Proceedings of International Joint Conference on Neural Networks, San Jose, California, USA, July 31 – August 5, 2011

A Self-Organizing Neural Scheme for Road Detection in Varied Environments

U. A. Malik, S. U. Ahmed and F. Kunwar

Abstract—**Detection of a drivable space is a key step in the autonomous control of a vehicle. In this paper we propose an adaptive vision based algorithm for road detection in diverse outdoor conditions. Our novel approach employs feature based classification and uses the Kohonen Self-Organizing Map (SOM) for the purpose of road detection. The robustness of the algorithm lies in the unique ability of SOM to organize information while learning diverse inputs. Features used for the training and testing of SOM are identified. The proposed method is capable of working with structured as well as unstructured roads and noisy environments that may be encountered by an intelligent vehicle. The proposed technique is extensively compared with the k-Nearest Neighbor (KNN) algorithm. Results show that SOM outperforms KNN in classification consistency and is independent to the lighting conditions while taking comparable classification time which shows that the network can also be used as an online learning architecture.**

***Keywords: Autonomous vehicle control, road detection, neural networks, unsupervised learning, self-organizing maps, feature based classification*.**

I. Introduction

With the application of intelligent agents for the purpose of transportation the fields of Mobile Robotics and Autonomous Vehicle Control (AVC) have been subject to extensive research. One of the principal components of any intelligent agent is the accurate perception of its immediate environment. For a Vision-Based Vehicle Control system this perception comes from the detection of drivable space and its edges. The overall *performance measure* of any such system is thus in direct correlation with the time efficiency and adaptability of the road detection algorithms.

Road detection approaches developed in the past decade can be broadly classified as either using classical computer vision functions or machine learning paradigms. Common texture based road segmentation techniques are presented in [1]-[4]. Whereas improved approaches for road detection in structured and unstructured environments are given by Thorpe et al [5], [6].

The above mentioned schemes are limited by the fact that they require extensive modeling of specific parameters. Parametric modeling cannot be effective in most scenarios since there will always be a large number of unrepresented variables in diverse environments which makes parameter inference a difficult, and in some cases an impossible task.

Machine learning techniques improve on the limitations of conventional computer vision techniques by intelligently modeling the visual information. Some widely utilized machine learning paradigms involve the use of evolutionary algorithms [7]-[10]. The Support Vector Machine approach has also been applied for road, non-road classification [11]. Machine learning algorithms mainly suffer because of the time inefficiency of their computations which renders them unfeasible for use in autonomous driving systems in most realistic environments where quick steering responses are required.

Artificial Neural Networks have been applied for the past two decades for fast classification of road surfaces in various environments [12]-[14]. Extensive research has been carried out in Carnegie Mellon University on autonomous road followers [15-16] using the Multilayer Perception (MLP) neural network. The main shortcoming of the neural network techniques is that they use the supervised learning paradigm. Training of such networks is a key concern because they require both the input and the desired output for each training instance and these input/target data pairs cannot easily be obtained in all situations. Recently a more efficient vision based driving system was developed [17] by using cellular neural networks.

It is evident from the above discussion that competitive learning approaches which work on the principle of unsupervised learning can be efficiently employed for road extraction. Moreover, these systems can be made to act as *lifelong* learners which can intelligently estimate environment parameters and adapt to change while they are in use.

Kohonen Self Organizing Map [18] is an unsupervised learning neural architecture that is a digital analog of the brain's self-organizing capability. Research has shown that topologically adjacent areas in the cerebral cortex perform similar cognitive functions. In such functions the relationship of physically similar stimulus is closely maintained. Likewise the nodes in Kohonen Network that represent input vectors close to each other in the input space have their weight vectors so oriented that they lie close by in the output space. In this way the weight vectors of the network will cluster the space such that they approximate the probability density function of the input vectors [18]. This ability of the SOM has been leveraged in this paper for the detection of a drivable surface. There is no such scheme in the authors' knowledge where SOM was adapted to the

Manuscript received February 18, 2011. This work was supported in part by the Department of Mechatronics Engineering, CEME, National University of Sciences and Technology (NUST), Islamabad, Pakistan. All authors are associated with this university.

U. Malik (phone: 0322-4614541; email: usman.malik88@gmail.com).
U. Ahmed (e-mail: s.usman87@yahoo.com).
K. Faraz (e-mail: k.faraz@ceme.nust.edu.pk).

problem of drivable space detection.

The rest of the paper is organized as follows: Section 2 provide an overview of our proposed algorithm. In Section 3 we discuss the training methodology of the SOM. The experimental observations and results are provided in Section 4. Finally, conclusions and future work is discussed in Section 5.

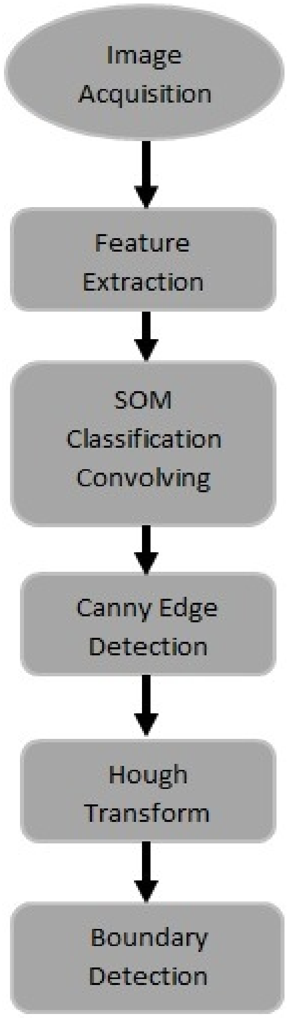

Fig. 1. Proposed Algorithm

Our proposed algorithm consists of five steps which include convolving of the entire cropped image and feature estimation, classification through SOM, employment of Edge Detection on the image, implementation of Hough Transform, and finally the evaluation of Boundaries as shown in Fig. 1. Each of the above mentioned steps are discussed in detail in the subsequent sections.

A. Feature Extraction

Feature extraction is the only preprocessing step in the presented scheme. For the purpose of classification and training the network, the acquired image is initially cropped at a suitable location. Subsequently, a 15x15 neighborhood of pixels is used to extract features and form a feature vector that forms the input to the SOM. Feature extraction consists of the following steps.

1) Mean RGB values: First the mean blue, green and red values are calculated in the neighborhood to be used as the first three features.

$$\overline{q}_i = \frac{1}{m \times n} \sum_{x=0}^{m} \sum_{y=0}^{n} s_i(x, y) \tag{1}$$

where q_i is the mean value. For our case $m = 15$, $n = 15$, i = red, green, blue, s_i is the pixel intensity value, x and y are the row and column number respectively. These features represent the light intensity of the environment.

2) Color Variation: The color variation is simply the standard deviation of the RGB values in the neighborhood.

$$v_i = \sqrt{\frac{1}{m \times n} \sum_{x=0}^{m} \sum_{y=0}^{n} [s_i(x, y) - \overline{q}_i]^2} \tag{2}$$

where v_i is the standard deviation and other variables are as given before. These features numerically represent the texture of the sub-image in the kernel window.

3) Color Saturation: Color saturation is calculated by dividing the sum of values of a particular color by the sum of all the color values of the pixels in the sub-image.

$$b_i = \frac{1}{T} \sum_{x=0}^{m} \sum_{y=0}^{n} s_i(x, y) \tag{3}$$

where b_i is the saturation of a specific color, T is the sum of all values in the neighborhood and other variables are as discussed before. The saturation level makes a distinction between road and non-road areas on the basis of color domination. In a sub-image with off-road grass, for example, the green color levels will dominate.

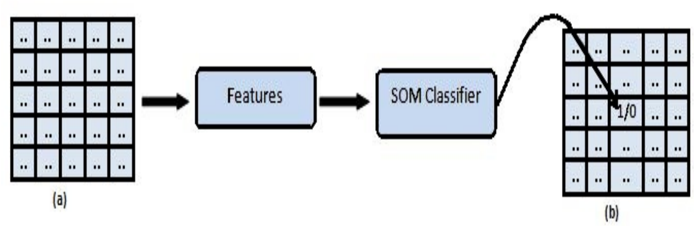

Fig. 2. Classification Process using a 15x15 Neighborhood

B. SOM Classification

Classification through SOM plays a pivotal part in our algorithm and is used for the sole purpose of estimating the drivable space in the image i.e. differentiating road area from non-road area. The feature vector formed in the previous step is given as input to the SOM classifier. The SOM then designates the center pixel of the original sub-image to be either road or non-road on the basis of the winner neuron for that sub-image. The winner neuron is selected on the basis of minimum Euclidean distance from the presented features. The feature vector is first normalized to ensure that no feature dominates the classification process. The process is repeated for the entire image in which each pixel is either given a value of zero or one to form a binary image. The process is shown in Fig. 2.

C. Edge Detection

Canny Edge enhancement is done on the binary image obtained in the previous step. Canny is a multistage edge detection algorithm which calculates the intensity gradient of the image. It takes a grayscale image as an input and returns a binary image with intensity at the edges. The binary images formed in the previous step do not vary much in their lightness continuity i.e. the noise in them is minimal so a static Canny parameter can be readily found.

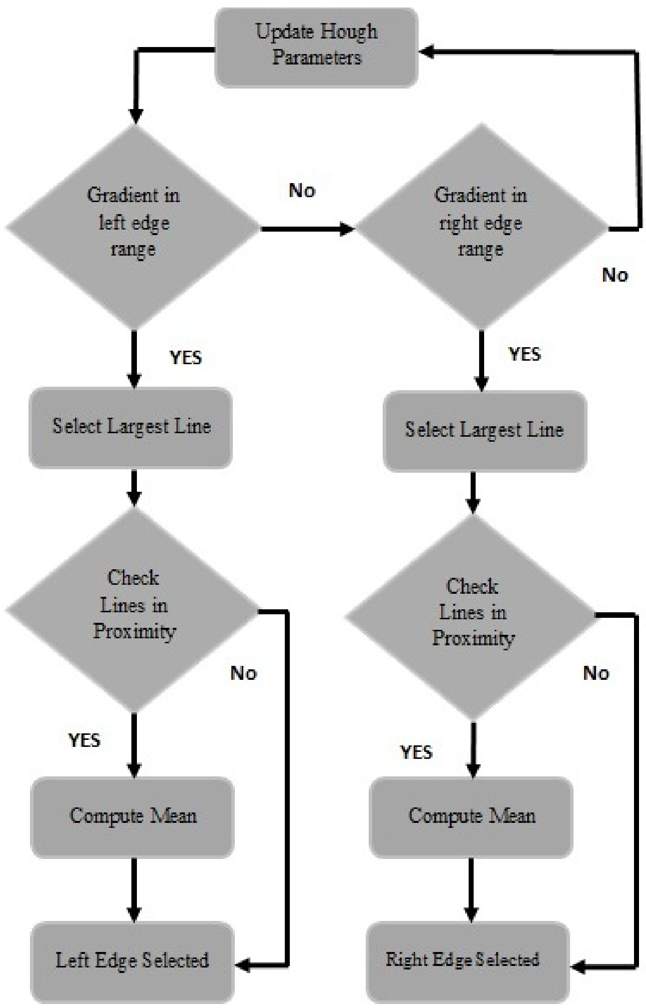

Fig. 3. Boundary Optimization

D. Hough Transform

Hough transform is the employed methodology for the detection of curves of any type including lines in an image. In our algorithm we have used the Probabilistic Hough Transform. It takes the edge image from the previous step and returns the starting and ending points of the lines. The Probabilistic Hough Transform can detect lines even in images with high noise content and also in case of outdoor images that contain few long linear segments and lack continuous lines. The Probabilistic Hough Transform will find a large number of lines in most cases. Therefore, we use an edge optimization technique to reduce the number of edges to the most likely candidates.

E. Evaluation of Boundaries

After the previous steps the lines obtained may not necessarily be our desired edges. Drivable space edges are found by first rejecting the lines not in a particular gradient range. The lines are divided into left and right bins on the basis of gradient. One longest line segment from each bin is chosen. Now the starting and ending points of the lines that are in close proximity and have small difference in orientation are averaged to obtain the edges. The process is shown in Fig. 3.

II. TRAINING METHODOLOGY

The proposed scheme uses SOM to estimate the drivable space in a given input image. For this estimation first the SOM network has to be trained to adjust its weights, which are distributed uniformly in the beginning, to the given dataset images. For this purpose a dataset consisting of diverse road environments was developed. The dataset contains images of both structured and unstructured roads. Also the dataset contains images in different lighting conditions such as roads with bright sunlight, roads with shadows and road images in poor light i.e. in the evening. After the dataset accumulation was completed, training of the network was carried out in the sequential mode using the 'Winner Takes Most' (WTM) approach. The WTM algorithm employed in this paper is described subsequently.

1) *Weight Initialization:* For the WTM training first the weight vectors of all neurons are initialized uniformly.

2) *Selection of Winner Neuron:* The input vectors are presented sequentially. For a given input vector the winner neuron is chosen on the basis of similarity between it and the neuron weights. We used Euclidean distance measure for winner neuron selection.

$$L_m = \sqrt{\sum_{i=0}^{N} [x_i - w(i, j)]^2} \tag{4}$$

Where L_m is the Euclidean distance, $w(i, j)$ is the weight associated with the input $x(i)$ and neuron j.

3) *Neuron Weights Modification*: The weights of all the neurons in the neighborhood are updated.

$$w_{n+1}(i, j) = w_n(i, j) + \eta(n)e(j) \tag{5}$$

$$w_{n+1}(k, j) = w_n(k, j) + \eta(n)h(d, \sigma)e(j) \tag{6}$$

$$\eta(n) = \eta_0 e^{-(n/\tau_2)} \tag{7}$$

$$e(j) = x(j) - w_n(i, j) \tag{8}$$

Where n is the iteration number i and j are the indices of the winner neuron and feature. The learning rate $\eta\ (n)$ is chosen to be a Gaussian function decreasing with the number of iterations, η_0 is the learning rate constant and τ_1 is the learning rate time constant. The neighborhood $h(d,\sigma)$ is a function of spread σ around the winner and distance d between the winner and the neighboring neuron k, $e(j)$ is the difference between the winner neuron weight and the given input.

4) *Neighborhood Adjustment:* The learning rate, neighborhood function, and the spread values are updated.

$$h(d,\sigma) = h_0 e^{-(d^2/\sigma^2)} \quad (9)$$

Here h_0 is the neighborhood constant and σ is given by

$$\sigma = \sigma_0 e^{-(n/\tau_2)} \quad (10)$$

Where σ_0 is the spread constant and τ_2 is the spread time constant.

5) Iterations: The above mentioned steps are repeated until the preset condition of minimum error or maximum number of iterations is reached.

The training is done in two phases. In the ordering phase the spread is kept almost as large as the size of layer for quick ordering. In the convergence phase the spread and the neighborhood constants are kept small. The learning rate is kept constant for the convergence phase.

III. EXPERIMENTAL RESULTS

The proposed architecture was tested on an actual video feed of the university roads. It demonstrated exceptional classification and road detection ability and computational time efficiency. The video feed was processed at 20 frames per second which is almost optimal for online use. The SOM classification showed invariability to all light compositions for example shadows or poor light conditions. A dataset consisting of almost all possible road environments for example midday and dusk etc. was developed for the training and testing purpose. The dataset comprises many images of unstructured roads and also roads without any markings. In addition the dataset includes ill-formed or coarse road textures. In all the dataset comprised of 450 images of which half of the images were used for training purpose. All images were 1200x1600 pixels.

For training, initially complete pictures were given to be convolved and for features to be extracted from the 15x15 sub-images. It was found after extensive experimentation with network parameters that the network exhibited sensitivity for only the holistic features of the images such as grass and vegetation, dirt, sky and road textures while it demonstrated a consistent response to variability in light composition. This type of variation in response with light conditions is what most road detection algorithms suffer from. Then in order to employ data reduction and effective training of network a new dataset consisting only of 15x15 sub-images was developed which contained sets of holistic features described above. Eight such sets of different categories in which an intelligent vehicle may find itself, were developed from the original image dataset. Each set contained 100 such sub-images of which 40 were used for training. The learning process of adjusting the weights (taken from a random uniform distribution) to the dataset images, took approximately 4 minutes. This means that the network can be initially trained with a road dataset offline once and then implemented as an online learner which learns from a frame in the video feed once every second to ensure consistency in classification accuracy. The algorithm was tested on a PC running a dual core with 2GB memory. Results of drivable space detection in various environments are shown in Fig. 4 through 6. It is to be noted that the drivable space is shown in dark while the non-road portions are shown in light shades for illustration.

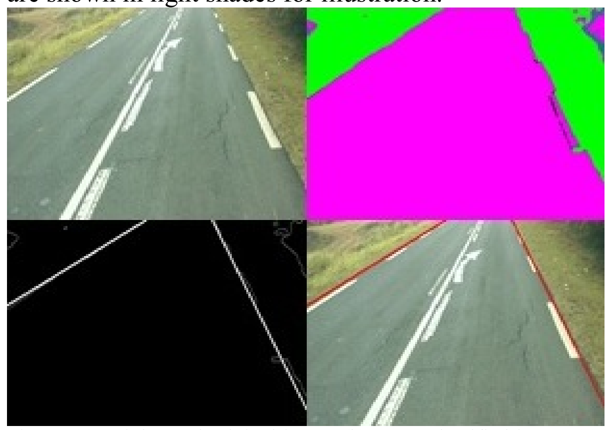

Fig. 4. (a) Bright road original image (b) SOM classified image (c) Image with detected edges (d) Optimized edge detected image.

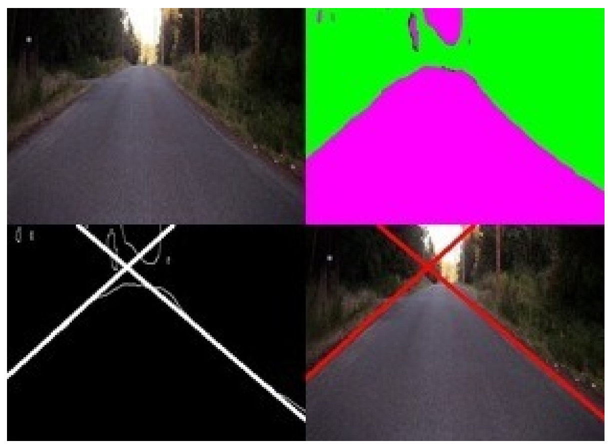

Fig. 5. Results with dark road

Fig. 6. Results with shadow road

One of the critical tasks in the SOM algorithm is the choice of layer topology and the number of neurons in the topology itself. To estimate these characteristics the network was trained with many different configuration of neurons. The most stable results were obtained by using a twodimensional topology with a 3x3 grid of neurons. The network accuracy was established by testing it with the remaining sub-images from the sets. It was also determined during the training that each of the neuron corresponded to one of three road classes which include the bright road pattern, road pattern with shadows and the dark or poor light road pattern. Likewise other neurons exhibited recognition for the off-road classes. The final network was tested with 480 sub-images with 160 sub-images for each road class. Table 1 shows the confusion matrix of the classification of the road pattern classes. The results show that the SOM only had some difficulty with classifying the dark road patterns where the environment has a low overall contrast.

TABLE 1 CLASSIFICATION ACCURACY OF PROPOSED SCHEME FOR ROAD PATTERNS

Group	Bright Road	Dark Road	Shadow Road	Accuracy
Bright Road	158	2	0	98.75%
Dark Road	4	155	1	96.87%
Shadow Road	0	1	159	99.37%

For the analysis of the SOM network the total classification error against the number of iterations is given in Fig. 7. Best results were obtained for a learning rate of 0.1 for the ordering phase and a learning rate of 0.001 for convergence phase. In all 6100 iterations were performed for each pattern in the training set.

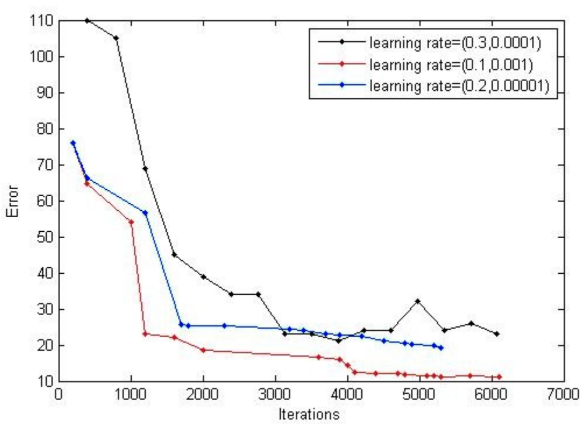

Fig. 7. Error vs. Iterations for different learning rates

To evaluate the efficiency of the scheme, it was compared with the standard K-Nearest Neighbors (KNN) [19] algorithm. The KNN algorithm first maps the feature vector into an N-dimensional space where N is the total number of features. Then each of its k Nearest Neighbors casts their vote in favor of their own class. So the class of a test instance is decided by a voting process which is weighted upon its distance from the test instance. The KNN algorithm was trained with the same sub-image dataset. Then the algorithm was tested on 480 instances randomly with 160 instances of each of the three road classes. The results are given in Table 2. The results indicate that while the KNN algorithm performed comparably in the case of bright road patterns, its accuracy significantly suffered in the case of shadow road and dark road patterns. This experiment clearly demonstrates the light invariability of the presented algorithm because its accuracy remains consistent in changing environmental light conditions.

TABLE 2 CLASSIFICATION ACCURACY OF KNN FOR ROAD PATTERNS

Group	Bright Road	Dark Road	Shadow Road	Accuracy
Bright Road	132	22	6	82.5%
Dark Road	27	115	18	71.87%
Shadow Road	14	36	110	68.75%

In addition to the accuracy, other aspects of the two techniques such as initial training time, learning update time, and pattern classification time are also compared. Results are given in table 3. It can be seen that while KNN has a better initial training and learning update times than the SOM classifier, the SOM outperforms the KNN in classification accuracy and consistency. Moreover, the pattern classification times of both the algorithms are comparable.

Classification and boundary estimation results for structured environments are given in Fig. 8. While the same results for unstructured environments are given in Fig. 9. It can be clearly seen that the algorithm works equally well in both cases. The proposed technique correctly classifies the cars and other urban obstacles as non-drivable space by the help of which the boundaries of drivable space can be evaluated.

TABLE 3 COMPARISON OF KNN AND THE PROPOSED METHOD

Attribute	KNN	SOM Classifier
Initial Training Time	0.71 s	240 s
Learning Update Time	0.0018 s	0.5 s
Pattern Classification Time	0.0012 s	0.0010 s
Classification Accuracy	74.3%	98.3%

IV. CONCLUSION AND FUTURE WORK

A robust approach is presented for clustering the road space according to the holistic features of the environment. The SOM exhibited remarkable classification accuracy. Our network took almost 0.5 seconds for learning a new pattern which suggests that the algorithm can be used as an online learner and its classification efficiency enhanced even further. The proposed approach was compared with KNN extensively. The results establish that the SOM achieved high invariability to environmental light conditions. Moreover, the sensitivity to textures and contours suggests that the network can be used as an adaptive segmentation technique as well. We are in the process of testing the algorithm to drive our university's Intelligent DRIving System (IDRIS) on both structured and unstructured roads.

Fig. 8. SOM Classified and Edge optimized images for Urban Environments

Fig. 9. Detection of edges in unstructured environments.

REFERENCES

[1] J. Zhang and H. Nagel,"Texture-based segmentation of road images," Proc. Intelligent Vehicles Symposium, Oct. 1994, pp. 260-265.

[2] J. M. Alvarez, A.M. Lopez and R. Baldrich, "Shadow resistant road Segmentation from a mobile monocular system," Pattern Recognitionand Image Analysis, Springer Berlin! Heidelberg, Volume 4478, 2007, pp. 9-16.

[3] Xiaowen Wu , YuxinPeng , Donghua Ding , Huaping Liu , Kezhong He and Fuchun Sun," Color Vision-based Multi-Level Analysis and Fusion for Road Area Detection,"Intelligent Vehicles Symposium,2008,pp. 602-608.

[4] Rasheed U. , Ahmed M. , Ali S. , Afridi J. and Kunwar F," Generic vision based algorithm for driving space detection in diverse indoor and outdoor environments," International Conference on Mechatronics and Automation (ICMA), 2010,pp. 1609-1615.

[5] J. D. Crisman and C. E. Thorpe, "SCARF: A color vision system that tracks roads and intersections," IEEE Trans. Robot. Automation, Feb. 1993, vol. 9, pp. 49-58.

[6] J. Crisman and C. E. Thorpe, "UNSCARF: A color vision system for the detection of unstructured roads," Proc. IEEE Int. Conf. Robotics and Automation, Apr. 1991, pp. 2496-2501.

[7] R. Arnay, L. Acosta, M. Sigut and J. Toledo, "Ant colony optimization algorithm for detection and tracking of non-structured roads,"Electronics Letters ,2008, Vol. 44 No. 12, pp. 725 – 727.

[8] M. Bertozzi and A. Broggi, " An Evolutionary Approach to Lane Markings Detection in Road Environments," Atti del 6 convegnodell'AssociazioneItaliana per I'IntelligenzaArtificiale, Siena, Italy, Sep.2002, pp.627-736.

[9] Byoung-Ki Jeon and Jeong-Hun Jang , "Road detection in SAR images using genetic algorithm with region growing concept," International Conference on Image Processing Vol. 2pp. 688-674 Sept. 2000.

[10] Tie Liu Nanning Zheng, " A novel approach of road recognition based on deformable template and genetic algorithm," Proceedings of Intelligent Transportation Systems, 2003, Vol 2 pp1251-1257.

[11] Shengyan Zhou, Jianwei Gong, GuangmingXiong, Huiyan Chen and Karl Iagnemma,"Road Detection Using Support Vector Machine based on Online Learning and Evaluation," IEEE Intelligent Vehicles Symposium San Diego, CA, USA,2010, pp. 256-262.

[12] Mike Foedisch and Aya Takeuchi Adaptive, "Road Detection through Continuous Environment Learning,"33rd Applied Imagery Pattern Recognition Workshop, 2004, pp. 16–21.

[13] Yang Xin and Huang Shunji, "The Global Road Extraction Approach from Synthetic Aperture Radar Images," IEEE Circuits and Systems International Conference on Testing and Diagnosis, 2009, pp. 1– 3.

[14] Xiping Tao, MuheGuo, Bo Zhang, "A Neural Network Approach to the Elimination of Road Shadow for Outdoor Mobile Robot," IEEE International Conference on Intelligent Processing Systems, 1997, vol.2, pp. 1302-1306

[15] Dean A. Pomerleau, "Efficient training of artificial neural networks for autonomous navigation," Neural Computation Journal, Volume 3 Issue1, Spring 1991 MIT Press Cambridge, MA, USA

[16] T. M. Jochem, D. A. Pomerleau, C. E. Thorpe. MANIAC: A Next Generation Neurally Based Autonomous Road Follower, IAS-3, Int. Conference on Intelligent autonomous Systems, Pittsburgh/USA, February 1993, pp. 15-18.

[17] GuobaoXu, Yixin Yin and Lu Yin, "Visual Information Processing Using Cellular Neural Networks for Mobile Robot," IEEE International Conference on Grey Systems and Intelligent Services 2007.

[18] T. Kohonen, "The self-organizing map, "Proceedings of the IEEE , 1990, vol. 78 , Issue:9, pp. 1464.

[19] Cover, T.; Hart, P.; , "Nearest neighbor pattern classification," Information Theory, IEEE Transactions on , vol.13, no.1, pp. 21- 27, Jan 1967