SERBIAN JOURNAL OF ELECTRICAL ENGINEERING
Vol. 6, No. 3, December 2009, 507 - 514

UDK: 004.932:[004.8.032.26:007.52

One Solution to Recognition of Artistic Pictures for Guide Robots by Using Artificial Neural Networks*

Luka Lukić[1]

Abstract: In this paper is presented one solution to efficient, robust and cheap recognition of artistic pictures on the walls of museums and exhibit halls that reveals satisfactory measure of universality in order to be applied in the areas of trade, process industry, quality control, etc. This solution can be used in a wide range of applications where there is a demand of classifying objects on basis of their visual properties in a large number of existing classes. Here is proposed a method of selective grouping of pattern vectors as training sets for classifiers (artificial neural networks in this case), providing a smaller number of hidden layers in networks, achieving more precise performances and significantly expanding a number of classes to be classified. Selection approach is used in the very classification as well – neural networks are fed with input pattern vectors chosen from subsets determined by additional coefficients.

Keywords: Artistic pictures, Guide robot, Recognition, Classification, Museum, Artificial neural networks.

1 Introduction

On one hand, assuming that colors and ratio of the image dimensions (height and width) are one of the most representative properties of the artistic image that define it unambiguously in the process of appearance based recognition, and, on the other hand, having in mind the most required attributes of a good image classifier: 1. independence of camera resolution, 2. independence of distance between a camera and an object, 3. independence of smaller and midsize rotations of the camera, 4. some degree of robustness to noise, a vector that fuses normalized histograms of R, G and B image channel is chosen as the pattern vector [1-2]. A quotient of image height and width is chosen as an additional factor that determines sets of training and classifying vectors. In this approach, feedforward neural nets are used as classifiers – exactly one neural network is assigned to the each picture.

[1]Graduate Student, Faculty of Technical Sciences, University of Novi Sad, Serbia
E-mail: L2Lukic@yahoo.com

*Award for the best paper presented in Section *Robotics and Flexible Automation*, at Conference ETRAN 2009, June 15-19, Vrnjačka Banja, Serbia.

2 Image Processing and Training of ANNs

A Image processing

After image $f(x, y)$ is captured, image processing is performed in the two groups of operations: 1. *Image processing and locating regions of interests* and 2. *Transformations of segments and filtering*. The first group of operations on the image is composed of seven steps: image capturing (step I), after that the image is transformed from RGB space to gray domain (step II). The following step (step III) is image thresholding by using the global threshold method proposed by Otsu [3], after which the values of obtained binary image are inverted (step IV). Morphological operation opening (step V) paves the way for region filling (step VI). After these steps, we can derive regions of interest and analyze them as separate images (step VII). Fig. 1 shows operation taken in order to distinguish regions of interest. Before we apply the second set of operations, we want to assure that false regions will not be analyzed which are created after thresholding and region filling. Algorithm for elimination of false regions in the image is given in Fig. 2.

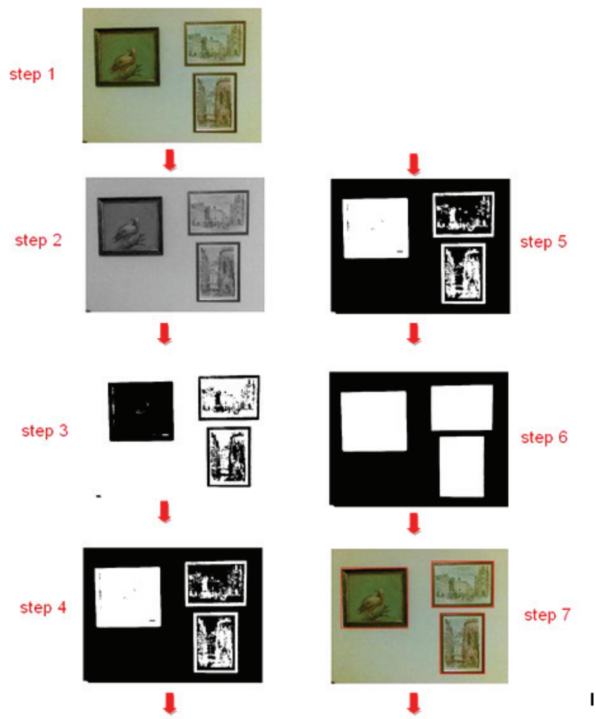

Fig. 1 – *The first set of operations on the image* (*Image processing and locating regions of interests*).

```
ELIMINATION_OF_FALSE_REGIONS
MAX_NUM_PIX ← MAX(REGIONS.PIX)
NORM ← REGIONS.PIX/ MAX_NUM_PIX
REAL_REGION_NUMS ← []
FOR i = 1: REGIONS.NUMBER_OF_OBJECTS
  IF NORM(i) > CERTAIN_PERCENTAGE_OF_THE_LARGEST
    REAL_REGION_NUMS ← [REAL_REGION_NUMS, i]
  END
END
```

Fig. 2 – *Algorithm for elimination of false regions.*

After applying the procedure for elimination of false regions "*regions candidates*" are used for further analysis. Due to rotations of the camera around line perpendicular to the plane of the picture, artistic pictures can be projected in one of the following scenarios: 1. as a rectangle (there are no rotations), 2. as a trapezoid (one angle rotation) and 3. as a quadrilateral (two or three angle rotations), which is the most common situation. Reconstruction of the original picture from a deformed one is done by using rubber-sheet transformations. Fig. 3 shows rubber-sheet transformations on the picture.

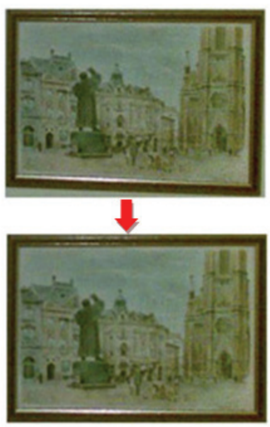

Fig. 3 – *Rubber-sheet transformations.*

In order to reduce stochastic noise acquired during image acquisition, a simple smoothing filter is applied:

$$g_i(x, y) = \frac{1}{9}\sum_{s=-1}^{1}\sum_{t=-1}^{1} f_r(x+s, y+t), \quad t = 1, \ldots, K \tag{1}$$

with:

K - number of valid regions for analysis,

f_r - image after rubber-sheet transformations.

The next step is calculation of image histograms over all the three channels: R, G and B channel:

$$h_{i,c}(m) = \left|\{(x,y) \mid g_{i,c}(x,y) = m\}\right|, \quad c = \{\mathrm{R},\mathrm{G},\mathrm{B}\}, \quad i = 1,\ldots,K\,. \tag{2}$$

Histogram sensitivity to the change in resolution of image is annulled by dividing it with its maximum value:

$$H_{i,c}(m) = h_{i,c}(m) \big/ \max(h_{i,c}(m)), \tag{3}$$

causing that histogram values are now within the range [0,1].

B Selection of Subsets of Pattern Vectors

Selection and determination of sets of training pattern vectors for training of neural network are based on analysis of the average value of k_{mi} that is a quotient of image height and width, and the maximal deviation r_i determined as following:

$$k_{m,i} = \frac{1}{n}\sum_{i=1}^{n} k_{i,j}\,, \quad j = 1,\ldots,n\,, \tag{4}$$

$$r_i = \max|k_{mi} - k_{i,j}|, \quad j = 1,\ldots,n, \tag{5}$$

then we form a similarity matrix, whose p-th column corresponds to the p-th candidate region and elements that belong to that column are indexes of the other region candidates whose average values are within the range $[k_{mp} - r_i, k_{mp} + r_i]$. Fig. 4 shows the algorithm for forming of similarity matrix.

```
FORMING_OF_SIMILARITY_MATRIX
MAT ← []
K ← LENGTH(REAL_REGION_NUMS)
FOR i = 1: K
  VECT ← [K_VECT >= D_LIMIT & K_VECT <= U_LIMIT]
   MAT ← [MAT, VECT^T]
 END
 MAT ← MAT – IDENTITY_MATRIX
```

Fig. 4 – *Algorithm for forming of similarity matrix.*

C Training of Neural Networks

Chosen architecture of neural networks is a feedforward two-layer (one input layer and one hidden layer) neural network with tangent sigmoid activation functions. Learning algorithm is scaled conjugate gradient algorithm [4]. Fig. 5 shows architecture of neural networks.

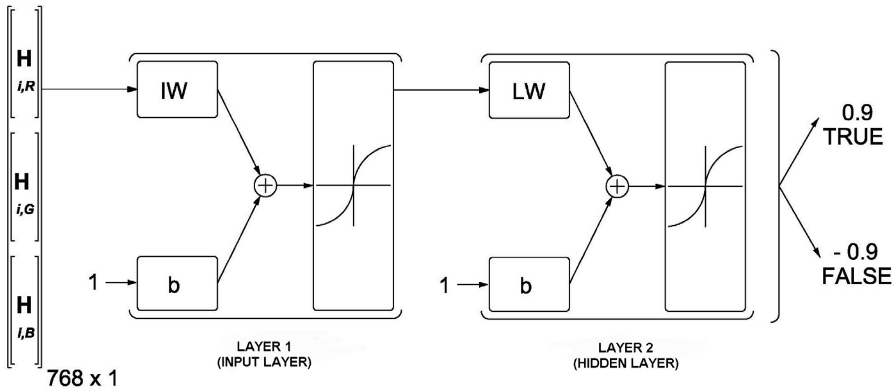

Fig. 5 – *Architecture of neural networks.*

Training of networks is performed in the following manner: for p-th picture, more precisely – for its normalized and transposed histograms of R, G and B channel that make one pattern vector with dimensions 768×1, p-th neural network is trained to give a value 0.9, but for all the other vectors that corresponds to pictures whose indexes are within the p-th column – to give a value –0.9.

3 Testing and Results

For the purpose of verification and testing of the proposed approach, we created a test platform in MATLAB® with a simple graphic user interface, shown in Fig. 6.

For training of neural networks, 16 different groups of pictures were taken, with 2, 3 or 4 different snapshots per one artistic picture, captured in different resolutions, under various light conditions and from different angles. Since training sets are relatively small, we took r_i as a 10% of k_{mi}, ignoring for a moment statement (5). For testing of responses to the change in light and angles of capturing, we took a picture *venice.* Fig. 7 shows results of recognition under various light and angular conditions.

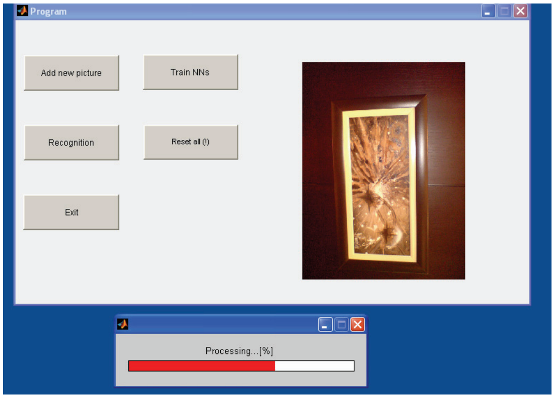

Fig. 6 – GUI *of a test platform in* MATLAB®.

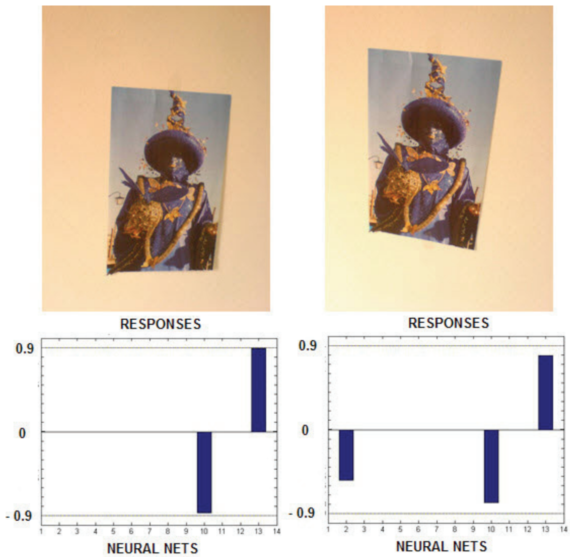

Fig. 7 – *Recognition under various light and angles*.

We can see from the figure that picture *venice* is similar to a picture with index 10, and that is *monastery*. Nevertheless, responses of neural networks that are assigned to them are significantly different, one has a large positive value (0.8893), which means "*hit*", but another one has a large negative value (-0.8708), which means "*miss*". The right part of the picture shows that in that case *venice* exhibits geometrical similarity to two other pictures – *monastery* and *churchyard*, but values of their responses show that appropriate recognition is achieved in this case as well. The second analysis related to recognition with artificially added noise, with $\sigma = \{5,10\}$. As an example, we took *mystique*.

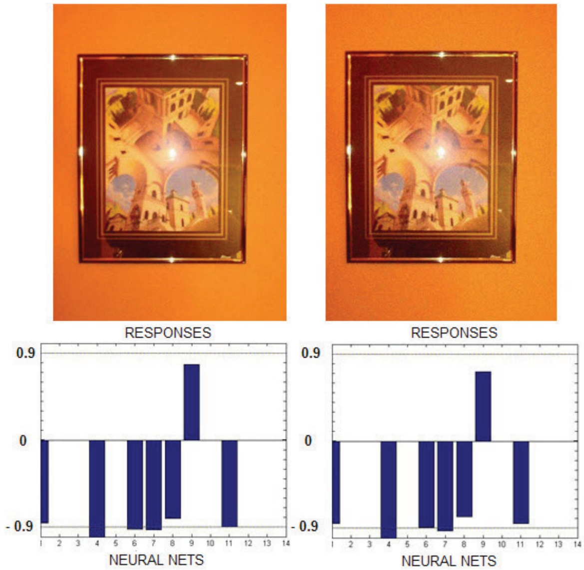

Fig. 8 – *Recognition with artificially added uniform noise.*

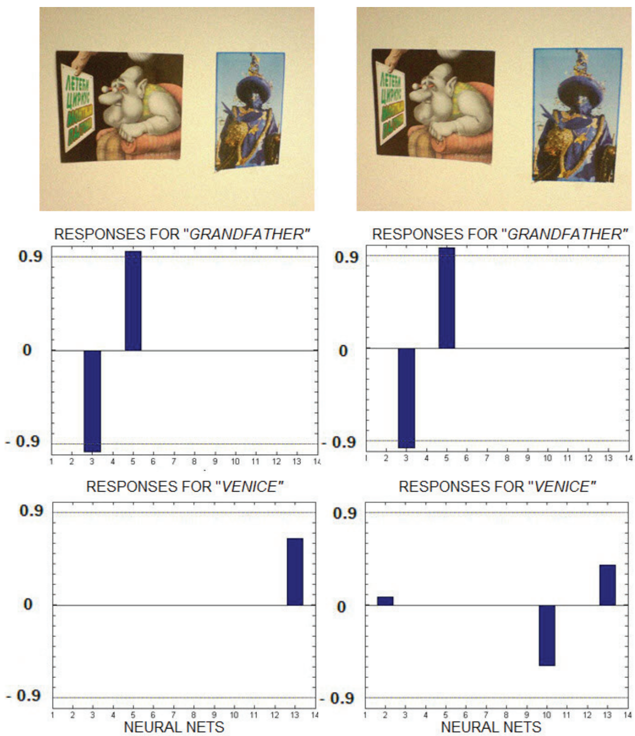

Fig. 9 – *Recognition in complex scenarios.*

Fig. 8 shows recognition of pictures with an artificially added stochastic noise. The left part is with noise $\sigma = 5$ and the right part is with $\sigma = 10$. We can see that we have pretty stable response to the added noise.

The last group of tests was related to complex scenarios in which there are two or more pictures, with different angles of capturing, various light conditions and artificially added uniform noise with $\sigma = 10$. Fig. 9 shows recognition of pictures in complex scenarios.

At the end, total percentage of recognition is 86%, e.g. successful recognition of 19 test pictures of overall 22 pictures. Pictures that were not appropriately classified were predominately monochromatic.

4 Acknowledgement

The author would like to acknowledge that he found inspiration for this work in a semester project given within a graduate course "Principles of Machine Vision" taught at the Department of Mechatronics by Professor Vladimir Crnojevic.

5 References

[1] V. Kecman: Learning and Soft Computing: Support Vector Machines, Neural Networks and Fuzzy Logic Models, MIT Press, Cambridge, USA, 2001.

[2] R.C. Gonzales, R.E. Woods: Digital Image Processing, 2nd Edition, Prentice Hall, New Jersey, 2001.

[3] N. Otsu: A Threshold Selection Method from Gray-Level Histograms, IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, No. 1, Jan. 1979, pp. 62 – 66.

[4] The Mathworks, Neural Network Toolbox™ 6, User's Guide, Online: www.mathworks.com