the environmental conditions, such as the type of the road the vehicle runs on, the surroundings or the lightening conditions.

Apart from their basic role of providing an automated steering, the road direction identification algorithms can also assist the human driver, proving useful as an active safety system for detecting potentially hazardous situations (for example, when the driver does not steer following the road due to lack of attention or lost of

---
[1] Prof., Depart. Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: victoremil@gmail.com
[2] PhD student, Depart. Electronics, Telecommunications and Information Technology, University POLITEHNICA of Bucharest, Romania, e-mail: tudorcristian80@gmail.com

consciousness). Such "lane-departure" warning systems are already integrated by some car manufacturers.

The automatic detection of the path to be followed by the vehicle is a difficult task, especially when dealing with outdoors scenes, requiring a system who is able to adapt to changing conditions. Most of the model-based road following systems are unable to cope with the change of the environmental variables like road width and lightning conditions, making them unreliable.

Since the artificial neural networks proved to be a promising solution for other pattern recognition problems (for example, in handwriting or face recognition), such conexionist techniques have been adopted for the autonomous navigation tasks, and specifically in autonomous road following.

One of the first successful implementation architecture for visual road following was ALVINN (*Autonomous Land Vehicle in a Neural Network*) developed by Pomerleau [2] at Carnegie Mellon University, Pittsburgh, USA. ALVINN is based on a feedforward network (multilayer perceptron), where the network is fed directly with image data at a low resolution level. ALVINN is a perceptron system which learns to control the NAVLAB vehicles by watching a person drive. ALVINN's architecture consists of a single hidden layer backpropagation network. The input layer of the network is a 30x32 unit two dimensional "retina" which receives input from the vehicle video camera. Each input neuron is fully connected to a layer of five hidden units which are in turn fully connected to a layer of 30 output units. The output layer is a linear representation of the direction the vehicle should travel in order to keep the vehicle on the road. ALVINN is the most successful development of the ARPA UGV (*Unmanned Ground Vehicle*) program. ALVINN has been demonstrated on several test vehicles driving at speeds of up to 70 mph, and for distances of over 90 miles without human intervention. ALVINN was originally designed as part of an unmanned vehicle for the modern battlefield, performing reconnaissance, surveillance as well as nuclear, biological, and chemical (NBC) detection missions [3]. However, it was adapted for civilian use, as part of the *Intelligent Vehicle Highway System* (IVHS) initiative.

The same team lead by Pomerleau designed an improved variant of ALVINN called MANIAC (*Multiple ALVINN Networks In Autonomous Control*) [4], which confers to the autonomous vehicle the ability to robustly and transparently navigate between many different road types. MANIAC is composed of several ALVINN networks, each trained for a single road type that is expected to be encountered during driving.

A few approaches for visual identification of road direction of an autonomous vehicle using radial basis function (RBF) neural networks have been performed and reported by Rosenblum and Davis [5] as well as by Neagoe et al [6].

The well known *road detection and tracking* algorithm (RDT), developed at the Universität der Bundeswehr München (UBM), has been adapted for following unpaved paths (dirt road detection) and contour lines [7].

Road detection is also a key issue for autonomous driving in urban traffic. He, Wang, and Zhang [8] have proposed a road-area detection algorithm based on color images.

Recently, Dahlkamp et al [9] have presented a method for identifying drivable surfaces in difficult unpaved and off-road terrain conditions as encountered in the DARPA Grand Challenge robot race. Instead of relying on a static, pre-computed road appearance model, this method adjusts its model to changing environments.

In this paper we present and evaluate an original algorithm for visual identification of road direction of an autonomous vehicle, based on a neural network classifier called *Concurrent Self-Organizing Maps* (CSOM) [10] introduced by V.E. Neagoe, representing a winner-takes-all collection of neural modules. Thus, we further extend the approaches presented in [11] and [12].

2. Algorithm Description

As shown in Fig. 1, the proposed model for visual autonomous road following consists of two main processing steps:

(a) **feature selection** using either a standard *edge detection* algorithm, or an algorithm based on the *Hough transform*, or *no feature selection* at all;

(b) **classification** using one of the four classifiers: *Concurrent Self-Organizing Maps (CSOM)*, *Multilayer Perceptron (MLP)*, *supervised Self Organizing Map (SOM)* [13], and *K-Means (Nearest Mean)* [14].

We have quantized the path to be followed in *three classes* (directions): *left*, *straight ahead* and *right*, so the output of the road following model is one of the above three directions.

2.1. Feature Selection

2.1.1. Edge Detection

The first considered feature selection algorithm is the classic Canny Edge Detector [15]. Basically, when applying this algorithm, a binary image of the same size as the input image is obtained, with the points belonging to edges marked as 1's and the others marked as 0's. This way, a large amount of useless information is filtered out, reducing the total amount of data to be further processed (by the classifier), while preserving the structural properties of the input image.
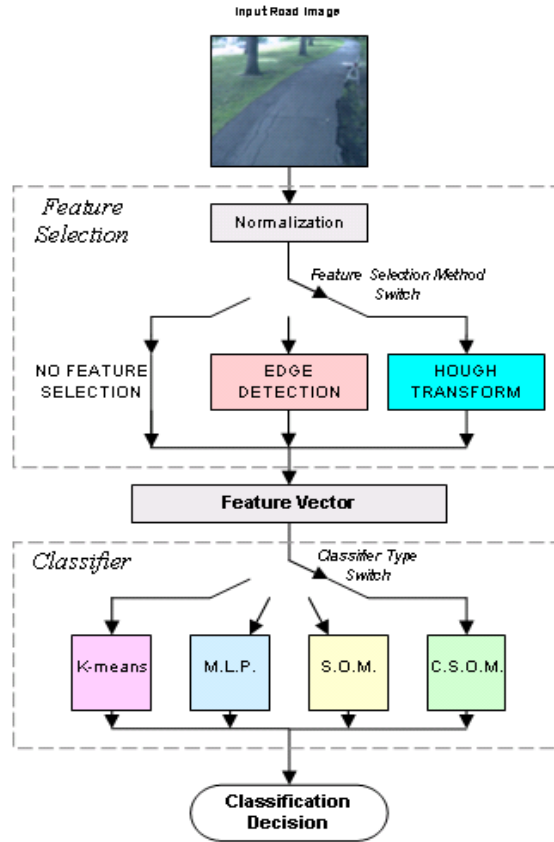
Fig. 1. Flowchart of the Road Direction Visual Identification algorithm

2.1.2. Hough Transform

The Hough transform is a computer vision technique used for identifying certain features (shapes) within a digital image. It consists of parameterizing a description of a feature at any given location in the original image's space. Because it requires that the desired features are specified in a parametric form, the Hough transform is most commonly used for the detection of regular curves such as lines, circles or ellipses. We have used its simplest form, namely the *Hough line transform*. The algorithm operates using the edge points detected with an embedded Canny Edge Detector.

In order to use the Hough transform, we must also choose a way of characterizing a line. One basic representation of a line is the slope-intercept form

$$y = mx + b, \tag{1}$$

where m is the slope of the line and b is the y-intercept (that is, the y component of the coordinate where the line intersects the Oy-axis).

This method cannot be used, however, to describe any given line, because when lines get more and more vertical, the magnitudes of *m* and *b* grow towards infinity. Another representation of a line, that solves the aforementioned problem, is the "normal form"

$$x \cos \theta + y \sin \theta = \rho. \tag{2}$$

This equation describes a set of lines passing through (x, y), where ρ is the length of a normal from the origin to this line and θ is the orientation of ρ with respect to the Ox-axis. For any point (x_i, y_i) on such a line, ρ and θ are constant.

In our image analysis application, the coordinates of the edge points (x_i, y_i) in the image are known and therefore serve as constants in the parametric line equation, while ρ and θ are the unknown variables we seek. The transform is implemented using an accumulator for the Hough parameter space (ρ, θ). This space is quantized into finite intervals (for both ρ and θ) that define the accumulator cells. By iterating through all possible angles for θ, we can compute the corresponding values for ρ (using equation (2)) and the corresponding accumulator cells are incremented. As the algorithm runs, each edge point (x_i, y_i) is transformed into a discretized (r, θ) curve (a sinusoid) in the Hough space and the accumulator cells which lie along this curve are incremented. The points which are collinear in the Cartesian image space can be viewed in the Hough parameter space as belonging to curves which intersect at a common (ρ, θ) point. Equivalently, the resulting peaks in the accumulator array represent strong evidence that a corresponding straight line exists in the image.

The Hough transform minimizes the effect of the noise present in the original image and/or in the binary image obtained after the edge detection stage.

2.2. CSOM Classifier

2.2.1. CSOM Architecture

The neural classification model called Concurrent Self-Organizing Maps (CSOM) represents a collection of small SOMs using a global competition strategy. The number of these modules equals the number of classes, so in our case there are three such identical modules, each being trained individually to provide best results for one class only, corresponding to a specific road direction.

2.2.2. Training the CSOM

As mentioned, each CSOM module is trained independently, using only the subset of training images corresponding to its assigned class (road direction), according to Fig. 2; this training is a supervised one.
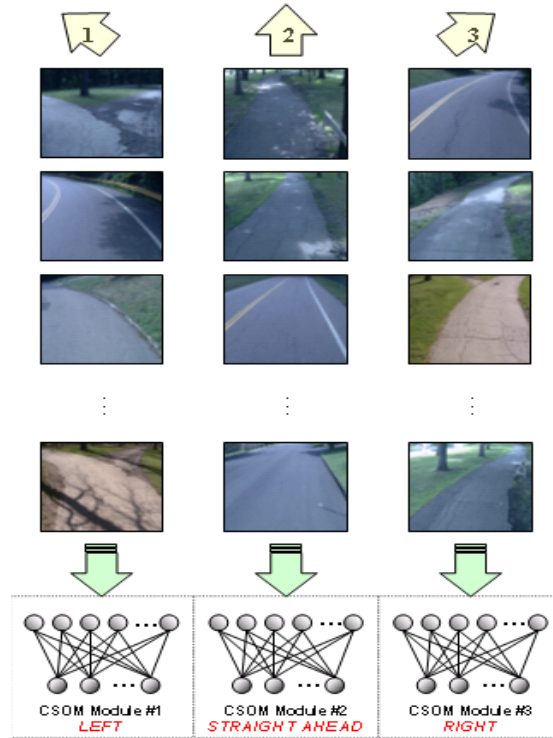
Fig. 2. The training phase of the CSOM model

2.2.3. Recognition Phase

After the training phase, the system should be able to correctly classify an unknown image into one of the three classes (road directions), using the information stored in the CSOM weights.

The image to be classified is applied to all the three modules of the CSOM system (see Fig. 3).

The distances between input vector and all the neurons of the three modules are computed and then the best matching neuron (over all) is determined as the one "closest" to the input vector. The unknown road image is classified as "belonging" to the minimum distance neuron, and the label of the module containing the best matching neuron is assigned to the input image.
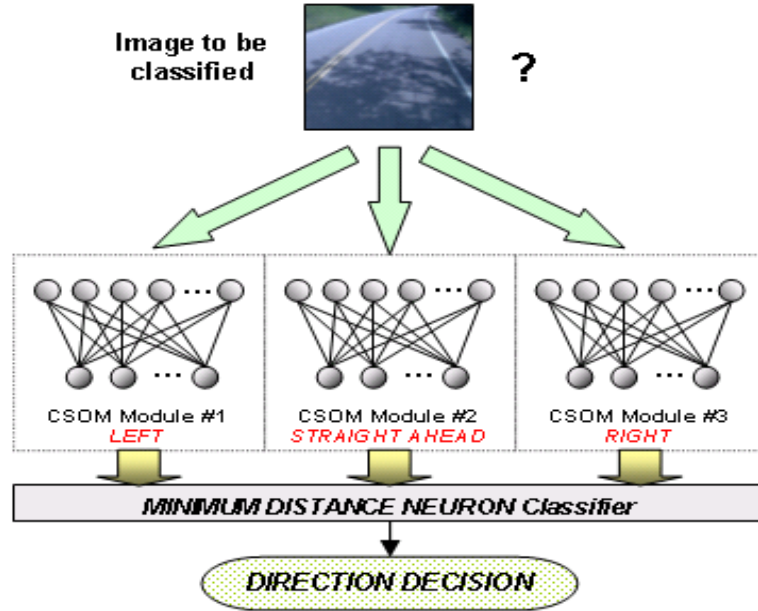
Fig. 3. The classification phase of the CSOM model

3. Experimental Results

3.1. Database

For our study, we have used road images from the *CMU (Carnegie Mellon University) Vision and Autonomous Systems Center's Image Database.* This database includes a large number of road images captured as part of the extensive research they conduct at their Robotics Institute, for the NAVLAB series of vehicles.

Aiming to ensure the maximum road-scenes diversity while keeping the number of train and test images reasonably low, we have extracted a dataset containing 210 color images of 256 x 240 pixels, equally divided into 3 road direction classes. From the 70 images available for each class, a varying number (minimum 5 and maximum 65 images) have been used for training the system, while the remaining pictures have been used for testing. Several samples from this database are presented in Fig. 2.

3.2. Performance evaluation as a function of classifier type

First, we investigate the influence of the classifier type on the correct road identification rate. The results, shown in Table 1 and Fig. 4, are obtained by

computer simulations using only the luminance (Y) component of the input pictures and considering the algorithm variant without feature selection. We have evaluated the following classifiers:

(a) the Concurrent Self-Organizing Maps (CSOM) system with SOM submodules of circular topology and a number of 10 neurons per module, 20 training epochs;

(b) the Multilayer Perceptron (MLP), with one hidden layer containing 30 neurons and with sigmoid activation functions, 50 training epochs;

(c) the Self-Organizing Map (SOM), in two variants, one with a square topology (SOM-s) having 10 x 10 neurons, and one with a circular topology (SOM-c) with 30 neurons, both trained for 20 epochs;

(d) the classical statistical classifier of the Nearest Mean (or prototype) – K-Means.

Table 1

Correct road identification rate [%] as a function of the classifier type (gray-scale images)

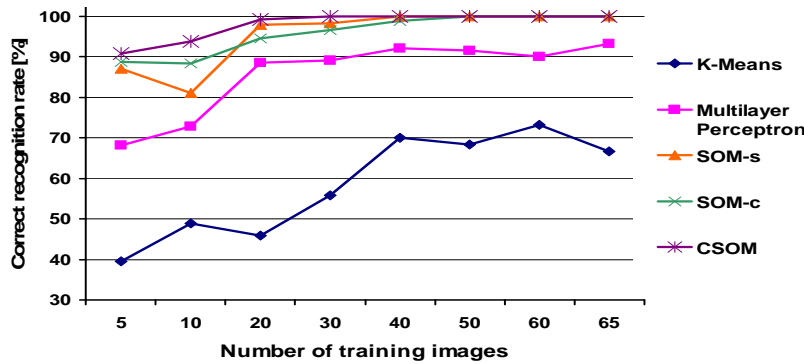| Classifier	Number of training images							
	5	10	20	30	40	50	60	65
K-Means	39.49	48.89	46	55.83	70	68.33	73.33	66.67
MLP	68.21	72.78	88.67	89.17	92.22	91.67	90	93.33
SOM-s	87.18	81.11	98	98.33	100	100	100	100
SOM-c	88.72	88.33	94.67	96.67	98.89	100	100	100
CSOM	90.77	93.89	99.33	100	100	100	100	100

Fig. 4. Road identification performances as a function of classifier type

3.3. Performance evaluation as a function of the selected color component

Table 2 and Fig. 5 present the road correct identification rate as a function of the selected color component: *red* (R), *green* (G), *blue* (B) or *luminance* (Y). The CSOM system contains SOM modules with a circular architecture, with 10 neurons per module, and no feature selection is employed. The chosen number of training epochs is 20.

CSOM correct road identification rate [%] as a function of the selected color component

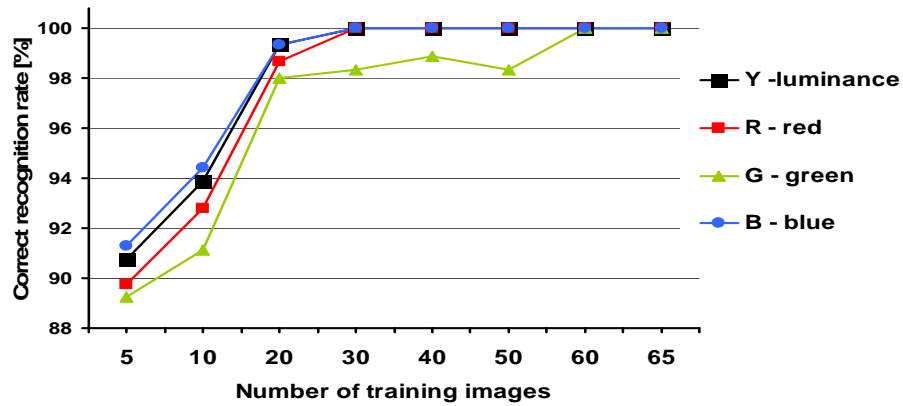| Color component	Number of training images							
	5	10	20	30	40	50	60	65
Y - luminance	90.77	93.89	99.33	100	100	100	100	100
R - red	89.75	92.78	98.67	100	100	100	100	100
G - green	89.23	91.11	98	98.33	98.89	98.33	100	100
B - blue	91.28	94.44	99.33	100	100	100	100	100

Fig. 5. CSOM performance as a function of the color component selection

3.4. Influence of the feature selection method

In Table 3 and Fig. 6 one can see the influence of the chosen feature selection method on the CSOM identification performance.

CSOM correct road identification rate [%] as a function of the feature selection method (luminance component; 20 training images)

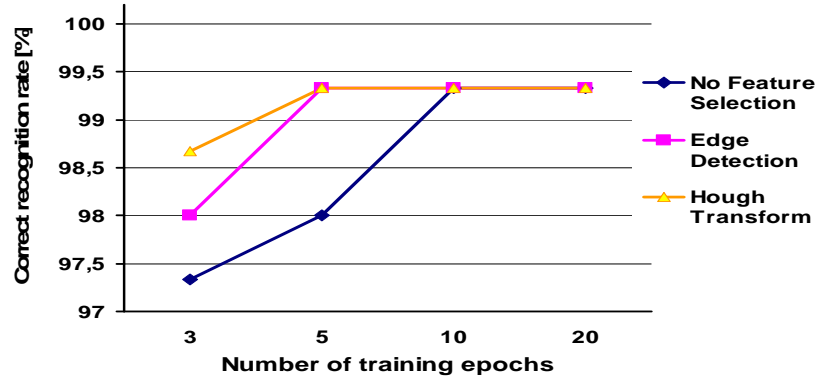| Color component	Feature Selection Method	Number of epochs			
		3	5	10	20
Y luminance	None	97.33	98	99.33	99.33
	Edge Detection	98	99.33	99.33	99.33
	Hough Transform	98.67	99.33	99.33	99.33
R red	None	96.67	97.33	98.67	98.67
	Edge Detection	97.33	98.67	98.67	99.33
	Hough Transform	98	99.33	99.33	99.33
G green	None	96	96.67	98	98
	Edge Detection	96.67	97.33	98.67	98.67
	Hough Transform	98	98.67	99.33	99.33
B blue	None	97.33	98.67	99.33	99.33
	Edge Detection	98	99.33	99.33	99.33
	Hough Transform	98.67	99.33	99.33	99.33

Fig. 6. CSOM performance as a function of the feature selection method, for 20 training images and using only the luminance (Y) component

4. CSOM for real-time path detection

We have used a small mobile robot ("police jeep" toy shown in Fig. 8) of sizes 30 cm (L) x 22 cm (W) x 19 cm (H). The machine vision system (Fig. 7) has two modules: the on-board module and the stationary module. We have used a wireless TV camera mounted on the mobile robot and the video signal is transmitted through the 1.2 GHz channel to the computer. The second communication channel (40 MHz) transmits the direction commands from computer to the robot.

In a preliminary phase, there have been performed the *acquisition and labeling* of the road picture set. A human trainer has driven the mobile robot to follow a specific path made by white sheets of paper on a red carpet (see Fig. 8 - *right*). The computer software automatically stores the corresponding data (image & its direction label) sequence, obtaining the labeled road image set.

The *training of the CSOM modules* has been performed using three classes of labeled road images: *left*, *straight ahead* and *right*. Each neural module has been trained with the image subset corresponding to its class label.

The artificial vision system identifies in real time the correct direction as follows. The video signal captured by camera is transmitted through the 1.2 GHz channel to the computer, where the input path picture is stored. The CSOM software computes the minimum distance between the input picture vector and every neuron. The minimum distance neuron is the winner and the direction label of the module to which the winner belongs is assigned to the input image. This direction is coded and transmitted through the 40 MHz channel to the steering system of the mobile robot, so that it can follow the correct path.
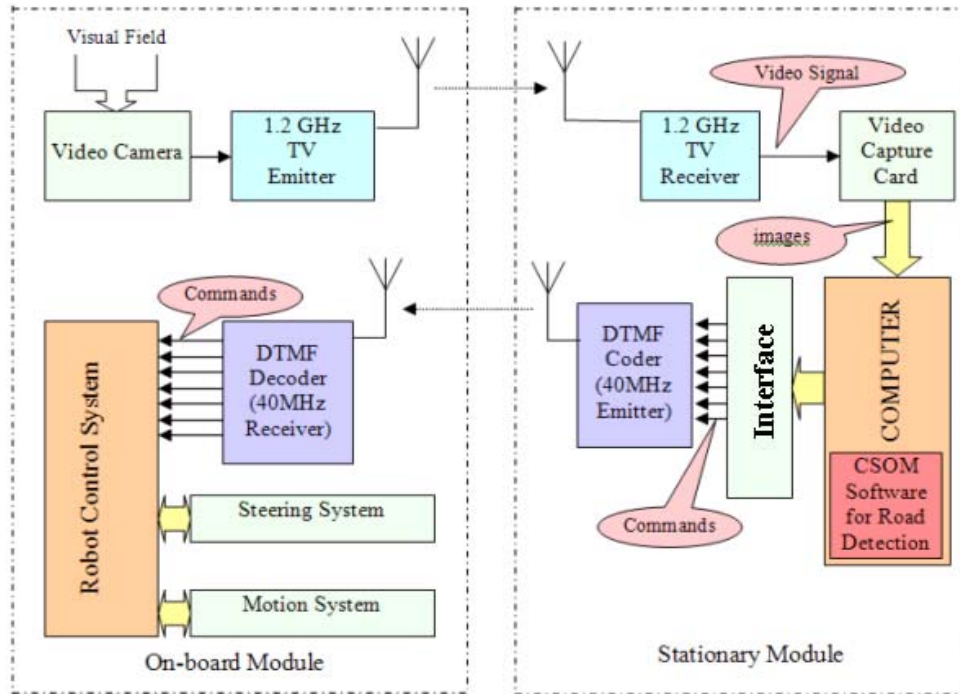
Fig. 7. General block diagram of the machine vision system for the automated guided mobile robot
(on-board module, stationary module)

Fig. 8. *Left*: Mobile robot including the on-board module; *Right*: Autonomous driving area

5. Conclusions

The experimental evaluation of the model has shown that the best road identification score is obtained by *CSOM* classifier. Moreover, CSOM architecture allows a faster and a more flexible algorithm implementation by comparison to other classifiers, a key feature for real-time applications.

For color images when choosing a small training lot, the best performances are obtained using the *blue* picture component.

For a small lot of training color images, the feature selection based on *Hough transform* leads to a better road identification score than that based on edge detection.

R E F E R E N C E S

[1] *P. Afonso, J. Azevedo, C. Cardeira, B. Cunha, P. Lima, V. Santos*, Challenges and solutions in an autonomous driving mobile robot competition, Proc. 7th Portuguese Conference in Automatic Control, Lisbon, Portugal, 11-13 September 2006

[2] *D.A. Pomerleau*, Efficient training of artificial neural networks for autonomous navigation, Neural Computation, **vol. 3**, no. 1, 1991, pp. 88-97

[3] *D.A. Pomerleau*, Defence and civilian applications of the ALVINN robot driving system, Proc. 1994 Government Microcircuit Applications Conference, San Diego, CA, Nov. 1994, pp. 358-362

[4] *T. Jochem, D.A. Pomerleau, C. Thorpe*, MANIAC: A next generation neurally based autonomous road follower, Proc. Internat. Conf. on Intelligent Autonomous Systems, February, 1993

[5] *M. Rosenblum, L.S. Davis*, An improved radial basis function network for visual autonomous road following, IEEE Trans. Neural Networks, **vol. 5**, 1996, pp. 1111-1120

[6] *V.E. Neagoe, M. Valcu, B. Sabac*, A neural approach for detection of road direction in autonomous navigation, in Computational Intelligence (Theory and Applications, ed. B. Reusch), Proc. Internat. Conf., 6th Fuzzy Days Dortmund, Germany, May 1999, Springer, Berlin, pp. 324-341

[7] *M. Lützeler, S. Baten*, Road recognition for a tracked vehicle, Proc. SPIE Conf. on Enhanced and Synthetic Vision, AeroSense 2000, Orlando, FL, USA, April 24-28, 2000

[8] *Y. He, H. Wang, B. Zhang*, Color-based road detection in urban traffic scenes, IEEE Transactions on Intelligent Transportation Systems, **vol. 5**, no. 4, Dec. 2004, pp. 309-318

[9] *H. Dahlkamp, A. Kaehler, D. Stavens, S. Thrun, G. Bradski*, Self-supervised monocular road detection, Proc. Robotics: Science and Systems, Aug. 16-19, 2006, Philadelphia, USA

[10] *V.E. Neagoe, A. Ropot*, Concurrent self-organizing maps for pattern classification, Proc. First IEEE International Conference on Cognitive Informatics, ICCI 2002, Aug. 19-20, 2002, Calgary, Canada, pp. 304-312

[11] *V.E. Neagoe, C.T. Tudoran*, Road following for autonomous vehicle navigation using a concurrent neural classifier, Proc. Int. Symp. Soft Computing in Industry (ISSCI'08), World Automation Congress, 2008 (WAC 2008), Sept. 28–Oct. 2, 2008, Hawaii, USA, ISBN: 978-1-889335-38-4

[12] *C.T. Tudoran, V.E. Neagoe*, A New Neural Network Approach for Visual Autonomous Road Following, Proceedings of the WSEAES 14th International Conference on Computers, July 23-25, 2010, Corfu Island, Greece, pp. 266-271

[13] *T. Kohonen*, Self-Organizing Maps, Springer-Verlag, 1995

[14] *M. Bishop*, Pattern Recognition and Machine Learning, Springer, New York, 2006

[15] *R.C. Gonzalez, R.E. Woods*, Digital Image Processing, Third Edition, Prentice-Hall, Upper Saddle River, New Jersey, 2008.