THERMAL IMAGERY-BASED RECOGNITION USED NEURAL NETWORKS AND JOINT TRANSFORM CORRELATOR

JACEK MAZURKIEWICZ
Institute of Engineering Cybernetics
Wroclaw University of Technology
Phone: +48-71-3202681, Fax: +48-71-3212677
E-mail: Jacek.Mazurkiewicz@pwr.wroc.pl
50-372 Wroclaw, ul. Janiszewskiego 11/17
POLAND

JOANNA BAUER
Institute of Physics
Wroclaw University of Technology
Phone: +48-71-3202825, Fax: +48-71-3212677
E-mail: Joanna.Bauer@pwr.wroc.pl
50-370 Wroclaw, ul. Wybrzeże Wyspiańskiego 27
POLAND

ABSTRACT

This paper focuses on two different implementations of infrared-based biometric system. Artificial neural network and computer simulated joint transform correlator realize the classification and identification of human face's thermograms. Two types of neural networks are used: multilayer perceptron and Kohonen network. The results of the classification and identification of human face's thermograms are discussed based on different sizes and topologies of the Kohonen layer as well as based on different methods of neurons distance calculation used during teaching process. The accuracy of multilayer perceptron system and computer simulated joint transform correlator are evaluated using standard ratios for biometric systems: FRR, FAR and EER The problem of the low stability of thermal pattern was also tested.

Keywords: Multilayer Perceptron, Kohonen neural network, thermogram, identification

1. INTRODUCTION

Authentication methods based on recognition and comparison of infrared face images are considered to guarantee a third best security level among other biometric methods. Infrared imagery cross-over error rate places it just behind two eye recognition methods: retina and iris scanning.

Infrared face recognition is based on the fact that thermal image of each human's face is individual and unrepeatable. Even identical twins have different infrared images. The IR approach posses some advantages over visible wavelength images. First, thermal image is not vulnerable to disguises, can't be altered or camouflaged. The systems based on facial thermogram technology are external lighting independent and passive. This means that authentication can be done in low light or even in the total darkness, without physical contact and human cooperation [1, 2].

This technology also has some drawbacks. The main disadvantage is a low stability of thermal pattern. Thermogram can be disturbed by many external and internal factors, like environment temperature, illnesses, emotional state etc.

Fig. 1. Visible wavelength and thermal photographs of the same face

2. RESEARCH MATERIAL

The experiment was performed on 10 adult healthy volunteers. The group consisted of five males and five females, aged 23 to 40. The faces were recorded with AGEMA 900 LW system, at the ambient temperature 21,5°C and the 28% humidity of the environment. Spectral region from 8 to 13 μm was chosen as a sensing range. For every subject 10 images in different conditions were taken. From original photographs recorded in format of 135x270 pixel bitmaps, the square sample of central face part was cropped. This way database of 100 images of size 85x85 pixels was created.

3. KOHONEN NEURAL NETWORK

Suppose that an input pattern has N features and is represented by a vector $\boldsymbol{x}$ in an n-dimensional pattern space. The network maps the input patterns to an output space. The output space in this case is assumed to be one-dimensional or two-dimensional arrays of output nodes, which possess a certain topological orderness. The question is how to train a network so that the ordered relationship can be preserved. Kohonen proposed to allow the output nodes interact laterally, leading to the self-organising feature map. This was originally inspired by a biological model. For example a random sequence of two-dimensional patterns can be mapped to an array of output nodes, with a preserved topology.

3.1. RETRIEVING PHASE OF KOHONEN NEURAL NETWORK ALGORITHM

During the retrieving phase all neurons from Kohonen map calculate the Euclidean distance between the weights and the output vector and the winner neuron is the one with the shortest distance [3]. So each neuron from Kohonen map calculates the output value according to the classical weighted sum:

$$Out(i,j) = \sum_{l=0}^{N-1} x_l w_{lij} \tag{1}$$

where:

Out(i, j) output value calculated by single neuron from Kohonen map indexed by *(i, j)* if Kohonen map is rectangular, for 1-D Kohonen map we have only single index *i*,

x_l component of *N*-elements size input vector,

w_{lij} weight associated with connection from component of input learning vector x_l and neuron indexed by *(i, j)* if Kohonen map is rectangular, for 1-D Kohonen map we have only single index *i*.

3.2. CLASSIC LEARNING ALGORITHM FOR KOHONEN NEURAL NETWORK

The most prominent feature is the concept of excitatory learning within a neighbourhood around the wining neuron. The size of the neighbourhood slowly decreases with each iteration. The selection of a winner can be modulated by the frequency sensitivity of the output nodes. The other possible way is to modulate the learning rate by the frequency sensitivity. It is hard to say which solution is better or more accurate. So for the discussion presented in this paper we decided for the second possibility [3].

The learning algorithm is based on the Grossberg rule. All weights are modified according to the following equation:

$$w_{lij}(k+1) = w_{lij}(k) + \eta(k)\Lambda(i^w, j^w, i, j)(x_l - w_{lij}(k)) \tag{2}$$

where:

k iteration index,

η learning rate function,

x_l component of input learning vector,

w_{lij} weight associated with connection from component of input learning vector x_l and neuron indexed by *(i, j)* if Kohonen map is rectangular, for 1-D Kohonen map we have only single index *i*,

Λ neighbourhood function, (i^w, j^w) - indexes related to winner neuron, *(i, j)* - indexes related to single neuron from Kohonen map.

The learning rate η is a decreasing function, for presented discuss we assume linear decreasing form. Learning rate function is responsible for the number of iterations - it marks the end of learning process. This way there isn't any factor to determine if number of iterations is satisfactory. The neighbourhood function - often called Mexican Hat - Λ could be realised in different ways. The main problem is to determine a group of neurons which are neighbourhoods of the winner neuron. These neurons increase their output value during single learning step. The maximum increase is related to the winner, if neighbourhood neuron is further to the winner this increase is less significant. Neurons which are not neighbourhood should decrease their output value or their output value doesn't change. The presented solution is based on the following description of the neighbourhood function [3]

$$\Lambda(i^w, j^w, i, j) = \begin{cases} 1 & for \quad r = 0 \\ \dfrac{\sin(ar)}{ar} & for \quad r \in \left(0, \dfrac{2\pi}{a}\right) \\ 0 & for \quad other\ values\ r \end{cases} \tag{3}$$

where:

a neighbourhood parameter, can be changed during learning algorithm,

r distance from winner neuron to each single neuron from Kohonen map, calculated by indexes of neurons as follow:

$$r = \sqrt{(i^w - i)^2 + (j^w - j)^2} \tag{4}$$

The learning procedure is iterative. The whole algorithm can be described by following steps:

1. All weights are initialised by random values generated from range (-1, 1).
2. The winner neuron for each learning vector is created by calculating the net output using random values of weights with ordinary Kohonen map retrieving algorithm.
3. All weights are modified using Grossberg rule (1) for single learning vector x_l using current value of learning rate function as well as current value of neighbourhood function assuming the proper winner neuron created in step 2.
4. The learning rate function value is modified, the neighbourhood parameter a (2) is modified and if the learning rate function value is greater than zero step 3 is executed for the next learning vector, else the learning algorithm stops.

4. ARTIFICIAL NEURAL NETWORK BASED SYSTEM

The system itself consists of three main blocks (Fig. 2). The first stage performs data acquisition and preprocessing. Artificial neural network is the second block - main processing is done in this stage. The third (output) block transforms output vector of the NN into response of the whole system.

Input to the system is presented as 85x85 pixel 8-bit grayscale bitmap. In the first stage size of the bitmap is rescaled to 25x25.

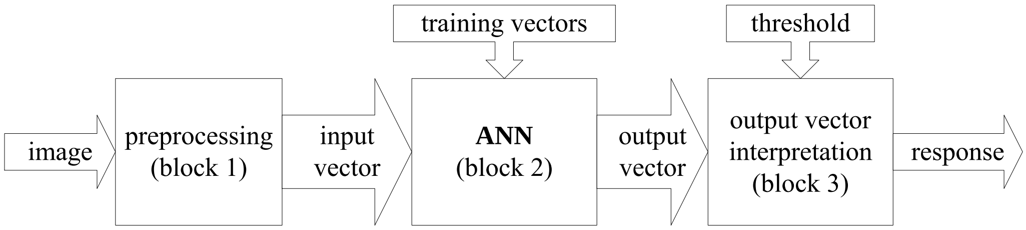

Fig. 2. Neural processing system model

As value of each pixel falls between 0 and 255 another rescaling has to be done in order to effectively utilize linear part of activation function. Also the system response should be invariant with respect to changes in environment temperature. Both problems can be solved with rescaling all pixel values in the input image to fall into range (0, 1). This is achieved by applying

$$x_i = \frac{p_i - min\{\, \boldsymbol{p}\,\}}{max\{\, \boldsymbol{p}\,\} - min\{\, \boldsymbol{p}\,\}} \qquad (5)$$

where:

x_i is i-th element of the input ANN vector $\boldsymbol{x}$,

p_i is i-th element of the input image vector $\boldsymbol{p}$,

to each pixel.

The second block is a Kohonen neural network or multilayer percetron. For Kohonen network eight examples of the net are prepared. The differences among them focused on topology and parameters related to the training process. The following names are used for network description:

- Epochs – number of training epochs,
- NeuronsX – number of neurons in horizontal line of network,
- NeuronsY – number of neurons in vertical line of network,
- Topology – kind of network topology – gridtop (GT) – rectangular or hextop (HT) – hexagonal,
- DistanceFunction – distance function between neurons: linkdist (LD),
- OPLearningRate – learning rate at the begin of training,
- OPSteps – number of training steps when the OPLearningRate is in use (1k = 1000),
- TPLearningRate – learning rate when the training process ends (2c = 0,02),
- TPNeighborHoodDistance – neighbourhood distance among the neurons when training process ends,
- Pattern2Learn – number of training vectors (photos) taken for each person.

The prepared examples of Kohonen neural network are presented in Table 1.

The multilayer percetron has 625 neurons in the input layer, one hidden layer consisting of 12 units (number of neurons in the hidden layer has been chosen experimentally). Response of the ANN is coded in the "1 of N" manner. Unipolar sigmoid function:

Table 1. Kohonen neural networks used during experiments

ANN	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**
Epochs	100	100	100	100	100	100	100	200
Neurons X	10	9	11	10	10	100	10	10
Neurons Y	10	9	11	10	10	1	10	10
Topology	GT	GT	GT	GT	GT	GT	HT	GT
Distance Function	LD	LD	LD	LD	LD	LD	LD	LD
OP Learning Rate	0,9	0,9	0,9	0,9	0,9	0,9	0,9	0,9
OPSteps	1k	1k	1k	1k	1k	1k	1k	1k
TP Learning Rate	2c	2c	2c	2c	2c	2c	2c	2c
TP Neighbor Hood Distance	1	1	1	1	1	1	1	1
Pattern 2Learn	5	5	5	3	6	5	5	5

$$f(u_i) = \frac{1}{1 + exp(-\beta u_i)} \qquad (6)$$

where:

u_i is weighted sum of i-th perceptron input values,

β is parameter defining “steepness” of activation function f,

was chosen for the nonlinear activation element, so all neuron output values fall in range (0, 1).

Training set consisted of 3 photos of five first subjects. The net was trained with use of backpropagation algorithm with momentum and variable learning rate. Mean square error measure (MSE):

$$MSE(\boldsymbol{o}, \boldsymbol{t}) = \frac{1}{N} \sum_{i=1}^{N} (o_i - t_i)^2 \qquad (7)$$

where:

o_i is i-th element of network output vector $\boldsymbol{o}$,

t_i is i-th element of target output vector $\boldsymbol{t}$,

N is the number of output neurons,

was used as the base of network energy function.

The third block of the system as an input takes an output vector of the ANN. It points the photo which is the real answer using the threshold level to cut off the lower output values from these neurons which are not the winners for the proper category for Kohonen network. For multilayer perceptron it performs MSE calculation of output vector o against the set of target output vectors T. Subject is considered authorized iff MSE of at least one of target output vectors ti $\in$ T with o is lower than threshold d. Index i of neuron with the highest response, max{oi}, indicates the class to which the corresponding input vector (and the subject itself) belongs.

5. JOINT TRANSFORM CORRELATOR APPROACH

The joint transform correlator (JTC) is one of the main optical architectures (except for the Van der Lugt correlator) that are being used for purposes of pattern recognition. There exists a comprehensive bibliography describing the applications of JTC, as e.g. [3, 4]

In this study we have used computer simulation of the JTC. As the basis of an optical correlator the fundamental relation for optical signal correlation was applied:

$$s * f = FT^{-1}(S \cdot F^{*}) \tag{8}$$

where:

s and *f* are functions (images) being correlated,
S and *F* are Fourier transforms of *s* and *f*, respectively,
FT^{-1} is the inverse Fourier transform.

For evaluation of the correlation quality two criterions were used: Discriminant Capability (DC) and Peak-to-Correlation Energy (PCE). DC characterizes the ability to recognize the target image against non-target (the recognized thermogram) and is defined as:

$$DC = \frac{CC}{A} \tag{9}$$

where:

CC is non-target correlation signal (cross correlation),
A is target correlation signal (autocorrelation).

Peak-to-Correlation Energy is defined as:

$$PCE = \frac{|c(0,0)|^2}{E_c} \tag{10}$$

where:

$c(0,0)$ is the highest value of the correlation peak,
E_c is the correlation plane energy:

$$E_c = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy |c(x,y)|^2 \tag{11}$$

The PCE calculated for a high and sharp correlation peak has a larger value in comparison to the case of low and broad correlation peak [5].

6. EXPERIMENTS RELATED TO KOHONEN NETWORK

The first set of experiments was focused on the classification of original infrared photos. The test set of photos includes 200 pictures: 100 of them were directly taken from infrared camera (some of them were used during training process), 100 others were changed by luminance and contrast regulations. The results of classification are presented in Table 2.

It is easy to notice that the number of correct classifications is quite similar for tested Kohonen neural network topologies – about 75 % of answers are correct. The number of epochs realized during training process as well as kind of topology (grid or hexagonal) has no significant influence for the results.

The second set of experiments was realized to test the stability of thermal patterns. The infrared photos were modified to simulate different illnesses which are reflected on humans face temperature picture: fever or problems with teeth for example. The same Kohonen networks – trained by original photos only - were used, as for the first set of experiments. We used 3 sets of 100 modified pictures for classification. The results are presented in Table 3, 4, 5.

Table 2. Results of classification for original infrared photos

ANN	#1	#2	#3	#4	#5	#6	#7	#8
NCC	150 75,0 [%]	147 73,5 [%]	148 74,0 [%]	158 79,0 [%]	135 67,5 [%]	148 74,0 [%]	161 80,5 [%]	140 70,0 [%]
NIC	50 25,0 [%]	53 26,5 [%]	52 26,0 [%]	42 21,0 [%]	65 32,5 [%]	52 26,0 [%]	39 19,5 [%]	60 30,0 [%]

Table 3. Results of classification for test of stability of thermal patterns – set #1

ANN	#1	#2	#3	#4	#5	#6	#7	#8
NCC	86 86,0 [%]	86 86,0 [%]	8 8,0 [%]	0 0,0 [%]	86 86,0 [%]	86 86,0 [%]	86 86,0 [%]	86 86,0 [%]
NIC	14 14,0 [%]	14 14,0 [%]	92 92,0 [%]	100 100 [%]	14 14,0 [%]	14 14,0 [%]	14 14,0 [%]	14 14,0 [%]

Table 4. Results of classification for test of stability of thermal patterns – set #2

ANN	#1	#2	#3	#4	#5	#6	#7	#8
NCC	88 88,0 [%]	88 88,0 [%]	0 0,0 [%]	0 0,0 [%]	88 88,0 [%]	88 88,0 [%]	88 88,0 [%	88 88,0 [%]
NIC	12 12,0 [%]	12 12,0 [%]	100 100 [%]	100 100 [%]	12 12,0 [%]	12 12,0 [%]	12 12,0 [%]	12 12,0 [%]

Table 5. Results of classification for test of stability of thermal patterns – set #3

ANN	#1	#2	#3	#4	#5	#6	#7	#8
NCC	91 91,0 [%]	91 91,0 [%]	10 10,0 [%]	10 0,0 [%]	91 91,0 [%]	91 91,0 [%]	91 91,0 [%]	91 91,0 [%]
NIC	9 9,0 [%]	9 9,0 [%]	90 90,0 [%]	90 90,0 [%]	9 9,0 [%]	9 9,0 [%]	9 9,0 [%]	9 9,0 [%]

where:
NCC number of correct classifications,
NIC number of incorrect classifications.

It is easy to notice that the number of correct classifications is identical for most of tested Kohonen neural network topologies – 86 % of answers are correct in set #1, 88 % of answers are correct in set #2, 91 % of answers are correct in set #3. The number of epochs realized during training process as well as kind of topology (grid or hexagonal) has no significant influence for the results. On the other hand for ANN #3 and ANN #4 the results are very bad. We can say that classification is completely incorrect. These results can be explained because of too low number of photos taken for training process (ANN #4) and too large number of neurons used to create the Kohonen layer (ANN #3).

The last set of experiments was focused on different norms to measure the distance among the neurons. Four norms were used: Euclid's norm, scalar product, norm L_∞, norm L_1. The Kohonen neural networks defined in Table 1 were tested by 100 original photos taken directly from infrared camera. The results of correct classifications are presented in Table 6.

The best results we noticed using Euclid's norm. Quite similar situation we can find for norm L_1 and there is rather no sense to used scalar product or norm L_∞ because of very poor results of correct classification. On the other hand the number of epochs realized during training process as well as kind of topology (grid or hexagonal) has no significant influence for the results.

Table 6. Results of correct classifications for different norms used to measure the distance among neurons

ANN	#1	#2	#3	#4	#5	#6	#7	#8
EN	68 68,0 [%]	70 70,0 [%]	71 71,0 [%]	62 62,0 [%]	57 57,0 [%]	68 68,0 [%]	72 72,0 [%]	69 69,0 [%]
SP	9 9,0 [%]	9 9,0 [%]	9 9,0 [%]	10 10,0 [%]	11 11,0 [%]	14 14,0 [%]	11 11,0 [%]	13 13,0 [%]
L_∞	9 9,0 [%]	17 17,0 [%]	25 25,0 [%]	21 21,0 [%]	25 25,0 [%]	22 22,0 [%]	19 19,0 [%]	20 20,0 [%]
L_1	8 8,0 [%]	61 61,0 [%]	58 58,0 [%]	64 64,0 [%]	63 63,0 [%]	67 67,0 [%]	60 60,0 [%]	59 59,0 [%]

where:
EN Euclid's norm,
SP scalar product,
L_∞ norm L_∞,
L_1 norm L_1.

7. ACCURACY MEASURES FOR BIOMETRIC SYSTEMS

Performance of biometric system is measured with two main parameters: false acceptance rate (FAR) and false rejection rate (FRR). FAR is the probability that a biometric system will incorrectly identify an individual or will fail to reject an impostor. FFR is the probability that a system will fail to identify a subject, or verify the legitimate claimed identity of an individual. Usually both ratios depend on one parameter. In ANN-based system it is the threshold of MSE (7) calculated for output vector o against desired output vector t. In the JTC case it's the threshold computed for described above criterions values, which are correlation measures. It is obvious that change in any of two measures implies change of another one in opposite direction. When threshold decreases the system becomes more rigorous, which results in decrease of FAR but an increase of FRR. Analogous situation occurs when threshold increases. Equal error rate (EER), sometimes called cross-over error rate, is evaluated at the intersection point of FAR and FRR plotted against threshold. Corresponding threshold value, called equal error point (EEP) is often used in a final system implementation. Another way to evaluate optimal threshold is finding minimum of weighted FAR+FRR function, which minimizes total number of authorization errors (regardless of their type).

8. ACCURACY ANALYSIS OF ANN-BASED SYSTEM

Error ratios have been averaged for 20 training sessions. For each rate sessions with different MSE training goals have been computed. To preserve clarity FAR and FRR have been presented on two different plots (Fig. 3., Fig. 4.).

One can notice that averaged FAR increases linearly with threshold value. This fact can be simply explained: during training sessions images of subjects to be rejected weren't presented to the network. It wouldn't be advisable, as during training ANN in a real implementation one cannot present images of all unauthorized subjects to the network.

It is obvious, that network response for images from classes, that haven't been used during training sessions, has a statistical character. In order to maximize system performance ANN should be trained to lowest possible MSE goal. Table 7 presents EER values for the same ANN configuration and for different training goals.

Table 7. Average neural system EER

	Training MSE goal	Average EER
1	0.010	0.1393
2	0.005	0.0858
3	0.001	0.0558

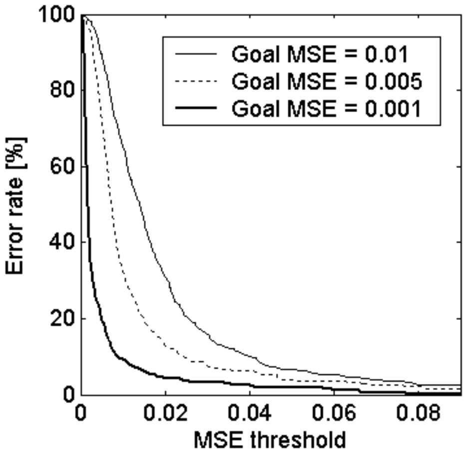

Fig.3. Average FRR of neural system

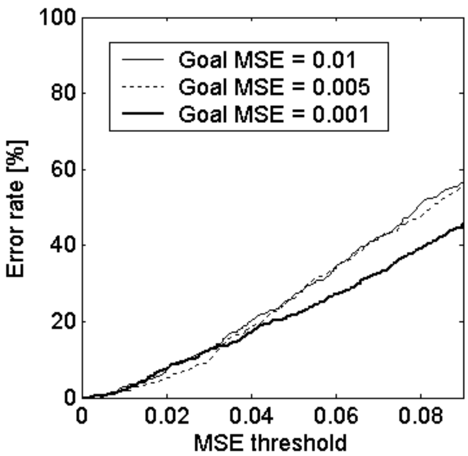

Fig. 4. Average FAR of neural system

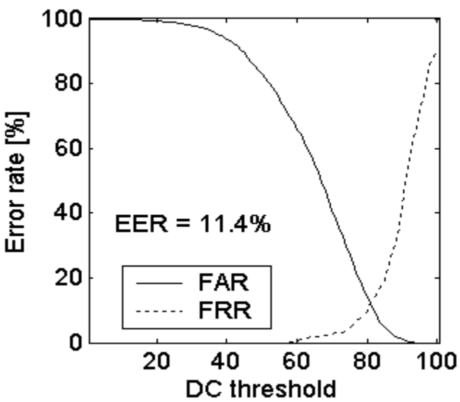

Fig. 5. JTC error rates based on DC criterion

9. ACCURACY ANALYSIS OF JTC BASED SYSTEM

JTC based thermogram recognition was used in identification mode, where the biometric system identifies a person from entire enrolled population by searching a database for a match. For accuracy evaluation DC (9) and PCE (10) criterions were used. The results are presented in Fig. 5. and Fig. 6.

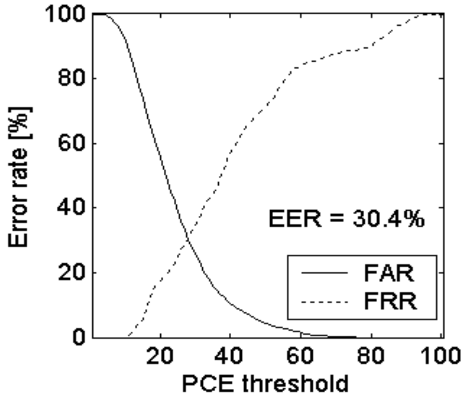

Fig. 6. JTC error rates based on PCE criterion

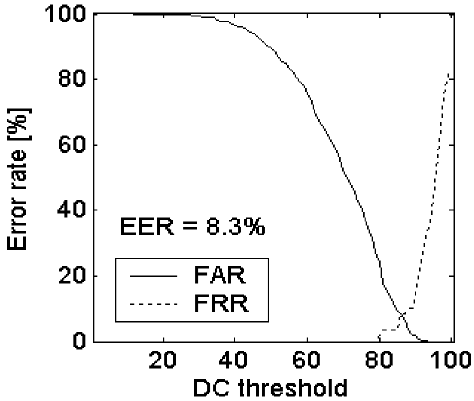

Fig. 7. JTC error rates based on DC criterion for averaged thermograms

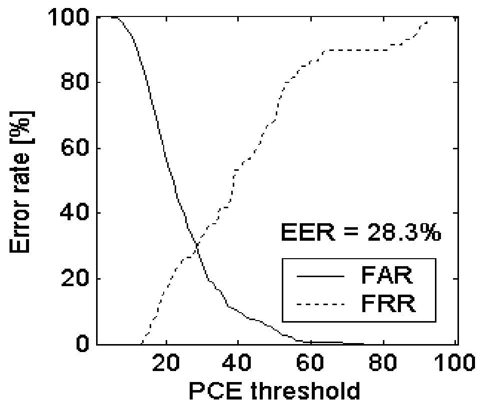

Fig. 8. JTC error rates based on PCE criterion for averaged thermograms

In the next step, for each person the mean images based on 5 different thermograms were created and accuracy evaluation was repeated (Fig. 7. and Fig. 8.). Generally in both cases using the DC criterion for correlation accuracy evaluation yield better results. Moreover the applying of thermograms averaging improved results of recognition from EER=11,4% to ERR=8,3% for DC and form EER=30,4% to EER=28,3% for PCE.

10. CONCLUSIONS

This paper focuses on implementation of infrared-based biometric system using neural networks. We have shown two complete neural biometric systems. Accuracy has been measured using standard ratios for this kind of applications.

The results of the classification and identification of human face's thermograms realized by artificial neural network are discussed based on different sizes and topologies of the Kohonen layer as well as based on different methods of neurons distance calculation used during teaching process.

The problem of the low stability of thermal pattern was also tested. The results seem to be very promising. In general the number of epochs realized during training process as well as kind of topology (grid or hexagonal) has no significant influence for the results. System which operate on infrared photos with artificial neural networks as a classifier devices can be used as very good alternative approach to classical methods focused on joint transform correlators.

Further research in this area will include testing the system with larger databases of subjects and development of better preprocessing methods in order to improve accuracy. Visible wavelength and infrared-based systems shall be compared. We will also make attempts to implement system in hardware.

REFERENCES

[1] D. A. Socolinsky, L. B. Wolff, J. D. Neuheisel, C. K. Eveland, Illumination Invariant Face Recognition Using Thermal Infrared Imagery, *IEEE Proceedings on Computer Vision and Pattern Recognition*, 3, 2001, pp. 156-165.

[2] J. Wilder, P. Phillips, C. Jiang, S. Wiener, Comparison of Visible and Infrared Imagery for Face Recognition*, Proceedings of the 2nd International Conference on Automatic Face & Gesture Recognition*, Killington, 1996, pp. 182-187.

[3] J. Colin, N. Landru, V. Laude et al., High-Speed Photorefractive Joint Transform Correlator Using Nonlinear Filters, *Applied Optics,* 1, 1999, pp. 283-285.

[4] K. Kodate, R. Inaba, E. Watanabe, T. Kamiya, Facial Recognition by Compact Parallel Optical Correlator, *Measurement Science and Technology*, vol. 13, 2002, pp. 1756-1766.

[5] B. V. K. Vijaya Kumar, L. Hassebrook, Performance for Correlation Filters, *Applied Optics*, 129, 1990, pp. 2997-3006.

[6] R. Hecht-Nielsen, *Neurocomputing*, Addison-Wesley, USA, 1990.

[7] J. Mazurkiewicz, J. Bauer, Thermal Imagery-Based Recognition Realized by Kohonen Neural Network, *10th International MENDEL 2004 Conference on Soft Computing*, Brno, June 16–18, 2004, pp. 117-122.

[8] J. Hertz, A. Korgh, R. Palmer, *Introduction to the Theory of Neural Computation*, Addison-Wesley, USA, 1991.

[9] S. Osowski, *Sieci Neuronowe w Ujeciu Algorytmicznym*, Wydawnictwa Naukowo-Techniczne, Warszawa, 1996.