ECONTECHMOD. AN INTERNATIONAL QUARTERLY JOURNAL – 2015, Vol.4, No.3, 73-78

DECISION-MAKING RULE EFFICIENCY ESTIMATION WITH APPLYING SIMILARITY METRICS

M. Anokhin, I. Koryttsev
Kharkov National University of Radio Electronics; anokmik@gmail.com

Received August 12.2015: accepted September 09.2015

Abstract. In article the short description of the most often used methods of classification at pattern recognition is given. The main attention is paid to the methods allowing development of a system for image recognition in a real time scale. The features formation method on the base of two-dimensional spatial spectrums of objects images is offered and application of similarity metrics in a decision-making rule for image classification is described. Experimental data of correct and erroneous recognition probabilities as well as image classification time depending on a number of features and on the identification threshold value are presented and analyzed.

Key words: image, classification, probability, recognition of images, Fourier transform, similarity criterion.

INTRODUCTION

Modern computing systems possess a sufficient performance level necessary for image recognition in real time, that allows to automate more effectively the processes preventing natural and antropogenic accidents, and hence to provide higher indicators of the population safety life and activity level. Also the mobile computing systems possessing a lower level of power consumption are under development, this allows creating compact and partially independent systems for image recognition. At present these developments are some of the most actual directions in the modern world, and new workings out of television systems for image recognition considering existing advantages and lacks of modern achievements in a science and engineering are necessary for a society. The modular structure of a system for image recognition allows to accelerate considerably the process of working out necessary modules and to simplify the process of updating components at occurrence of more productive technological decisions. The structure of a television system for objects recognition in real time assumes presence of modules realizing certain functions. Basic duties of such a system are recording and transmission of a video information stream through the data link, object detection, formation of features and classification of input images on the basis of their similarity to images of classes from a database, as well as displaying the result to inform an operator of the recognition system.

Now there are existing developed and maintained systems which are using various video analysis technologies for objects recognition. Such systems instances are VOCORD Tahion, Hawk-Eye and MOBILEYE. VOCORD Tahion carries out video surveillance outline for population vital activity provision in urban conditions. The Hawk-Eye system has been developed by the research engineers from British company Airplane-RADAR and intended for identification of landing ball location around separating lines. The advanced warning system MOBILEYE is the intellectual system for driver assistance utilized for notification about potential accident situation on a road.

THE ANALYSIS OF LAST RESEARCHES AND PUBLICATIONS

The human sight perception of surrounding space is the unique neurophysiologic process. Processing a light stream is carried out in common by an eye iris and a pupil. In the presence of bright light sources the eye iris extends, and the pupil is accordingly narrowed and on the contrary under opposite conditions. After passage of a pupil the light stream of a certain wavelength influences a retina and forms neuroimpulse influencing a specific part of a brain. The result of the given transformation is that the person is capable to see the objects which are in a surrounding space [1]. However the given process possesses some lacks. The surrounding space includes a number of radiated light streams, distinction of which is difficult to the human sight. Thus, there is a situation when the radiated light streams which have close values of a wave length are identified after transformation as the same color.

The method of objects classification used in the image recognition system is defined substantially by a source of input data. Methods of image recognition using image power indicators are based on selection of local features or on the use of a complete image description [2].

Character local image features are selected after preliminary processing of an image, that leads to reduction of the image description. Then the standard statistical approach is applied to recognition of images. For the first time the given method has been applied to extraction of sixteen face parameters of the person, and for classification the Euclid distance is used [3]. The given realization has been improved by increasing the features space up to thirty five elements that has allowed improvement of the recognition quality [2]. Also Hough transformation [4], Reisfeld operator of symmetry [5], filtering and morphological operations [6] are used for image recognition. At manual features extraction it is possible to reach more quality recognition, but at such approach system productivity worsens considerably at features formation [7]. Another widely known method of local features formation is based on the use of dynamically connected structures [8]. For each sample a graph is forming on the image according to following algorithm. On each image the set of reference points is selected, each of which is a link of a full connected graph and marked with the response of the Gabor filter applied to the area around a reference point. A set of such graphs represents the mesh structure capable to distinguish input images effectively. A serious lack of the given method is that formation of the initial set of graphs is necessary to make manually. Application of the parametrical models using deformable templates allows define automatically reference points, that decreases the lack specified above [9]. For recognition quality improvement instead of the Gabor filter, which is forming features, histograms of orientated gradients may be used [10]. An image texture allows define the qualitative parameters necessary for recognition. Object classification on the image is made by application of the structural technique. For this technique formation of the features space may be done with application of large values filters. Thus each filter is a matrix $N \times N$ containing binary values, and the set of filters defines the alphabet of features of image structure [11]. Definition of statistically dependent links allows to classify objects on the image. Holisitc classification methods are one of the most simple since they are comparing directly corresponding elements of the two-dimensional array of intensity values from input image and image from database. Though this approach has been shown as operating, but it is very sensitive to changes in image orientation or environment parameters [12]. This problem can be solved by utilizing statistical methods for extracting the most significant features thus reducing dimension of features that describe the image. For the first time such approach was implemented in the principal conponent analysis [13]. This method shows that any object image can be effectively represented in the space of its eigen images and recovered as very closed to original on using a reduced image desciription. When image eigenvectors are used for image classification, image features space can be significantly reduced thus eliminating external factors influence on recognition quality [14]. Alternative approach is to use the Fisher's linear discriminant analysis that maximises dispersion correlation between database images and hypothetically provides better image classification than the principal component analysis or method that uses difference images which are defined as difference between corresponding elements of two images [15]. Also in pattern recognition systems methods based on the usage of neural networks or machine learnings techniques has found their application. There is a method that extracts fifty basic features and then auto-associative neural network is used for transforming features into five-dimensional features space. After that standard multi-layer perceptron is used for image classification [16]. Similarly hierarchial neural networks are used which was grown automatically and not trained on gradient descent method [17]. Hybrid neural networks which are utilizing local image sampling, Karhunen-Loeve transformation and multi-layer perceptron has improved images recognition systems performance [18, 19]. Feed forward neural networks are used for classification after features space was generated using principal component analysis [20]. Other approach assumes image decomposition into three components using wavelet transform. The decision rules for every component are fused using radial basis function neural network, for the further classification of input images [21]. Similarly, an input image can be divided into defined amount of regions and for every region a module of the neural network is assigned. The output from all modules is combined using fuzzy Sugeno integral [22]. Also another method was proposed in which similarity function is trained using certain level of confidence that two images belong to the same person. Features space is formed by acquiring subregions local binary pattern histograms and Chi-square distances between corresponding histograms are used as discriminative features. AdaBoost algorithm is applied for the most efficient features extration and similarity function formation [23]. Considerable amount of classes presented in recognition system database causes some problems at classification stage. Possible decision for this problem is decomposing it into a set of binary classification problems in which every classifier is trained for corresponding pair of classes, ignoring all other classes, and then all classifiers are fused into one global classifier [24]. Also support vector machine method, in which all binary classifiers are transformed into one high-dimensional feature space, can be used for solving this problem[25]. It is also supposed that optimum hyperplane exists that can divide various classes and maximize distance between each class and hyperplane that will allow correct recognition of an input image.

THE PURPOSE

Given article is directed on studying of modern methods of image similarity determination in decision-making rules for the subsequent objective choice of a direction for engineering an image recognition system, capable to function in real time. Object achievement is carried out by fulfillment of experimental researches of the developed module for image classification and the analysis of the most important characteristics at image classification.

BASIC RESULTS OF RESEARCHES

The estimation of the developed module efficiency is made for classification of two-dimensional objects by using images from the MUCT database [26]. The module contains two components, realising the feature formation method and the decision-making algorithm.

According to conditions of image recognition system performance the images arriving to an input of the system or being in its database, preliminary pass a processing stage. The square image form is created for this purpose, and the image will be transformed to a black-and-white format as the system will work over the infra-red range.

As formation of the most informative features is the important process for maintenance of input object recognition quality, therefore the developed method makes double transformation of an image data file. Preliminary the two-dimensional discrete Fourier transform is applied for transition of an image data file into the frequency domain [27, 28].

The Fourier transform allows to form more exact image representation, as the Fourier factors are more exact approximation of an image. This results from the fact that the initial and transformed images are connected only with complex constants, describing changes in an amplitude scale and a phase shift of initial image components. Also presence of a bigger number of algorithms for fast calculation of the discrete Fourier transform allows to process images more effectively unlike other methods.

For forming more informative features, the secondary orthogonal transformation is applied, allowing to allocate eigen vectors of a transformed image data file. It will allow to limit the image feature space, that will simplify requirements to a memory configuration in the image recognition system.

The classification module component realising decision-making algorithm is constantly in a waiting mode and after reception of an input data of objects makes definition of their similarity to classes from a database. The developed algorithm of the decision-making rule for realisation of the described problem uses the metrics of similarity presented by the valid function defining similarity of two objects.

There is a theoretical assumption that a similarity value is in inverse proportion to a distance between two objects. This may be explained by the fact that each object being described with a feature set in a corresponding multidimensional space and the objects located at rather short distance, are perceived as having greater similarity, and accordingly the converse is logical [29]. Such metrics allow to calculate image similarity more effectively, as the calculation needs for insignificant computing capacities.

The decision-making rule algorithm provides initial definition of input object image similarity to all images of classes from a database. Therefore an average similarity value of the input object image to each image of a class from a database is calculated for the number of feature vectors defined by the classification module configuration under the formula:

$$S_{mean} = \frac{1}{N}\sum_{i=1}^{N} S(a_i, b_i), \quad (1)$$

where: a_i, b_i - accordingly vectors of an input object image and a class from a database, N - quantity of eigen vectors.

After forming the list of similarity the maximum similarity value is allocated. This value and a class are added in the similar classes list, provided that the similarity value exceeds an identification threshold, defined by a configuration of the classification module.

Then the maximum similarity value of the similar classes list is repeatedly defined, and the information about the corresponding class of a database is deduced on the display for the further decision-making by the operator of the computing system. In the absence of similar classes the corresponding notice is also deduced.

The similarity definition between input object image and a class from a database has been made during experimental researches with the help of similarity metrics [30]. Each object may be described with a feature set in the corresponding multidimensional space.

Hence, objects located close one to another, are perceived as having the bigger similarity, and on the contrary. Similarity between two images can be estimated on using the Dice similarity criterion:

$$D_{Dice} = \frac{2 \times \sum_{i=1}^{N} A_i B_i}{\sum_{i=1}^{N} A_i^2 + \sum_{i=1}^{N} B_i^2}, \quad (2)$$

where: A_i, B_i are data files of the input object image and image of a class from a database, and N – characterises dimension of the image.

The similarity metrics, which is close to the Dice similarity criterion, is called as the Cosine metrics and is defined by the expression:

$$D_{Cos} = \frac{\sum_{i=1}^{N} A_i B_i}{\sqrt{\sum_{i=1}^{N} A_i^2}\sqrt{\sum_{i=1}^{N} B_i^2}}. \quad (3)$$

The Tanimoto factor is also the similarity metrics:

$$D_{Tan} = \frac{\sum_{i=1}^{N} A_i B_i}{\sum_{i=1}^{N} {A_i}^2 + \sum_{i=1}^{N} {B_i}^2 - \sum_{i=1}^{N} A_i B_i}. \qquad (4)$$

Consideration of the several similarity metrics allows to define the recognition system, which performance is more rezultive and provides the best quality. During modelling it has been found out that the results received at the use of the Cosine metrics and Dice similarity criterion are identical, therefore the combined graphs will be resulted further. Preliminary experimental researches were carried out with representation of the input object image by a singl image. It has allowed to estimate the image classification module performance quality in the stably formed space.

Fig. 1 and 2 shows the correct recognition probability versus the number of eigen vectors, with different identification thresholds used in a decision-making rule.

In fig. 3 and 4 the graphs of recognition error probability are given as functions of the same parameters.

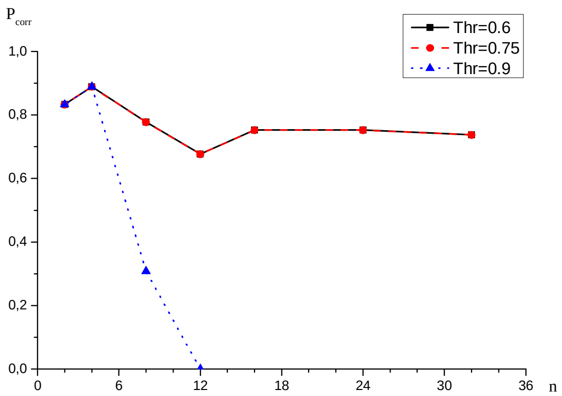

Fig. 1. Recognition correct probability for the Cosine metrics and the Dice similarity criterion

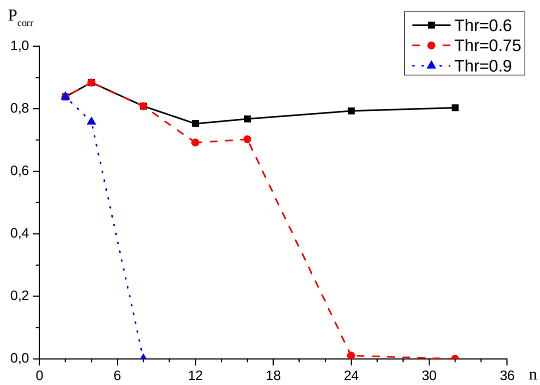

Fig. 2. Recognition correct probability for the similarity Tanimoto metrics

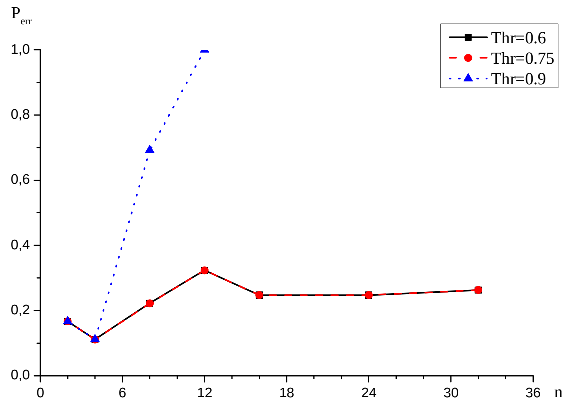

Fig. 3. Recognition error probability for the Cosine metrics and the Dice similarity criterion

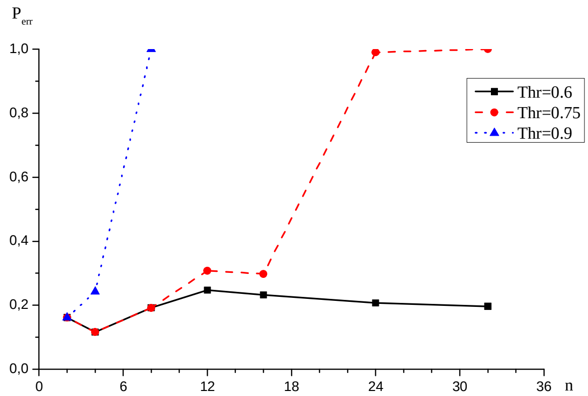

Fig. 4. Recognition error probability for the Tanimoto similarity metrics

Figures 5, 6 and 7 represent the time spent for recognition of one object with application of different similarity metrics.

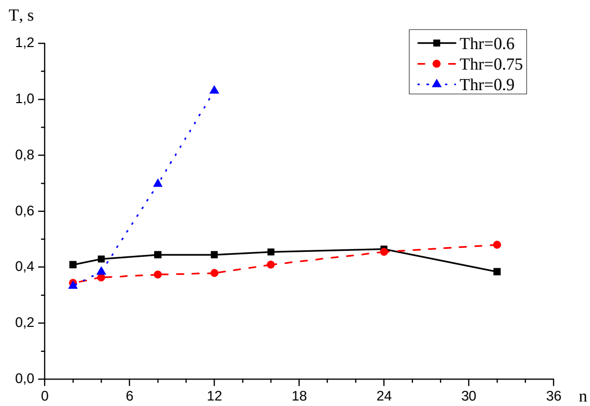

Fig. 5. Recognition time for the Cosine similarity metrics

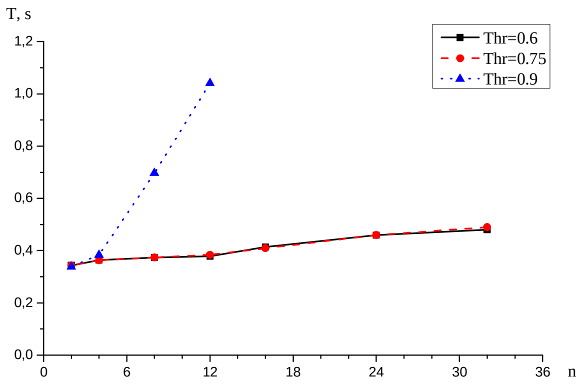

Fig. 6. Recognition time for the Dice similarity criterion

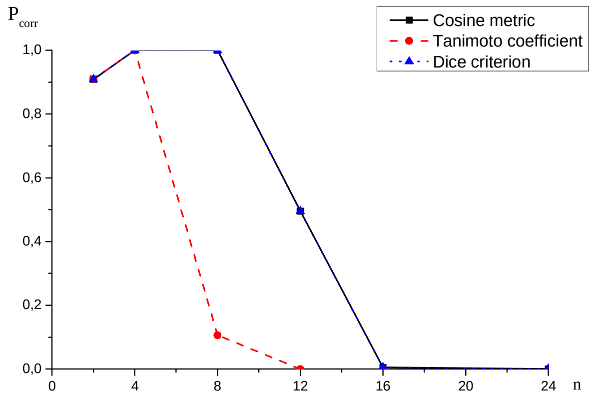

Fig. 8. Recognition correct probability for a set of video frames

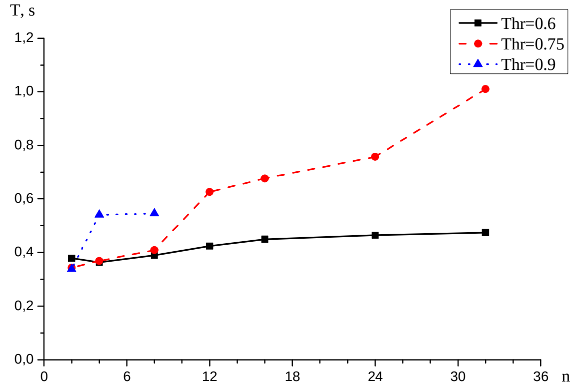

Fig. 7. Recognition time for the similarity Tanimoto metrics

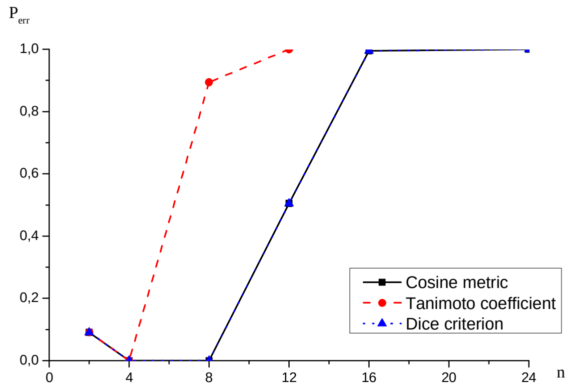

Fig. 9. Recognition error probability for a set of video frames

The analysis of the experimental graphs shows that at an increase in the identification threshold value the number of correctly distinguished objects starts to decrease. This will allow to reduce the time for recognition of input objects, as the database classes with similarity below an identification threshold will be excluded from the list of similar classes.

The minimal values of error probability coinside with a small number of feature vectors - from 2 to 8. This will allow to short the feature space and reduce the recognition time as well.

Current expermental results of the recognition system prototype efficiency allow to increase the productivity of the developed classification module considerably and start further researches.

The image classification module has been also approved at modelling the system performance in real time. For this purpose the individual image of an input object has been transformed to display its moving in a video stream.

In fig. 8 and 9 graphs of correct and error probabilities of input object recognition received for a set of video frames are shown.

Experimental researches show that the use of video frames allows to raise considerably the recognition quality.

CONCLUSIONS

1. Comparison of various similarity metrics shows that the cosine metrics and the Dice similarity criterion are less exposed to erroneous recognition. Also the cosine metrics and the Dice similarity criterion, unlike the Tanimoto similarity metrics, just slightly influence productivity of the developed method. Hence in the decision-making rule it is necessary to give preference to the cosine metrics or the Dice similarity criterion.

2. The important fact is that the correct recognition of input images is possible if the features space will be reduced to increase productivity of the image recognition system. This confirms that the most informative features are in the first eigen vectors of the image. Thus a considerable number of following nonsignificant eigen vectors is perceived by the recognition system as the raised noise level increasing a recognition error value.

REFERENCES

1. Semenets V., Natalukha Yu., Taranukha O., Tokarev V. 2014. About One Method of Mathematical Modelling of Human Vision Functions. ECONTECHMOD. An international quarterly journal Vol. 3, №3, 51-59.

2. Jafri R., Arabnia H. 2009. A Survey of Face Recognition Techniques. Journal of Information Processing Systems. 5 (2), 41-68.

3. Kanade T. 1973. Picture Processing System by Computer Complex and Recognition of Human Faces. Kyoto University, Japan, PhD Thesis.

4. Nixon M. 1985. Eye spacing measurement for facial recognition. SPIE Proceedings. 279-285.

5. Reisfeld D. 1994. Generalized symmetry transforms: attentional mechanisms and face recognition. Tel-Aviv University. PhD Thesis.

6. Graf H., Chen T., Petajan, E., Cosatto E. 1995. Locating faces and facial parts. International Workshop on Automatic FACE - and Gesture-Recognition. 41-46.

7. Cox I., Ghosn J., Yianilos P. 1996. Featurebased face recognition using mixture-distance. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 209-216.

8. Lades M., Vorbrüggen J., Buhmann J., LangeJ., Malsburg C., Würtz R., Konen W. 1993. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Computers. 42, 300-311.

9. Campadelli P., Lanzarotti R. 2005. A Face Recognition System Based on Local Feature Characterization. Advanced Studies in Biometrics. 3161, 147-152.

10. Albiol A., Monzo D., Martin A., Sastre J. 2008. Face recognition using HOG–EBGM. Pattern Recognition Letters. 29, 1537-1543.

11. Brytik V.I., Zhilina O.Yu., Kobziev V.G. 2014. Structural Method of Describing The Texture Images. ECONTECHMOD. An international quarterly journal Vol. 3, №3, 89-98.

12. Baron R. 1981. Mechanisms of Human Facial Recognition. International Journal of Man-Machine Studies. 15, 137-178.

13. Sirovich L., Kirby M. 1987. Low-dimensional Procedure for the Characterization of Human Faces. Journal of the Optical Society of America A: Optics, Image Science and Vision. 4, 519-524.

14. Turk M., Pentland A. 1991. Face Recognition Using Eigenfaces. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 586-591.

15. Moghaddam B., Nastar C., Pentland A. 1996. A Bayesian Similarity Measure for Direct Image Matching. Proceedings 13th International Conference on Pattern Recognition. 350-358.

16. DeMers D., Cottrell G. 1993. Non-linear dimensionality reduction. Advances in Neural Information Processing Systems. 5, 580-587.

17. Weng J., Ahuja N., Huang T. 1993. Learning recognition and segmentation of 3-D objects from 3-D images. Proceedings of the International Conference on Computer Vision (ICCV 93). 121-128.

18. Lawrence S., Giles C., Tsoi A., Back A. 1997. Face Recognition: A Convolutional Neural Network Approach. IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition. 1-24.

19. Turk M., Pentland A. 1991. Eigenfaces For Recognition. Journal Of Cognitive Neuroscience. 3, 71-86.

20. Eleyan A., Demirel H. 2005. Face Recognition System Based on PCA and Feedforward Neural Networks. Computational Intelligence and Bioinspired Systems. 3512, 935-942.

21. Li B., Yin H. 2005. Face Recognition Using RBF Neural Networks and Wavelet Transform. Advances in Neural Networks. 3497, 105-111.

22. Melin P., Felix C., Castillo O. 2005. Face recognition using modular neural networks and the fuzzy Sugeno integral for response integration. International Journal Of Intelligent Systems. 20, 275-291.

23. Zhang G., X., Li S., Wang Y., X. 2004. Boosting local binary pattern (LBP)-based face recognition. Advances In Biometric Person Authentication, Proceedings. 3338, 179-186.

24. Krebel U. 1999. Pairwise classification and support vector machines. Advance in Kernel Methods – Support Vector Learning. 255-268.

25. Burges C. 1998. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery. 2, 121-267.

26. Milborrow S., Morkel J., Nicolls F. 2010. The MUCT Landmarked Face Database. Pattern Recognition Association of South Africa.

27. Brigham E. 1988. Fast Fourier Transform and Its Applications. Prentice Hall. 448.

28. Gonzales R., Woods R. 2007. Digital Image Processing. Prentice Hall, 976.

29. Yaegashi Y., Tateoka K., Fujimoto K., Nakazawa T., Nakata A., Saito Y., Abe T., Yano M, Sakata K. 2012. Assessment of Similarity Measures for Accurate Deformable Image Registration. Journal of Nuclear medicine and Raidation Therapy. 3(4).

30. Sung-Huyk C. 2007. Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions. International Journal of Mathematical Models and Methods in Applied Science. 1(4), 300-307.