A New Recognition Method for Natural Images

WEIREN SHI

College of Automation

Chongqing University

No.174, Shazheng Street, Shapingba District, Chongqing

People's Republic of China

wrs@cqu edu.cn

ZUOJIN LI †

College of Automation

Chongqing University

No.174, Shazheng Street, Shapingba District, Chongqing

People's Republic of China

cqulzj@gmail.com

XIN SHI

College of Automation

Chongqing University

No.174, Shazheng Street, Shapingba District, Chongqing

People's Republic of China

meetshixin@gmail.com

ZHI ZHONG

The Smartech Institute

Room 1808, Anhui Building, No.6007, Shennan Road, Futian District, Shenzhen，

People's Republic of China

zzhong@mae.cuhk.edu.cn

† corresponding author

Abstract: - Natural images recognition is an important area of machine vision. This paper presents a novel approach for natural images recognition, based on the non-Gaussian distribution property of natural images. In this new method for recognition, first supervised classification is conducted to the natural images based on their label value, then independent components linear transforms are conducted to each category of samples, high-dimensional data are transformed to irrelevant independent components, and finally the probability distance between independent component subspaces is used for unsupervised classification. This classification tree also shows some features of signal processing of biological optic nerve. Experiment on ORL Face Database identity recognition shows that this method is featuring high recognition rate and low time consumption; meanwhile, another experiment is conducted on direction determination of intelligent robots autonomous navigation, also producing a good result.

Key-Words: - natural image, independent component subspace, hierarchical discriminant regression, recognition of face， robot navigation

1. Introduction

The core of natural images recognition is how to transform the high-dimensional data collected by the machine to low-dimensional spaces and enable it to recognize the natural goals rapidly and efficiently.

Generally, the foremost problem is to extract features of natural images so as to lower the dimensionality of data. Natural images refer to images that can be observed by human optic nerve system or input visual information of human beings or machines. For machines, natural images are high-dimensional vectors. There are two existing methods for extracting features or lowering dimensionality, namely parametric model and linear transform. Extracting features based on parametric model requires different models for different features [1-4]. Although, this method applies well in recognizing natural images, medical images and handwriting [5-7], its dependence on subjective modeling limits its generalization ability; for example, it is hard to apply to computer visual recognition. The basic principle of the linear transform method is to extract features through linear transform to high-dimensional data, and its basic requirement is to choose suitable transform to reduce or remove redundant information. Now there are various methods to finish this task [8-10], such as Karhunen-loeve (K-L) transform, singular value decomposition, discrete Fourier transform, and wavelet transform, etc. [11]. Principal Component Analysis (PCA) is the optimum transform to extract features and compress data when the mean square error is at its minimum value, so it applies widely [12]. Although the PCA has wider adaptability than the parametric modeling method in application, the second-order statistics based method also has some shortcomings. For example, the extracted principal component is only determined by second-order statistics of the data, namely the self-correlation matrix, which can only describe fully the stationary Gaussian distributed data. However, the natural images collected by machine vision don't have Gaussian distribution property [13-15], which limits application of PCA to machine vision. But it is exciting that PCA shows its advantages in analyzing the high-order statistical characteristics of natural images and also in processing pictures and sound, digital watermarking and medical signals [16-21].

Next problem is how to recognize different objects in the extracted feature data. In this process, our goal is to realize high recognition rate and low time consumption. In recent years, Classifiers based on decision trees have been widely used due to its flexibility and generalization ability [22-25]. Hwang, W. (2000) and J.Y Weng, et al. (2001) [26-27] have proposed a classification and recognition idea featuring autonomous mental development through simulating the mechanism of human autonomous development and put this idea into practice through improving the decision tree, which has opened a new path in machine recognition. Compared with traditional decision trees, the method proposed by Hwang, W. (2000) [26] has enhanced the ability of autonomous learning and recognition of machine. However, during the node split process in this method, subspaces only display the general distribution of samples rather than to reflect the feature information of sample data; that is to say not much clustering information is reflected, which will impact on the recognition rate. On this account, this paper proposes a classification and recognition method of hierarchical discriminant regression (HDR) tree based on independent components subspace (ICS-HDR). Experimental results show that the proposed ICS-HDR has a competitive performance in ORL face recognition and robot navigation over HDR

2. Methodology

2.1 Independent component feature of natural images

There is no uniform definition for natural image, which is put forward by scientists in order to study the optic neural response to the external environment observed by humans. As for the source of information, natural images are information observed by human optic nerve system or natural environmental information collected by machines. It is evident that the processing objects of machine vision are natural images.

Before explaining the superiority of extracting independent features of natural images, we would better discuss on the necessity of analysis on high-order statistical features of natural images. It is necessary because different images may have the same second-order statistical features. For instance, the natural image in Figure 1 was conducted Fourier transform to get Fourier coefficients, and then we changed their order at random and conducted Fourier counter-transform, generating Figure 2. Through this reconstruction process, we know that though the Fourier coefficients of the two pictures are in different orders, their second-order statistical features are the same as their values remain unchanged and so does their energy spectrum.

Fig.1. Original image

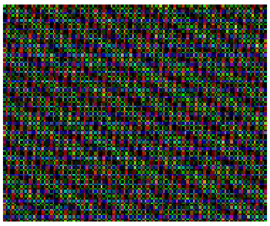

Fig.2. Counter-transform result of randomly ordered Fourier coefficients

Independent components can give good expression to the high-order statistical features of data, whose computation process is as follows:

Step 1**:** center unification. For a given training sample $\boldsymbol{X}$, its average value $\boldsymbol{m} = \mathrm{E}(\boldsymbol{X})$, so $\boldsymbol{X} = \boldsymbol{X} - \mathrm{E}(\boldsymbol{X})$. Let each sample $\boldsymbol{X}$ and required independent component $\boldsymbol{S}$ be zero-mean vectors.

Step 2: conduct whitening to vector $\boldsymbol{X}$ following the equation: $\boldsymbol{V} = \boldsymbol{D}^{-1/2}\boldsymbol{EX}$, and let $\mathrm{E}(\boldsymbol{XX}^{\mathrm{T}}) = \boldsymbol{I}$. Among them; $\boldsymbol{D}$ and $\boldsymbol{E}$ are respectively the eigenvalue matrix and eigenvector matrix of $\boldsymbol{X}$'s covariance matrix

$$\boldsymbol{R}_V = \mathrm{E}((\boldsymbol{X} - \mathrm{E}(\boldsymbol{X}))(\boldsymbol{X} - \mathrm{E}(\boldsymbol{X}))^T) \quad (1).$$

Step 3: determine the number m of the estimated independent components. In order to reduce the human interference in machine, we can use the number of label values and explain it in the classification tree later. Let $p \leftarrow 1$.

Step 4: choose initial vector $\boldsymbol{w}_p$ with unit norm.

Step 5: renew $\boldsymbol{w}_p$ by flowing:

$$\boldsymbol{w}_p \leftarrow \mathrm{E}\{\boldsymbol{V}\,\mathrm{g}(\boldsymbol{w}_p^T\boldsymbol{V})\} - \mathrm{E}\{\mathrm{g}'(\boldsymbol{w}_p^T\boldsymbol{V})\}\boldsymbol{w}_p; \quad (2)$$

among them, g is a non-linear function $\mathrm{g}(y) = \tanh(ay)$, the constant $1 \le a \le 2$ and usually $a = 1$.

Step 6: orthogonalization:

$$\boldsymbol{w}_p \leftarrow \boldsymbol{w}_p - \sum_{j=1}^{p-1}(\boldsymbol{w}_p^T\boldsymbol{w}_j)\boldsymbol{w}_j$$

Step 7: standardization of $\boldsymbol{w}_p$: $\boldsymbol{w}_p \leftarrow \boldsymbol{w}_p / \|\boldsymbol{w}_p\|$

Step 8: if $\boldsymbol{w}_p$ does not converge, go back to Step 5

Step 9**:** $p \leftarrow p + 1$, if $p \le m$, go back to Step 4.

Step 10: obtain the independent component $\boldsymbol{S}$ following the equation: $\boldsymbol{S} = \boldsymbol{w}_p\boldsymbol{X}$

Now, $\boldsymbol{S}$ is the independent space of the sample $\boldsymbol{X}$. Later in the classification regression tree, we only need to choose its subspaces to present the features of the original samples instead of using all the independent components to extract image features. This will be validated in the experiment later.

2.2 Hierarchical Discriminant Regression Tree based on Independent Component Subspace

Suppose the sample aggregate is $L = \{(\boldsymbol{x}_i, \boldsymbol{y}_i)\}, i = 1,2,\cdots n$; $\boldsymbol{x}_i$ is the vector corresponding to the natural information of training samples, and $\boldsymbol{y}_i$ is the vector marking the category features of $\boldsymbol{x}_i$. The presented process for natural image recognition is as follows:

Step 1: y value based clustering. Determine the leaf node based on the maximum distance between $\boldsymbol{y}_i$ under the present knot $d_{y\max}$: if $d_{y\max} > \delta$, it is not a leaf node, requiring further division; otherwise, it is a leaf node and needs no further division. Usually, let $\delta = 0.01$ and group the sample aggregate L to k categories: $L = \{l_1, l_2, \cdots l_k\}$.

Step 2: sample dimensionality reduction. Compute the independent subspace of each category of $l_i, i = 1,2,\cdots k$ following the above independent component analysis procedures(2.1), and then work out k-1-dimensionality subspace supported by independent components and get $\boldsymbol{x}' \in R^{k-1}$ after dimensionality reduction, and calculate m'_{ix}, the center of each category after dimensionality reduction.

Step 3: subspace clustering. Re-conduct clustering to the sample x' after dimensionality reduction according to the distribution of the present category center m'_{ix}. That is to allocate the sample x' to category i^*, which has the minimum discriminate distance, when

$$i^* = \arg\min(d_{SDNLL}(x', i)),\ i = 1, 2, \cdots k$$

and

$$d_{SDNLL}(x', i) = \frac{1}{2}(x' - m'_{ix})^T o_i^{-1}(x' - m'_{ix}) + \frac{k-1}{2}\ln(2\pi) + \frac{1}{2}\ln(|o_i|) \ .$$

Among them, $o_i = w_e \rho_i^2 I + w_m s_m + w_g \Gamma_i$, $\rho_i = \|x' - m'_{ix}\|$, s_m is the covariance matrix of all the samples after dimensionality reduction, and Γ_i is the covariance of the samples in i Category after dimensionality reduction. w_e, w_m, w_g are weighted coefficients and meet the condition: $w_e + w_m + w_g = 1$. It is clear that two times of clustering are aiming at different objects; this process can be briefly shown with Figure 3.

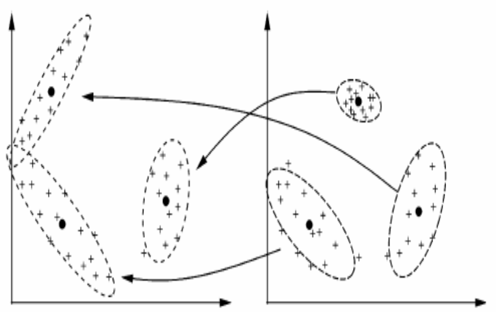

Fig. 3 Y-clusters in space y and the corresponding X-clusters in space X

Step 4: sample distribution. Based on the new clustering of samples in independent subspaces, original samples (x_i, y_i) will be allocated to the child node corresponding to its category. If a node has the same y value, it is a leaf node; otherwise, it should be further split following Steps 1 to 3 until the leaf node is got. As such, we get the classification regression tree as in Figure 4.

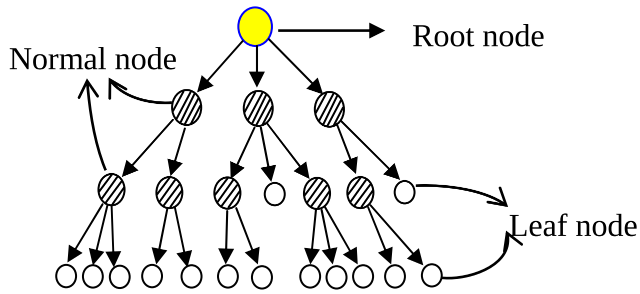

Fig. 4 The structure of ICS-HDR tree

Step 5: testing. The algorithm of trial recognition conducted to samples by independent component hierarchical regression tree is similar to the searching process of ordinary classification trees. That is starting from the root node and locating n childe nodes closest to the testing sample according to their SDNLL distance. These n childe nodes will be further searched until the closest leaf node is found. Then all the samples in these n childe nodes will be searched for the one that has the shortest European-style distance to the testing sample. The y value of this testing sample is considered as the result of recognition.

Characters of this method:

1) On account of the non-Gaussian distribution property of natural images, it employs independent components which can reflect high-order statistical properties to extract features, whose eigenvectors have statistical independence.

2) During classification and recognition process, it adopts supervised learning and unsupervised

clustering, enhancing the difference between categories and correlation within categories.

3) By reducing the dimensionality of samples and employing tree-mode classification structure, it improves the speed of searching.

3. Experiment and Result

Two ORL-face recognition experiments are conducted with this method and their recognition rate and time consumption are compared with those of the HDR method proposed by Hwang, W. et al. [26].

3.1 Experiment on identity recognition to Britain ORL Face Database

This database contains 10 images for each of 40 distinct subjects, taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). Figure 5 is a sample of one person, 112×92 pixel each picture. To facilitate the calculation by computer, we split the images into segments of 8×8 pixel. In this method, the extracted features have 40 dimensions. Can these 40 independent components reflect the original images? Figure 6 is the reconstructed image of the 40-dimensionality independent components of Figure 5, which reflects the features of the original picture. Then in order to display the applicability of this method, we chose different numbers of training and testing samples to make recognition experiment to ORL Face Database and compared their recognition effects and time consumption. The result shows in Table 1.

Fig.5. Original image

Fig.6.Reconstructed image of 40-dimension independent components

Table 1 ID recognition results of ORL-face database

Training samples vs. testing samples	Error rate		Training-time(sec)		Testing-time(msec)	
	HDR	ICS-HDR	HDR	ICS-HDR	HDR	ICS-HDR
320 vs.80	0/80	0/80	2.843	3.048	1.851	1.901
280 vs.120	1/120	0/120	2.625	2.712	1.912	1.992
200 vs.200	3/200	1/200	2.137	2.201	2.030	2.003
160 vs.240	15/240	11/240	2.012	2.032	2.092	2.025

Furthermore, we analyzed the average recognition rate with different component number to the above training and testing samples, shown in Figure 7. This experiment demonstrated that the recognition rate of independent component based HDR trees is higher than that of traditional HDR trees overall, showing that natural image ICA produce more sample information during classification and recognition. Independent component analysis of natural images contains more information than PCA. As for time consumption, these two methods didn't show notable difference, because the whitening process in independent component algorithm takes some time.

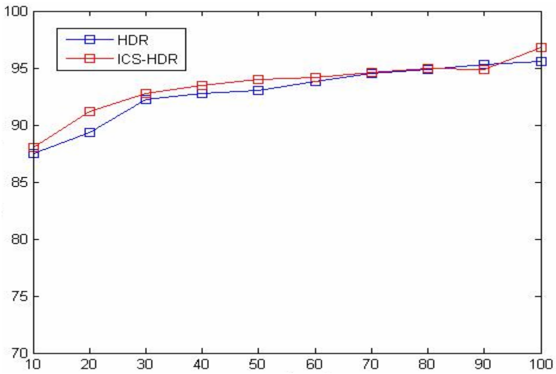

Fig. 7 Comparison of recognition rate between ICS-HDR and HDR with different component number

3.2 Experiment on direction determination of intelligent robot visual navigation

There are three choices in the autonomous navigation of intelligent robot, namely left turn, right turn, and stop. As we designed: when it encounters an obstacle on the left-front, right-front and front, we will train it to turn right, turn left and stop respectively. Figure 8 demonstrates the obstacle locations. The experiment platform we used is the Pioneer3-AT outdoor mobile experiment platform made by US-based Mobile Robots, as shown in Figure 9. We have taken 90 pictures for three situations, 30 for each. Some samples are used to train the robots' behaviors and the others for testing. We have made 5 experiments to different numbers of training and testing samples, whose simulated results are shown in Table 2.

Fig.8: Images of three different obstacle locations (left-front, right-front and front respectively)

Fig. 9 Intelligent robot in the experiment

Table 2 Result of mobile robot visual navigation

Training samples vs. testing samples	Error rate		Training-time(sec)		Testing-time(msec)	
	HDR	ICS-HDR	HDR	ICS-HDR	HDR	ICS-HDR
81 vs.9	1.5/9	1.3/9	1.601	2.009	1.051	1.154
72 vs.18	3.2/18	2.9/18	1.525	1.812	1.300	1.392
60 vs.30	7.9/30	7/30	1.420	1.461	1.751	1.703
45 vs.45	12.3/45	12/45	1.312	1.332	1.801	1.925

We can see from this table 2 that hierarchical regression trees based on independent components can also produce good result in machine visual direction determination. However, we should also notice that in this experiment recognition rate and time consumption of this method are not as satisfying as those in face recognition experiment. This is because the natural situational images we collected have some noises and large pixels; besides, differences between the natural situations are much greater (which is true to the situations robots will encounter in navigation). At the same time, the samples we used are acquired by artificial partition of a picture. If the partition is conducted autonomously by robots, the recognition effect will be even worse. This is also the difficulty exists in visual information based autonomous navigation of intelligent robots.

4. Conclusion

This paper proposes that independent component based hierarchical regression trees can be used in robot visual learning and recognition. From the two experiments above, we can see that HDR trees based on independent components have better effects than traditional HDR trees. This is because traditional HDR trees only cluster the shadow of original data in a certain discriminate subspace instead of original data or the feature information of original data. This kind of subspace learning only reflects general distribution not restricted by feature information of original data. Therefore, traditional HDR trees can only reduce dimensionalities rather than improve the classification accuracy and recognition rate. The experiments have proved it has better effects than the method in literature [26]. Especially, this method has even better generalization ability in robot visual recognition.

1) This paper adopts independent component analysis to extract image features, showing the high-order statistical features of images and more local information, which are crucial to recognition. As Experiment 1 shows, this method has higher recognition rate than traditional HDR trees.

2) This paper uses independent component analysis to analyze natural images, giving consideration to their non-Gaussian distribution property. As the invariance of independent component subspaces is similar to the quality of complicated visual cells [28], we say this method is inspired by biology.

3) ICS-HDR trees adopt supervised and non-supervised classification method, improving the accuracy of recognition. Meanwhile, as it is conducted to feature subspaces, the speed of classification is increased.

4) As for its functions, the classification regression tree not only sets up a basic data storage form but also reflects its distribution property and the connection between different data types. Moreover, as for its adaptability, training data of any task types can use the tree to construct reflection relationship, as long as it can be expressed by input and output vectors. Therefore, this method has strong generalization ability and adaptability.

5) However, in the newest study of image processing and recognition [29-31], pictures of different pixels, feature dimensionalities, information collection environments and experiment platforms show different effects. As Experiment 2 shows, this method has lower recognition rate in machine visual direction determination than in face recognition; therefore, searching for an efficient and feasible method for machine visual navigation is our task in future.

6. Acknowledgment

It is a project supported by National Basic Research Program of China (Program 973), No: 2007CB311005-01. And I would like to especially thank the ORL Face Database provided by Cambridge University.

References:

[1] Y. Lamdan and H.J. Wolfson. “Geometric Hashing: A General and Efficient Model-Based

ter Vision, 1988, pp.238-249.

[2] K. Ikeuche and T. Kanade. "Automatic Generation of Object Recognition Programs". *Proc. IEEE*, Vol. 76, No. 8, 1998, pp. 1,016-1,035.

[3] D.J. Kriegman and J. Ponce. "On Recognizing and Positioning Curved 3-D Objects from Image Contours". *IEEE Trans. Pattern Analysis and Machine Intelligence,* Vol. 12, No. 12, 1990, pp. 127-137.

[4] Chellapa R. and Kanal L.N.eds. "Two dimensional discrete Gaussian markov random field models for image processing". *Progress in Pattern Recognition*, Vol. 2, 1985, pp.79-112, North Holland.

[5] Sarkar A.,Sharma K.M.S. and Sonak R.V. "A new approach for subset 2-D AR model identification for describing textures". *IEEE Transaction on image processing*, Vol. 6. No. 3, 1997, pp. 407-414.

[6] Kaplan L.M. "Extended fractal analysis for the texture classification and segmentation". *IEEE Transaction on image processing*, Vol. 8, No. 11, 2000, pp.1572-1585.

[7] Pentland A. "Fractal based decomposition of nature scenes". *IEEE Transactions on pattern Analysis and Machine Intelligence*, Vol. 7, 1997, pp.170-179.

[8] M. Turk and A. Pentland. "Eigenfaces for Recognition". *J. Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp.71-86.

[9] H. Murase and S.K. Nayar. "Visual Learning and Recognition of 3-D Objects from Appearance". *Int'l J. Computer Vision*, Vol. 14, No. 1, 1995, pp.5-24.

[10] Mojsilovice A., Popovic M.V. and Rackov D.M. "On the selection of an optimal wavelet basis for texture characterization". *IEEE Transactions Image Processing*, Vol. 9, No. 12. 2000, pp.18-33.

[11] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition*, Second Edition, Publishing House of Electronics Industry. Beijing, China. 2004.

[12] A. Pentland, B. Moghaddam, and T. Starner. "View-Based and Modular Eigenspaces for Face Recognition". *Proc. IEEE Conf. Comp. Vision Pattern Recognition*, 1994, pp.84-91.

[13] Simoncelli E P,Olshausen B A. "Natural image statistics and neural representation". *NeuroScience, Annual Review*, Vol. 24. 2001, pp.1193-1216.

[14] Vinje W E, Gallant J L. "Sparse coding and decorrelation in primary visual cortex during natural vision". *Science*, Vol. 287, 2000, pp.1273-1276.

[15] Olshausen B A. and Field D J. "Emergence of simple-cell receptive field properties by learning a sparse code for natural images". *Nature*, Vol. 381, 1996, pp.607-609.

[16] Hyvärinen A. "Sparse code shrinkage: denoising of nongaussian data by maximum likelihood estimation". *Neural Computation*. Vol.11, No.7, 1999, pp.1379-1768.

[17] Bartlett M S, Lades H M and Sejnowski T J. "Independent component representations for face recognition". *Proc of the SPIE, 3299: conference on Human Vision and Electronics Imaging III*, 1998, pp.528-539.

[18] Yu D, Sattar F. "Robust image watermarking based on independent component analysis".

IEEE International Conference on Acoustics, Speech, and Signal processing, Vol. 4, 2002, pp.13-17.

[19] R. Vigário, J. Särelä, V. Jousmäki, M. Hämälinen, and E. Oja. "Independent component approach to the analysis of EEG and MEG recordings". *IEEE Trans. Biomedical engineering*, Vol. 47, No. 5, 2000, pp.589-593.

[20] Stone J V. "Independent component analysis: an introduction". *Trends in cognitive sciences,* Vol. 6, No. 2, 2002, pp.59-64.

[21] Bartlett M. S., Movellan J. R., Sejnowski T. J. "Face recognition by independent component analysis". *IEEE Transaction on neural networks*, Vol. 13, No. 6, 2002, pp.1450-1464.

[22] S.K. Murthy. "Automatic Construction of Decision Trees from Data: A Multidisciplinary Survey". *Data mining and knowledge discovery*, Vol. 2, No. 4, 1998, pp. 345-389.

[23] D. Swets and J. Weng. "Discriminant Analysis and Eigenspace Partition Tree for Face and Object Recognition from Views". *Proc.Int'l Conf. Automatic Face- and Gesture-Recognition*, 1996, pp.192-197.

[24] D. Swets and J. Weng. "Hierarchical Discriminant Analysis for Image Retrieval". *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 21, No. 5, 1999, pp. 386-401.

[25] Yang CC, Prasher SO, Enright P, et al.. "Application of decision tree technology for image classification using remote sensing data". *Agricultural Systems*, Vol. 76, No. 3, 2003, pp.1101-1117.

[26] Hwang, W., Weng, J. "Hierarchical Discriminant Regression". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, 2000, pp.1277-1293.

[27] J.Y Weng, J. McClelland, A. Pentland, O. Sporns, I. Stocklllan, M. Sur, and E.Thelen. "Artificial intelligence: Autonomous mental development by robots and animals". *Science*, Vol. 291, 2001, pp.599-600.

[28] Hoyer P.O. and Hyvärinen A. "Independent component analysis applied to feature extraction from colour and stereo images". *Network*, Vol. 11, 2000, pp.191-210.

[29] S. Akle, M. E. Algorri and A. Salcedo, Use of Wavelet-basis Functions to Extract Rotation Invariant Features for Automatic Image Recognition, *WSEAS Transactions on Information Science and Application*, Vol.5, No.5, 2008, pp.664-673.

[30] S. Makdee, C. Kimpan and S. Pansang, Invariant Rang Image Multi-pose Human Faces Recognition Using Volume Feature and Fuzzy Neural Network, *WSEAS Transactions on Systems*, Vol.6, No.3, 2007, pp.468-474.

[31] A. Lgbal and M. Yaacob, Advanced Computer Recognition of Aesthetice in the Game of Chess, *WSEAS Transactions on Computer*, Vol.7, No.5, 2008, pp.497-510.

[32] K. Adnan and D. Kamil, Image Compression Using Neural Networks and Haar Wavelet, *WSEAS Transactions on Signal Processing*, Vol.4, No.5, 2008, pp.330-339.