Cent. Eur. J. Comp. Sci. • 2(3) • 2012 • 236-244
DOI: 10.2478/s13537-012-0019-3

VERSITA

Central European Journal of **Computer Science**

Algorithm of traffic signs recognition based on the rapid transform

Research Article

Jan Gamec*, Daniel Urdzík†, Mária Gamcová‡

Deptartment of Electronics and Multimedia Communications,
Technical University of Košice, Park Komenského 13, 04120 Košice, Slovak Republic

Received 30 January 2012; accepted 17 August 2012

Abstract: This paper presents a model of a system for invariant object recognition, which consists of five stages. The first stage shifts the object so that the centroid of the object coincides with the center of the image plane. The second stage is an application of the polar-coordinate transforms used to obtain N-dimensional vectors-representations of the input object. In this stage, any rotation of the input object becomes a cyclic shift of the output value of this stage. The third stage employs CT (Certain Transform), a class of shift-invariant transformations to provide invariant representations for cyclically shifted inputs. The next stage normalizes the outputs of the previous stage to obtain scale invariance. The final stage realizes a classification.

Keywords: symbol recognition • rapid transform

1. Introduction

Symbol detection and recognition in an image, independent from its position, size and orientation – otherwise invariant symbol recognition ability – is a very important aspect of automatic symbol recognition [2, 3]. Many techniques developed for this purpose are based on existence of available libraries, which contain saved images (features) and recognition problem rests in searching for match between one of these features and symbols in the image. Searching for matches means comparison of features characteristics file and characteristics of input image. This procedure can take very long time, due to huge amount of input information, as well as information about known symbols (features). One very important task is to build up a system of detection and symbol recognition, which reduces this amount and which is able to reach fast recognition. One approach is the application of transformations of an input image and consecutive characteristics (features) selection for next processing which appropriately represent the symbol. Process of recognition based on extraction of features can be divided into several simpler processes:

* *E-mail: jan.gamec@tuke.sk*
† *E-mail: daniel.urdzik@tuke.sk (Corresponding author)*
‡ *E-mail: maria.gamcova@tuke.sk*

- Acquiring of input data intended for recognition in a form which allows further processing (digitalization, preprocessing).
- Input information analysis: symbol detection in input image, extraction of features, which characterize input symbol.
- Classification: input symbol recognition, i.e. class definition of known (learned) symbols, which mostly matches against the characteristics of an unknown symbol using selected criteria.

Neural networks were successfully used at pattern extraction which is used for object recognition. Within the technology of neural networks, there is a group of images containing known objects placed into the pattern library. Network extracts characteristics from this group of pattern pictures, which are later used for comparison with characteristics of extracted unknown objects images. Despite the results reached in this area and existence of many various structures of these networks, their practical use has limitations in many cases. Most of already existing networks require many connections and processing units, the amounts rapidly increasing within enlargement of input image, due to invariant object recognition. For example invariance towards rotations and object size change can be reached in many cases only by integrating rotated and enlarged versions of these objects into the group of pattern images [1, 9].
For these reasons, we can say that symbol recognition invariant on shift, rotation or scale variation isn't trivial problem. Within the process of recognition, there are used fast translation invariant transform of the class CT (Certain Transform). Submitted model is intended for binary images and thus system is able to process large image areas.

2. Model of the system for invariant symbol recognition

Model of the system for invariant symbol recognition consists of five levels each of them has set its own specific function (Fig 1). In the first level symbol in the input image area shifts to the centroid of the symbol and matches the display area. In the second level coordinates are converted into polar ones by polar transformations. Outputs of this level are generally N-dimensional vectors which characterize the symbols. Operations accomplished in this level cause that any output symbol rotation is expressed by the cyclic shift of outputs of this level. In the third level scale factor is removed from the outputs of the previous level by normalizing of these outputs to obtain the invariance to change the size of the symbol. In the next level fast translational invariant transformation (rapid transform, modified rapid transform or other CT class transformation), which provides invariant representations of cyclically shifted inputs, is applied on the results of the third level. The last level is the classification network whose purpose is to identify the unknown input symbol based on selected criteria.

2.1. Symbol centering in the input image

Elimination of the information about position of an symbol in the input image, or to do operations with input image, which cause the information to be not essential, is needed for succeeding in symbol recognition invariant on shift of the symbol [1, 4]. This process is running in the first level of described model of the system by centering of the symbol in the input image that means symbol is shifted, so its centroid is matched to the center of image area. Module of discrete Fourier transform shows the invariance to shift too, despite that DFT usage in this level is not appropriate for the following reasons:

- DFT needs complex arithmetic operations, which complicates the network.
- Phase spectrum DFT is not invariant on shift and so it is not possible to use it in this level. This causes the loss of the information about symbol, which DFT phase spectrum does not contain.
- Symbol centering network is more preferable in comparison with the DFT to use for acquiring invariance to shift, which has quite simple implementation and which saves all information about input symbol.

Subsystem for symbol centering contains two directional sub-networks: shift network for horizontal direction and shift network for vertical direction. Each of these sub-networks contains K lower levels and on each of these levels, there is a symbol shifted by distance (in image elements - OP) equal to second power. This also means that on any j-*th* lower

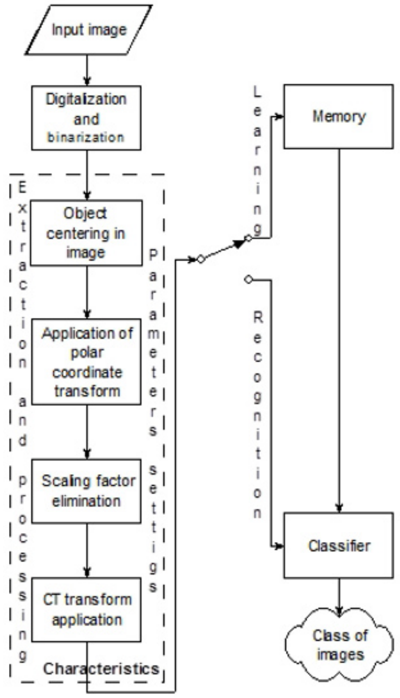


Figure 1. Block diagram of the system for invariant symbol recognition.

level the symbol is not shifted by distance $2j$ OP or any other shifting. Function of this level, shifting the symbol in vertical direction by 2 OP.

Symbol management calculate coordinates of symbol centroid in actual input image, which are needed for calculation of relating shifts in directional sub-networks of subsystem for symbol centering. Coordinates of the symbol centroid $(\bar{x}, \bar{y})$ can be calculated by formulae:

$$A = \sum_x \sum_y f(x, y)\,, \tag{1}$$

$$\bar{x} = \sum \sum x.f(x, y) \,/ A, \tag{2}$$

$$\bar{y} = \sum \sum y.f(x, y) \,/ A, \tag{3}$$

where $f(x, y)$ is value of pixel with coordinates (x, y), which can be 0 or 1 (binary images) and A is number of image elements with value 1 or can also express area (scale) of the symbol. After finding the $(\bar{x}, \bar{y})$ required shifts are calculated, which need to be done, for matching the symbol centroid to the center of image area. These shifts are later encoded into binary codes, based on which are for each lower level of directional shift networks generated by management circuit three managing signals [9].

Function of directional networks is similar to binary codes transfer mechanism into decimal representation. After calculation of values required for shifting in mentioned direction, will be whole operation of shift realized on K lower levels of the shift network for this direction. Centroid of the symbol doesn't have to be integral, but then its coordinates are not going to be matching the center of the image area. However, if the symbol is big enough, then this inaccuracy will not have very large impact on result of the recognition.

2.2. Application of a polar coordinate transforms

For the purpose of analyzing the symbol rotation around the axis going through its centroid, are in the second level of the described model applied polar coordinate transforms on outputs of the first level [1, 9–11]. Transformation of the information about the symbol into axis θ and r allows extracting of such a characteristic, within which rotation of the symbol affects as cyclic shift of their values.
General implementation of discrete polar transform has two dimensional structure of MxN points, organized as two dimensional space with axis r andθ. Value of the point $p(r, \theta)$ in polar plane is derived from corresponding value of image element (x, y). Each circle in image plane, independent from its scale, is represented by same amount of points N in polar plane. Because of the certain amount of image elements, which creates perimeter of a smaller circle is shown as higher amount of element in perimeter of big circle. For example circle with radius r=3 has perimeter 18 OP, so $p(3, \theta) = f(3, 0)$ for $-N/(2.18) < \theta < N/(2.18)$, where $f(3, 0)$ is value of image element (3, 0). Each point $p(r, \theta)$ can be acquired same.
Amount of N discrete values of angle θ needs to be chosen so, that within this transform were the least losses of the information about the symbol. For representation of the symbol are used characteristics – vectors, which reduce information included in MxN points of polar plane into N dimensional vectors set, enough characterizing symbol, where symbol rotation affects as cyclic shift of their values. Each elements of given vectors are calculated by formulae:

$$p_1(\theta) = \sum_r p(r, \theta), \tag{4}$$

$$p_2(\theta) = \sum_r r.p(r, \theta), \tag{5}$$

or

$$p_3(\theta) = \sum_r r^2.p(r, \theta), \tag{6}$$

etc.
Due to the similarity of these formulas with this one for calculation of moments are these vectors called moment of the rows.
Number of discrete angle θ values has great significance on this model's level. This parameter influences the quality of transforming process in terms of losing the information of the symbol and on the other side determines the size of entering characteristics i.e. degree of reduction of these information and so it determines the speed of the distinguishing process. Due to the facts, setting of the optimal value of this parameter is difficult and possible by setting up the experiment. Reaching the perfect recognition is in most cases possible only by using characteristics – moments of zero and first orders. Sometimes is necessary to use moments of higher orders for increasing the recognition ability. Better recognition ability of these moments is resulting from their structure, where the values of each image elements are "balanced" by higher orders of relevant values of r(6), which causes their increased sensibility for distortion of the symbol. Influence of usage of these moments on the results of the recognition was subject of many experiments.

2.3. Elimination of scaling factor

Elimination process of influence of scale variance is based on next formulas. Integrals are used to simplify, in sum operations in formulas 4 and 5. Let:

$$\int_r p(r, \theta)\, dr = p_1(\theta) \tag{7}$$

and

$$\int_r r.p(r, \theta)\, dr = p_2(\theta), \tag{8}$$

then

$$\int_r p(r/n,\theta)\,dr = n.p_1(\theta) \tag{9}$$

and

$$\int_r r/n.p(r/n,\theta)\,dr = n^2.p_2(\theta), \tag{10}$$

where n is mentioned scale factor. By next application of logarithm on results of integrals:

$$\log(n.p_1(\theta)) = \log(n) + \log(p_1(\theta)) \tag{11}$$

and

$$\log\left(n^2.p_2(\theta)\right) = 2.\log(n) + \log(p_2(\theta)). \tag{12}$$

Resulting from these formulas we know, that by eliminating of constants log(n) and 2.log(n) we acquire characteristics invariant on symbol scale shift. Removing of these constants can be reached by using:

$$p_i'(\theta) = \log(p_i(\theta)) - \max_\theta(\log(p_i(\theta))) + C, \tag{13}$$

where C is such a constant, where elements p_i' are always positive. In such a case, where $p_i(\theta)$=0, and $p_i'(\theta)$ is equal zero. It's not necessary to calculate the maximum value of log($p_i(\theta)$) in formula 13, because maximal element of vector p_i provides the previous level of this model. Realization of this model isn't very complicated.

2.4. Application of rapid translational invariant transform

Class of rapid translational invariant transforms (CT - Certain Transforms) [8] was derived from rapid transform (RT) by generalizing of operators pair $fs(a, b)$ where (s=1, 2) in original algorithm. RT is rapid, non-orthogonal, non-linear, translational and partly rotational invariant transform. Transforms of CT can be divided by using operators $fs(a, b)$. CT transforms are used within the processing of image information in robotics, in problematic of symbol recognition, can be used within estimating of movement [8] etc. One of their most important properties is speed, translational invariance, partial invariance on rotation and scale variance and simple technical realization. They were described in [5, 6, 8]. The most frequently used transforms of CT class are presented in Table 1.

Table 1. Frequently used transforms of CT class.

	RT	NT	MT	QT
$f_1(a,b)$	$a+b$	$max\{a,b\}$	$a \bigvee b$	$(a+b)^2$
$f_2(a,b)$	$\|a\text{-}b\|$	$min\{a,b\}$	$a \bigwedge b$	$(a-b)^2$
Input values	*real*	*real*	*binary*	*real*

2.5. Modified rapid transform

Resulting from properties of CT class transforms, they are not able to recognize reversed symmetrical inputs and in the process of recognition nor the reversed symmetrical symbols. For curing this unwanted invariance, there was designed modified rapid transform (MRT -Modified Rapid Transform), which preserves property of invariance to translation.
Modification of original rapid transform was done by putting the k (in general) modified steps before calculation of RT itself.
Operator $f0$, used in preprocessing MRT can be realized by next simple formula:

$$x_i' = f_0(x_i, x_{i+1}, x_{i+2}) = x_i + |x_{i+1} - x_{i+2}| \tag{14}$$

It's important for removing of unwanted invariance to reflection of input data, that the operator f0 is asymmetrical. Two 1-dimensional MRT for each direction can be used for 2-dimensional input signal, or one 2-dimensional MRT. Goal of this modification is removing of unwanted invariance to reflection of input data. For 2-dimensional MRT can be used in steps of preprocessing next modified operations:

$$
\begin{aligned}
f_0^x : x'(i,j) &= x(i,j) + |x(i+1,j) - x(i+2,j)|\,, \\
f_0^y : x'(i,j) &= x(i,j) + |x(i,j+1) - x(i,j+2)|\,, \\
f_0^{x+y} : x'(i,j) &= x(i,j) + |x(i+1,j) - x(i+2,j)| + |x(i,j+1) - x(i,j+2)|\,, \\
f_0^{xy} : x'(i,j) &= x(i,j) + |x(i+1,j) - x(i,j+1)|\,.
\end{aligned}
\tag{15}
$$

From consideration and computer simulations are resulting next results:

- MRT using f_0^x cannot suppress the invariance to reflection of input data in y direction.
- MRT using f_0^y cannot suppress the invariance to reflection of input data in x direction.

Both operators $f_0^{x+y}$and f_0^{xy}suppress the invariance to reflection in both direction. Due to their simplicity is operator f_0^{xy}preferable to operatorf_0^{x+y}.
In this model it's preferable to use NT and modified NT with functions $max(a,\ b)$ and $min(a,\ b)$. One of the reasons is that in next level of the model, it will be necessary to know maximal components of characteristics of the symbol - vectors $pi(\theta)$, i= 1, 2, ... and transform NT reorganize this components of vectors, so the maximal component will be the first in a row. Network doesn't have to be more complicated by functions, which search for maximal component of the vector.

2.6. Classification

Complete classification within the process of symbol recognition we describe as identification of unknown input symbol, which means to find class of symbols based on selected comparing criteria, which characteristics are mostly matching the characteristics on unknown symbol. It's possible to divide the methods into syntactic and mathematical. Mathematical methods of classification can be then divided into deterministic methods and statistic. Some of the simplest classifiers based on statistic methods are classifiers that are using Euclidian distance:

$$1 - dimensional \ : \ d_E^{(1)} = \sqrt{\sum_{i=1}^{N} (x(i) - \tilde{x}(i))^2} \tag{16}$$

$$2 - dimensional \ : \ d_E^{(2)} = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M} (x(i,j) - \tilde{x}(i,j))^2}, \tag{17}$$

where N, M are dimensions of 1-dimensional $\bar{X}$ or 2-dimensional $[X]$ of input signals and $\tilde{\bar{X}}$ or if you like $\left[\tilde{X}\right]$ is 1-dimensional or 2-dimensional reference signal of classifier.
Due to the character of extracted features - vectors pi, i= 1, 2,...N, the 1-dimensional Euclidian distance was used for classification of input of the previous levels $d_E^{(1)}$ 16.

3. Experimental and simulation results

We test several properties of the system for invariant symbol recognition using program tools, created in Matlab. In the experiments the system was tested whether it can recognize the most important prohibitory, priority and warning traffic signs, which were represented by several binary images with the size of 347x347 pixels (Figs. 3-5). Overall there were

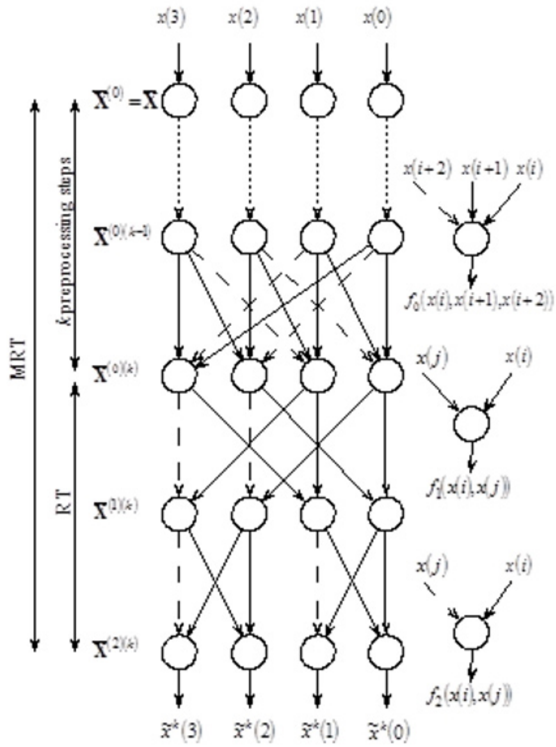


Figure 2. Signal graph of MRT calculation (for N=4).

two sets of binary images. The first set of the images was obtained by binarization of the images without the edge detection (Fig. 5a). The second set of the images was obtained by binarization of the images which were processed by edge detection algorithm (Fig. 5b). The goal of the experiments was also to test the system's ability to recognize the 15 traffic signs under hardened conditions such as changed scale and shifted positions in the image. Each of the traffic signs was changed in scale from 0.25 to 1.26 of the original. In addition, the position of the scaled (scale 0.25) traffic signs in the image was significantly shifted (multiple of the size of recognized object horizontally and vertically) as can be seen in Fig. 4. The complete number of the images with shifted and scaled images resulted in set of 91 tested images. The traffic sign recognition was tested with the RT and MRT. The successfulness of traffic sign recognition with RT or MRT was tested with the number of four moments. However the MRT as the more sophisticated algorithm which has the option of preprocessing was set to different number of preprocessing. The number of preprocessing was set from 1 to 5. In this way we were able to examine the successfulness of the traffic sign recognition with MRT under different settings.

The results of the experiments were as follows. The recognition with the RT transform was able to recognize 85% of the overall set (91 images) of traffic signs. The results of the recognition procedure with MRT are different in the manner of how many steps of preprocessing were selected. The influence of different settings of the number of preprocessing steps on the successfulness of the recognition with MRT is shown in the Table 2.

Table 2. Succesfulness of MRT.

Number of preprocessing steps	1	2	3	4	5
Successfulness	82%	87	91%	96%	78%

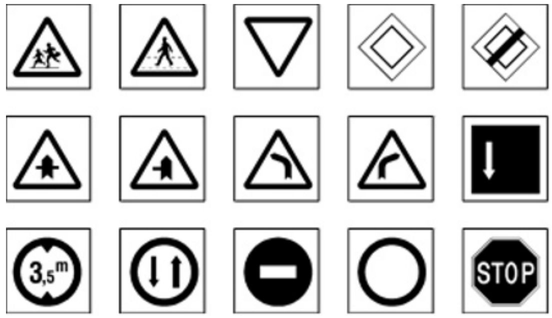


Figure 3. Images with traffic signs.

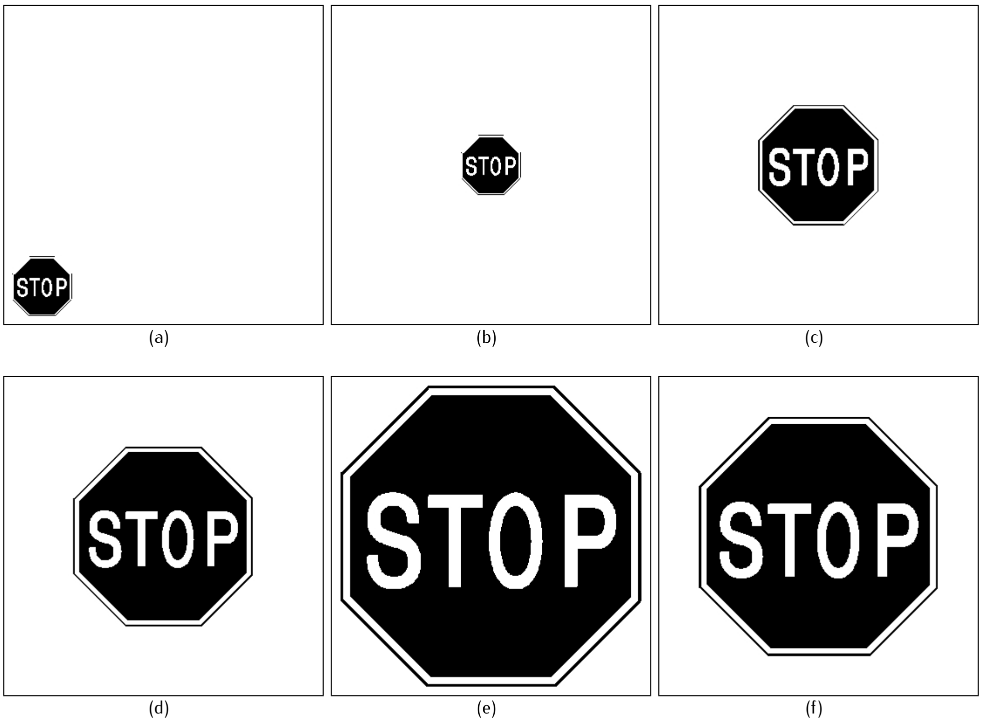
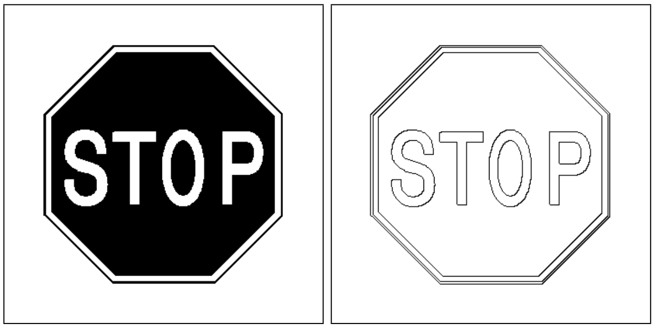


Figure 4. Shiftation and spatial tralnslation of the testing images: (a) shifted and scaled by factor 0.25, (b) scaled by factor 0.25, (c) scaled by factor 0.5 (d) scaled by factor 0.75, (e) scaled by factor 1.26, (f) original image.

4. Conclusion and future work

From the results of the presented experiments it is significant, that the RT and MRT have comparable successfulness, if the number of preprocessing steps of MRT is set to 1 or 2. With the increasing number of preprocessing steps the reliability of the MRT algorithm is increasing. However, if the number of the preprocessing steps is high, the MRT cannot work properly which results in a lower rate of success of the traffic sign recognition. The optimal number of the preprocessing steps appears to be 4, where the MRT was able to recognize the 96% of the tested traffic signs. The successfulness of the proposed algorithm was significantly decreased when the binarization of the input images were performed after the edge detection (Fig. 5b). The binary images which contain the information about the contours of the recognized objects contain less points of significance and hence are harder to be recognized by the algorithm. Also the experiments confirmed our notion, that the algorithm whether RT or MRT were not able to recognize the traffic signs



(a) (b)

Figure 5. Testing binary images: (a) without edge detection, (b) with edge detection.

scaled to 0.25 of the original. Such small images with traffic signs had relatively low quantity of information and hence could not be recognized. In the future work we want to focus on the traffic sign recognition on the gray scale images.

Acknowledgements

We support research activities in Slovakia / Project is co-financed from EU funds. This paper was developed within the Project "Centrum excelentnosti integrovaného výskumu a využitia progresívnych materiálov a technológií v oblasti automobilovej elektroniky", ITMS 26220120055

References

[1] Eason G., Noble B., Sneddon I.N., On certain integrals of Lipschitz-Hankel type involving products of Bessel functions, Phil. Trans. Roy. Soc. London, 529-551, 1955

[2] Ren F.X., Huang J., Jiang R., Klette, R., "General traffic sign recognition by feature matching", Image and Vision Computing New Zealand, 2009. IVCNZ '09. 24th International Conference, 23-25, 409-414, 2009

[3] Ruta A., Li Y., Liu X., "Detection, Tracking and Recognition of Traffic Signs from Video Input," Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference, 55-60, 12-15 Oct. 2008

[4] Shams S., Translation-, Rotation-, Scale-, and Distortion- Invariant Object Recognition Through Self-Organization, Int. J. Neur. Syst. 8, 173-179, 1997

[5] Turán J., Fast translation invariant transform and their applications. Košice: Elfa, 1999

[6] Turán J., Acoustic Object Recognition with Use of Rapid Transform, First Japanese-Czechoslovak Joint Seminar on Aplied Electromagnetics, 87-91, 1992

[7] Turán J., Recognition of Printed Berber Characters Using Modified Rapid Transform, J. Of Communications XLV, 24-27, 1994

[8] Verma, B. Recognition of Rotating Images Using an Automatic Feature Extraction Technique and Neural Networks, Int. J. Neur. Syst., 8, 201-207, 1997

[9] You S.D, Ford G.E., Connectionist Model for Object Recognition, Applications of Artifical Neural Networks III, 1709, 200-207, 1992

[10] You S.D, Ford G.E., Object Recognition Based On Projection, Proceedings of 1993 International Joint Conference on Neural Networks, 31-36, 1993

[11] You S.D, Ford G.E., Network model for Invariant Object Recognition and Rotation Angle Estimation, International Joint Conference on Neural Networks, 1993