11th Symposium on Neural Network Applications in Electrical Engineering, NEUREL-2012
Faculty of Electrical Engineering, University of Belgrade, Serbia, September 20-22, 2012

Recognition and Classification of Geometric Shapes using Neural Networks

Sofija S. Spasojević, Marko Z. Šušić, and Željko M. Đurović

Abstract—**The research presented in this paper refers to classification of geometric shapes (cubes, pyramids and cylinders) using multilayer neural network. The input data of the algorithm are the images of shapes placed in different positions and distances from the camera. The classification is based on feature vectors that are obtained using methods of digital image processing. Feature vectors are inputs of neural network. Supervised training of neural network is performed. Reduction algorithm was used in aim of dimension reduction of feature vectors, so the classification results can be displayed graphically. Recognition and classification of geometric shapes may be of interest for realization of many robotic tasks, especially those related to catching of objects with robotic arm or movement of a robot with a set of obstacles.**

Index Terms—**feature vectors, classification, pattern recognition, dimension reduction, neural networks.**

I. Introduction

In recent years, neural networks became very popular technique for solving practical problems with successful applications in many fields. The majority of these applications are concerned with problems in pattern recognition and make use of different network architectures such as the multilayer perceptrons and the radial basis function networks [1].

Pattern recognition, classifier design and decision-making systems are very represented in many areas of science. These techniques are primarily intended for automatic classification of concrete objects. The range of objects and signals that can be recognized and classified by machine is very large.

Some examples of pattern recognition are the detection and face recognition, recognition of voice commands and other audio signals, fingerprints recognition etc. These examples are used in surveillance systems to enhance security. Recognition of traffic signs, license plates of cars and detecting edges of roads aims to improve the quality and safety of traffic. The signal recognition is also present in industrial processes, where it is necessary to make a decision on whether a system works properly on the basis of output signals and parameters that define the work mode. Beside mentioned examples, other objects that can be recognized are geometric shapes, letters, numbers, characters and handwriting [2]-[5].

The processes of recognition and classification can be successfully implemented and applied to various purposes.

Each pattern recognition process begins with an adequate data acquisition of shapes, which is followed by the algorithmic processing. The goal is to generate feature vectors based on object characteristics, so the shapes can be classified [6].

The task of this paper is recognition and classification of geometric shapes (cubes, pyramids and cylinders) based on the pictures that were generated using the camera. Shape recognition from the picture belongs to specific class of pattern recognition problems. Besides the use of mathematical tools, these problems require the use of digital image processing. Efficient digital image processing is a key factor in generating characteristics of geometric shapes and leads to successful classification results.

There are a lot of techniques that solve shape classification problems using different approaches. One of them is design of linear classifiers presented in [7]. In this paper, classification of shapes is done using multilayer neural network.

Results of this paper can be used for the realization of many robotic tasks. Robots are machines that needs to learn how to behave, move and perform tasks in their environment. Most of robotic tasks have some kind of implementation of pattern recognition or decision-making system, especialy those related to catching of objects with robotic arm and movement of a robot through trajectory with set obstacles. In robotic tasks implementation of pattern recognition system leads to better interaction and communication of robot with its environment and more efficient realization of set tasks.

II. Description of Geometric Shapes

Each shape recognition process begins with an adequate description of the given shapes and that is the most important part of classification process.

Input data of classification algorithm are images of the geometric shapes – cube, pyramid with a square base and cylinder. The shapes on the images stand upright, in several positions and at different distances from the camera. A homogeneous background was used. It was taken 30 snapshots

S. S. Author is with the Institute Mihailo Pupin, University of Belgrade, Serbia (e-mail: sofija.spasojevic@pupin.rs).
M. Š. Author is with the Institute Mihailo Pupin, University of Belgrade, Serbia (e-mail: marko.susic@pupin.rs).
Ž. Đ. Author is with the University of Belgrade, Serbia (e-mail: zdjurovic@etf.rs).

978-1-4673-1572-2/12/$31.00 ©2012 IEEE

of each object for neural network training set and 30 snapshots of each object for neural network testing set.

Examples of these snapshots are shown in Fig. 1.

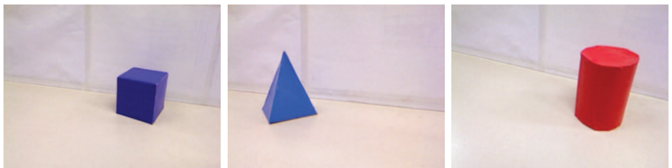

Fig. 1. Examples of cube, pyramid and cylinder snapshots.

Proper selection of features plays a key role in every classification process. The features of the given shapes need to be informative and representative, and at the same time sufficiently reliable and discriminatory for classification purposes. Appropriate image processing is central to classification based on snapshots, given that image processing is extremely important for subsequent feature extraction.

The first step in image processing is image segmentation. The idea is to extract the shape from the background and to analyze the shape as an individual region of the image. The result of such segmentation is a binary image, where white pixels represent the shape and black pixels the background. Images are usually segmented by analyzing the image pixel brightness histogram. This technique is very successful if the background brightness statistics are different from pixel brightness statistics of the object which is being extracted [8].

Following image segmentation, the edges are detected applying Canny's edge detector [9], to obtain the contour of the region. Fig. 2 - 4 show the outcomes of segmentation and edge detection of cube, pyramid and cylinder obstacles. Further processing is based on an analysis of the contour.

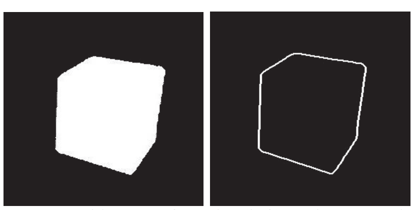

Fig. 2. Segmented image of a cube (left) and extracted edge (right).

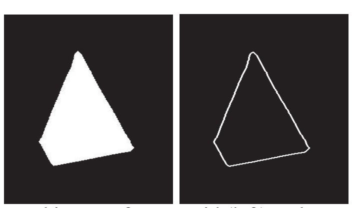

Fig. 3. Segmented image of a pyramid (left) and extracted edge (right).

Fig. 4. Segmented image of a cylinder (left) and extracted edge (right).

The first feature which describes the shapes is the ratio of eigenvalues computed for a covariance matrix derived on the basis of a pattern vector, denoted by X_1. The pattern vector is defined by the white pixel coordinates on the segmented image. The dimension of each such vector is N x 2 , where N is the number of pixels which represent the shape and this number varies depending on the size of the shape on the image. The individual pixels are defined by their x and y coordinates, such that the second dimension of the pattern vector is set at 2. In view of the above, the dimensions of the covariance matrix are 2 x 2. The end result is comprised of two eigenvectors and two eigenvalues corresponding to them. The specific value of the feature is the ratio of the first to the second eigenvalue.

After the center, height and width of the contour are defined, the second feature – the shape height to width ratio – is derived and will hereafter be denoted by X_2, Table I. The results obtained for X_2 are consistent with real ratios, in view of the fact that a pyramid and cylinder are more elongated than a cube and, as a result, their height-to-width ratio is greater.

TABLE I
FEATURE VALUE INTERVALS BY SHAPE

	Cube [min, max]	Pyramid [min, max]	Cylinder [min, max]
X_1	[1.02,1.50]	[1.17,2.50]	[1.66,2.84]
X_2	[0.89,1.22]	[1.28,2.07]	[1.38,1.75]
X_3	[7.48,33.28]	[35.09,49.27]	[13.59,33.53]
X_4	[0.66,0.91]	[0.50,0.64]	[0.65,0.89]
X_5	[0.65,0.91]	[0.57,0.78]	[0.68,1.00]
X_6	[1.31,1.59]	[1.75,3.22]	[1.40,1.89]
X_7	[1.07,1.37]	[1.35,1.54]	[1.10,1.28]

The next three features are related to the description of the smallest rectangular box which encloses the shape, the so-called bounding box. These features represent: the percentage of the bounding box not occupied by the shape, the shape to bounding box area ratio, and the shape to bounding box circumference ratio.

The first feature of the bounding box is the percentage of the bounding box not occupied by the shape, hereafter denoted by X_3. It is apparent in Fig 5 - 7 that this percentage is the greatest in the case of the pyramid, and much smaller for the cube and cylinder. The dimensions of the bounding box depend to a large extent on the position of the shape on the image.

The next two features associated with the rectangle which encloses the shape are the shape to bounding box area ratio and the shape to bounding box circumference ratio, denoted by X_4 and X_5, respectively. The area is computed as the number of pixels which comprise the shape (number of white pixels on the segmented image), while the circumference of the shape is the number of pixels which comprise the contour.

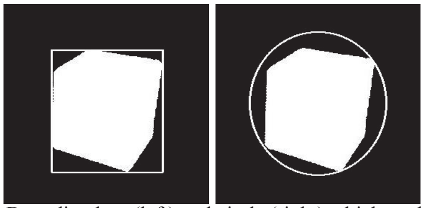

Fig. 5. Bounding box (left) and circle (right) which enclose a cube.

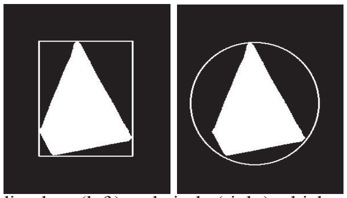

Fig. 6. Bounding box (left) and circle (right) which enclose a pyramid.

Fig. 7. Bounding box (left) and circle (right) which enclose a cylinder.

The last two features are related to a circle which encloses the shape and they are represented by the circle to shape area ratio and the circle to shape circumference ratio, denoted by X_6 and X_7, respectively. As in the case of the other features, the ratios between the circle and the shape are computed. Similar to the bounding box, the largest part of the circle which is not occupied by the shape is found in the case of the pyramid, and thus the resulting area/circumference ratios of the circle and the pyramid are the greatest, Table I. Fig. 5 -7 show the bounding boxes and circles which enclose obstacles of different shapes.

The values of all above explained features are computed for each input shape and grouped into vectors of dimensions 7x1, given that there are 7 different features. As a result, each input shape is described by a corresponding feature vector. Obtained feature vectors are later used as inputs into neural network for purpuses of training and testing neural network. However, classification results cannot be clearly presented in 7D space. According to this, reduced 2D vectors are used for graphical representation of classification results. Reduced 2D vectors are obtained using method of dimension reduction, which is explained later.

III. Feature Vector Dimension Reduction

In view of the fact that the input shapes are described by seven features, the resulting feature vectors represent the shapes in 7-dimensional space. Those vectors are used for classification using multilayer perceptron. However, it would not be possible to clearly visualize the shapes in 7D space and the results of classification.

For this reason, the dimension of the feature vectors are reduced to 2D.

The task of reduction is to find a linear transformation which would highlight the differences between the presentations of shapes which belong to different classes. In other words, the space of the shape needs to be projected into another coordinate system, such that the new coordinate axes will be in the directions of the information relevant to the classification. If X is the vector whose dimension n needs to be reduced, and Y is the linear transformation of that vector of dimension m, the task is to find a transform matrix A of dimension n x m, which satisfies the following relation:

$$Y = A^T X \tag{1}$$

There are several ways in which dimension reduction can be accomplished and each of them defines an optimum criterion that needs to be minimized. The criterion used in this research is based on the measure of scatter [10]. Compared to other criteria, it yields the best results which are consistent with the goal of the reduction – improvement of classification efficiency. The procedure of this method is as follows:

The expected vector M_i and the covariance matrix Σ_i are added to each class. Based on these parameters, the within-class scatter matrix S_W and between-class scatter matrix S_B are defined as:

$$S_W = \sum_{i=1}^{L} P_i E\{(X - M_i)(X - M_i)^T / \omega_i\} = \sum_{i=1}^{L} P_i \Sigma_i \tag{2}$$

$$S_B = \sum_{i=1}^{L} P_i (M_i - M_0)(M_i - M_0)^T \tag{3}$$

where P_i is the share of each class in the total number of all shapes. It is computed as the ratio of the number of shapes in a single class to the total number of all shapes. ω_i represent the *i*-th class. M_0 is the expected vector of mixture distribution for all the classes together:

$$M_0 = \sum_{i=1}^{L} P_i M_i \tag{4}$$

In Equations (2), (3) and (4), the value of L is 3 since there are three different classes.

The criterion J (5) is defined by means of the matrices from Equations (2) and (3) and the minimization of this criterion yields the transform matrix:

$$J = S_W^{-1} S_B \tag{5}$$

A random n-dimensional vector X is approximated by the m-dimensional vector Y, applying the minimization criterion J, whose end result is the transform matrix A:

$$A = \begin{bmatrix} \Psi_1 & \Psi_2 & \dots & \Psi_m \end{bmatrix} \tag{6}$$

where Ψ_i ($i = 1,...,m$) are the eigenvectors of the matrix $S_W^{-1} S_B$ to which the highest eigenvalues correspond. In other words, the following relation is true:

$$(S_W^{-1} S_B)\Psi_i = \lambda_i \Psi_i \quad i = 1,...,n \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \tag{7}$$

where λ_i ($i = 1,\ldots,n$) are all eigenvalues of the matrix $S_W^{-1} S_B$.

The transform matrix A contains weight coefficients. The multiplication of the transposed matrix A by 7D feature vectors X results in reduced 2D vectors, where a greater weight as assigned to more representative features and vice-versa. The order of the features in vector X is given by (8), and the resulting transform matrix A is presented in (9).

$$X = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & X_5 & X_6 & X_7 \end{bmatrix}^T \tag{8}$$

$$A = \begin{bmatrix} -0.0465 & -0.0716 \\ 0.0110 & -0.0244 \\ -0.0041 & 0.0011 \\ -0.6546 & 0.5152 \\ 0.5620 & -0.6786 \\ -0.0149 & 0.1357 \\ 0.5013 & -0.4999 \end{bmatrix} \tag{9}$$

An analysis of the results leads to the conclusion that the most representative are the fourth, fifth and the seventh feature, X_4, X_5 and X_7, or the shape to bounding box area ratio, the shape to bounding box circumference ratio and the circle to shape circumference ratio. Compared with these three, the other features carry considerably less weight and play a smaller role in the formation of reduced vectors. By representing shapes in 2D space, the results of the classification of the shapes can be graphically displayed.

IV. Classification of Shapes using Multi-layer Neural Network

There are lot of techniques which are used for classification problems. One approach for solving problem of classification geometric shapes is presented in [3]. That approach use linear classifiers, more precisely structure known as piece-wise linear classifier [11]-[12]. In this paper, for purpuse of classification of geometric shapes myltilayer neural network is used.

Neural network is designed in few steps. At the beginning, it is important to choose appropriate structure of neural network and to adjust activation functions of nodes in the hidden and output layer. Afterwards, desired output vectors which are used for training of neural network are formed. Process of training of neural network is done with 90 input feature vectors of cube, pyramid and cylinder (each shape per 30 feature vectors). Testing of neural network is done with completely new set of shapes and their feature vectors. Testing set contains also 90 feature vectors.

Results of testing neural network shows how much neural network is capable to learn and generalize in case of different input vectors.

Due to the fact that obtained feature vectors are 7D vectors, input layer of neural network has seven nodes. Number of nodes of output layer is the same as number of existing classes, hence three. Depending on which class belongs to the shape, the corresponding output node should have the highest value. The number of hidden layers and nodes in them are mostly determined experimentally. If the network does not converge to the exact solution of the problem, it is necessary to increase the number of nodes in the hidden layers. However, the number of nodes in the hidden layers should not be overmuch increased, because consequence can be overtraining of neural network. In that case, neural network works perfect for training set and very bad for new testing set. After numerous tests and experiments conducted structure of neural network which gives very good results is chosen. Finally, neural network has only one hidden layer with two nodes. Activation functions of all nodes of the input, hidden and output layer are bipolar sigmoid function.

Structure of used neural network is given on Fig. 8:

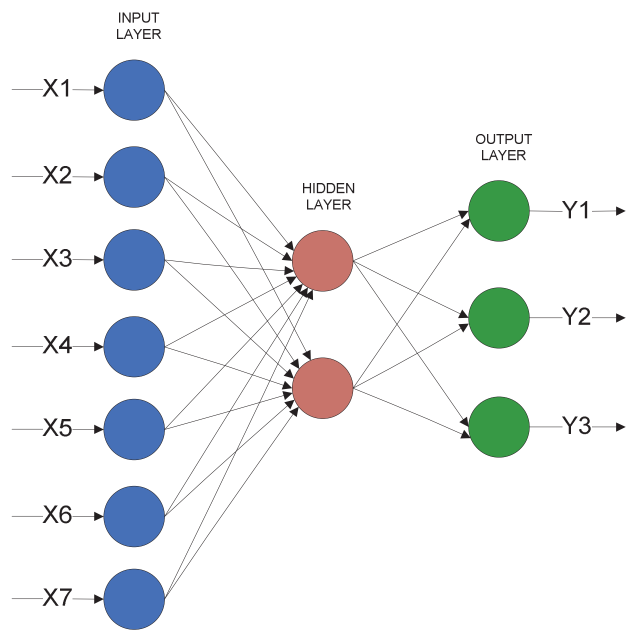

Fig. 8. Structure of neural network.

There are three possible ways to train neural network – supervised learning, unsupervised learning and reinforcement learning. For training of neural network in this work supervised learning is applied. In case of supervised learning for each input signal the appropriate desired output signal is given. In this way the neural network is trained using pairs of input/desired output.

If the feature vector that corresponds to cube shape is the input of neural network, appropriate output vector is:

$$d = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \tag{10}$$

In case that the feature vector that corresponds to pyramid shape is the input of neural network, appropriate output vector is given by (11):

$$d = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \tag{11}$$

When input of neural network is the the feature vector that corresponds to cylinder, appropriate output vector is given by (12):

$$d = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \qquad (12)$$

The difference between actual output and desired output is measured in each iteration by the error signal generator, which then generates an error signal. The role of the error signal of the neural network is to correct its weight to small differences between actual and desired output. When the new feature vector is set as input, depending on the output node that is the highest, neural network decide about affiliation of shape to some class. If all outputs are small, neural network can not decide which class belongs the shape.

Training of neural network is done through 150 epochs. At the end of training process, error is less than 10^{-8}. Graph of the training of neural network is shown on Fig. 9:

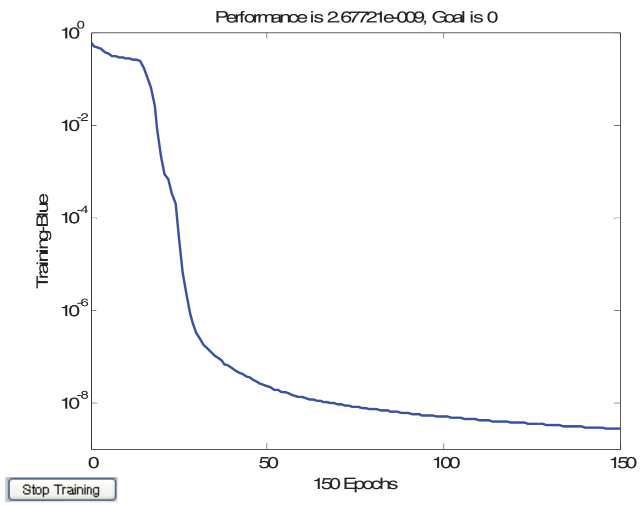

Fig. 9. Training of neural network.

Initial weight values of multilayer network have strongly influence to the final solution. The weights are usually initialized with small random values and these values should be different. The initial values of weight should not be large, because of the possibility that the system gets stuck in a local minimum. Equal weight of the initial value can not properly train the neural network, because in many cases, the solution requires that the weights are different.

The data for training should be selected on way to cover the entire space of input variables. Training set must be sufficiently informative and representative in order of classification.

In this case, direct inputs are 7D feature vector shapes. On this way all the features are represented and equally affect on the classification. Reduced 2D feature vectors are used only for visual representation of classification results.

In case of applying linear classifiers both the classification process and representation of classification results are using only the reduced 2D feature vectors. This procedure largely favors two features, while the others features are significantly less presented.

Fig. 10 shows the results of classification for the training set of shapes. In this case neural network exactly recognized all shapes, as expected due to the fact that the highest percent of success of neural network is represented in the case of training set. Especially because of the fact that the process of training is done using desired outputs and on that way neural network knows which output is desired for each input.

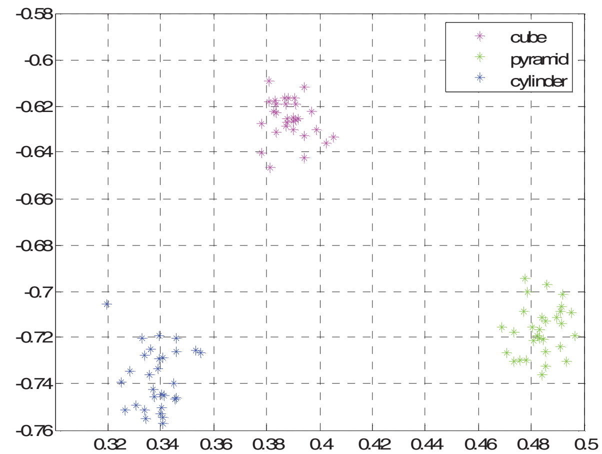

Fig. 10. Results of classification for training set of shapes.

Table II shows success of classification for training shapes in sense of number of correctly classified shapes. Neural network in this case has all shapes correctly classified. Percent of success of classification is 100%. This result is expected due to the fact that neural network has been trained using those shapes.

TABLE II
MATRIX OF SUCCSESS FOR TRAINING SHAPES

	Cube	Pyramid	Cylinder
Cube	30	0	0
Pyramid	0	30	0
Cylinder	0	0	30

The testing procedure is performed using a completely new set of input vectors that represent new shapes. Success of testing process shows how much neural network has learned in the training process and the extent of its ability to generalize in the case of new shapes. Results of classification for the testing set of shapes are shown in Fig. 11:

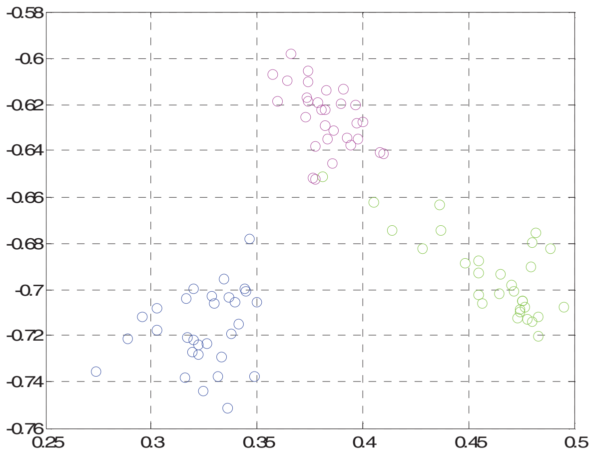

Fig. 11. Results of classification for testing set of shapes.

In the case of testing set, results of classification by neural network are slightly worse, as expected. Classes are less linear separated, and a few shapes are not classified correctly.

Table III shows success of classification for testing shapes in sense of number of correctly classified shapes. Neural network in this case has two shapes wrongly classified. Percent of success of classification is 97,78%. Overall conclusion is that the results of the classification using neural network are extremely good, considering the proper classification of most shapes. This means that the neural network showed a remarkable ability to generalize in the case of new entry shapes.

TABLE III
MATRIX OF SUCCSESS FOR TESTING SHAPES

	Cube	Pyramid	Cylinder
Cube	29	1	0
Pyramid	1	29	0
Cylinder	0	0	30

V. CONCLUSION

Thanks to new technologies, the field of pattern recognition is constantly evolving, improving and often is used in solving practical problems. Many systems used today in science and everyday life, can not imagine without some form of implementation of pattern recognition algorithms or embedded decision rules. Pattern recognition has especially found application in the field of robotics, given the fact that recent researches in robotics develop in the direction of increasing robot autonomy and independence in working tasks.

Robot observe the world around itself by one or more cameras that are fixed on its construction.

Therefore, data from the camera in the form of images or image sequences are input of many algorithms related to the functioning of the robot in its environment. This paper deals with the task of recognition of geometric shapes based on images that are generated by the camera using multilayer neural networks.

Neural networks are due to its characteristics of learning, memory and generalization chosen as a suitable method for solving problems of recognition and classification of geometric shapes. In this paper multilayer neural network has generated very good results and demonstrated a high degree of generalization in the case of new shapes. In subsequent researches, other types of neural networks can be applied, such as radial neural network. Different type of training could be also applied and analized. Further research can be related to the recognition of shapes in different environments of the robot in purpuse of its efficient interaction with the environment during the execution of working tasks.

ACKNOWLEDGMENT

This paper was partially funded under projects III44008 and TR35003 sponsored by the Ministry of Education and Science of the Republic of Serbia and under the CARE-Robotics project supported by the SNSF SCOPES IP program within the scope of IZ74Z0_137361/1.

REFERENCES

[1] Christopher M. Bishop, "Neural Networks for Pattern Recognition", Oxford University Press, 2005.

[2] Plamondon R, Srihari, "Online and Off-line Handwriting Recognition: A Comprehensive Survey", IEEE Transactions on Pattern Analysis and Machine 22(1):63–84, 2000.

[3] Wenyin L, "On-line Graphics Recognition", State-of the-art. In: GREC 5th IAPR International Workshop on Graphics Recognition. Springer. pp. 291–304, 2003

[4] Arica N, Yarman-Vural, "An Overview of Character Recognition Focused on Off-line Handwriting", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 31(2):216–33, 2001.

[5] Stallings W, "Approaches to Chinese Character Recognition". Pattern Recognition Society Monographs 8(2):87–98, 1976.

[6] Sing-Tze Bow, "Pattern Recognition and Image Preprocessing, Second Edition", Marcel Dekker, Inc, 2002.

[7] Sofija Spasojević, Željko Đurović, "Identification and classification of geometric shapes in 2D plane", 56. Conference ETRAN, Zlatibor, 2012.

[8] Tsai W-H, "Moment-preserving Thresolding: A New Approach. Computer Vision, Graphics, and Image Processing", 29(3):377–393, 1985.

[9] Canny J. A, "Computational Approach to Edge Detection". IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679–98, 1986.

[10] Fukunaga Keinosuke, "Introduction to Statistical Pattern Recognition", Academic Press, 1990.

[11] Christopher M. Bishop, "Pattern Recognition and Machine Learning", Springer, 2006.

[12] J. P. Marques de Sa, "Pattern Recognition, Concepts, Methods and Applications",Springer,2001.