Local Accuracy Measurement for Face Recognition System using Numerous Classifiers (PCA, GA and ANN)

Md. Shahjahan Kabir
Lecturer, Dept. of CSE,
Dhaka Int. University(DIU)
Dhaka, Bangladesh

Kh. Mohaimenul Kabir
Dept. of CSE
Dhaka Int. University (DIU)
Dhaka, Bangladesh

Rabiul Islam, Ph.D.
Associate Professor,
Dept. of CSE,RUET
Rajshahi, Bangladesh

ABSTRACT
Between the various biometric methods, Face Recognition has become one of the most burning topic tasks in the pattern recognition field during the past decades. In This Work a Face Recognition System has been developed By applying different multiple classifier selection schemes on the output of three different classification methods namely Artificial Neural Network, Genetic Algorithm And Euclidean distance measure based on Principal Component Analysis dimensionality reduction technique. Dynamic classifier selection technique using classifier of local accuracy measurement is to look up the individual correctness of individual classifier and select the best one from them. Here it is proposed a classifier local accuracy measurement technique which is used to dynamic classifier selection algorithm. From the result and performance analysis it can be said that numerous classifier selection schemes give better performance than single classifier and dynamic classifier selection algorithm using proposed classifier local accuracy measurement technique gives stable high and better performance.

General Terms
Pattern Recognition, Multiple Classifier, Protection.

Keywords
Feature Selection, Face Recognition, Principal Component Analysis (PCA), Genetic Algorithm (GA), Artificial Neural Network (ANN), AT&T Cambridge face database.

1. INTRODUCTION
Face Recognition can be used in law enforcement agencies for security purpose, check cashing security, criminal investigations. It can be used for authentication and identification. [1]

It is now proved that combination of multiple classifiers rather than single classifier gives better performance. There are two basic approaches a Combination of Multiple Classifiers (CMC) algorithm may take: classifier fusion and dynamic classifier selection. The system is tested on AT&T Cambridge Face Database in which the proposed system is robust against illumination and expression as well as increasing the recognition performance. In classifier fusion algorithms, individual classifiers are used in parallel and their outputs are compounded in some manner to attain a "group consensus." Dynamic Classifier selection attempts to predict which single classifier is most likely to be correct for a given sample. Only the output of the selected classifier is considered in the final decision. [2]

Theoretical and experimental results reported in the literature have clearly shown that classifier fusion is effective if the individual classifiers are "accurate" and "diverse", that is, if they exhibit low error rates (at least lower than 50%) and if they make different errors , it has been shown that the combination.

Unfortunately, the reported experimental and theoretical results have pointed out that the creation of accurate and diverse classifiers is a very difficult task. In real applications, the most likely situation is to have reasonably accurate but "positively" dependent classifiers i.e., classifiers that make many identical errors. [3]

On the other hand, it can be verified experimentally that it is easier to design a classifier ensemble, where on considering each pattern; at least one classifier can classify it correctly, while the remaining classifiers could make the same error. Accordingly, the authors and other researchers have proposed an alternative approach to classifier combination, based on the concept of "dynamic classifier selection" (DCS). DCS is based on the definition of a "function" that for each pattern selects the classifier that is more likely to classify it correctly. [2]

This paper deals with the performance analysis of multiple classifier selection schemes using face recognition system.

2. MULTIPLE CLASSIFIER SELECTION

2.1 Basic Concept
Multiple classifier combination can be explained briefly as: to derive the final classification decision by integrating the output of multiple learning machines according to certain Combination approach. In pattern recognition and classification, the algorithm that is effective for one feature set may be unsuitable to other feature sets and numerous classifiers can provide the complementary information about the classified pattern on hand, so numerous classifier combination may outperform any individual classifier by integrating the advantages of various classifiers. [7]

A classifier divide the feature space according to its classification scheme that means each classifier classifies or divide the feature space into mutually exclusive subspaces or class. When an input pattern comes, it assigns that pattern into a suitable subspace whether the input is totally correct or partially correct to that subspace. In multiple classifier selection scheme the subspaces divided by the single classifier is combined in a way that it reflect the Bayesian optimal subspaces.

2.2 Optimality of Numerous Classifier Selection System
Without losing generality, each decision region R_i^j can be considered subdivided into the regions $R_{i+}^j = R_i^j \cap R_i^B$ and $R_{i-}^j = R_i^j - R_{i+}^j$. Accordingly, $R_i^j = R_{i+}^j \cup R_{i-}^j$. The decisions made by each classifier C_j are equal to those of the optimal

Bayes classifier withinR_{i+}^{j}. Non-optimal decisions are made within R_{i-}^{j} and those are shown in Fig 1 and Fig 2. [2]

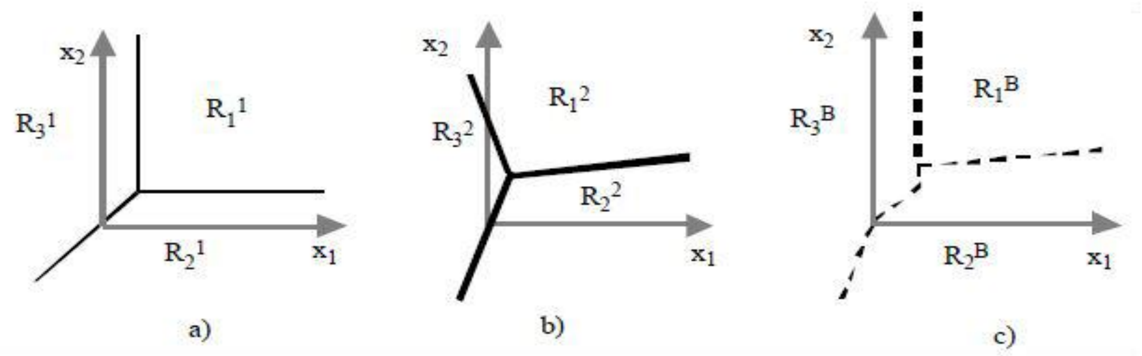
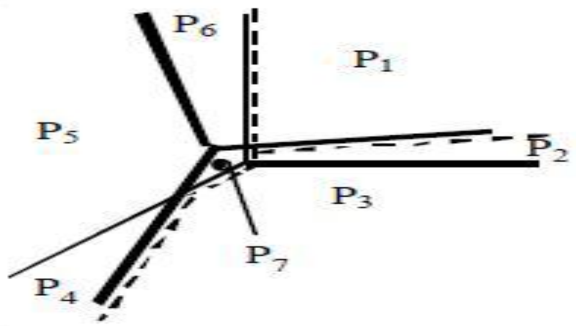


Fig 1: Example of a two-dimensional classification task with three data classes: a) boundaries of the decision regions of classifier C_1; b) boundaries of the decision regions of classifier C_2; c) boundaries of the optimal Bayes decision regions.



$P_1 = R_{1+}^{1},\ P_1 = R_{1+}^{2}$
$P_2 \subset R_{2+}^{2}$
$P_3 = R_{2+}^{1},\ P_3 \subset R_{2+}^{2}$
$P_4 \subset R_{3+}^{2}$
$P_5 \subset R_{3+}^{1},\ P_5 \subset R_{3+}^{2}$
$P_6 \subset R_{3+}^{1}$
$P_7 \subset R_{3+}^{1}$

Fig 2: Feature-space partitioning generated by two classifiers in Figures 1. a and 1.b.

From the Fig 1 and Fig 2 it can be said that every classifier divides the feature space in sub region, but they are always not fulfilling the Bayesian optimal decision boundary or region. Each classifier gives optimal solution for some region. But if combined the classifier, the Bayesian optimal region can be achieved as the definition of optimal region given above and the figure shown above.

3. IMPLEMENTATION

A Facial Recognition system is a computer application which automatically justifies or ascertains a person from a digital image or a video frame from a video source. Compare from selected facial features from the image and a facial database is one of those to do this.

3.1 Feature Extraction

An image can be represented as a two dimensional matrix which can be represent as a set of different unique ghost image called principal component or Eigen face. This process of finding principal component is referred as feature extraction using PCA. The process of feature extraction is as follows:

Step 1: Get some data, Let X={x_1,x_2……..x_n}be the set of data

Step 2: Subtract the mean. Where mean is arithmetic mean defined as $\boldsymbol{\bar{x}} = \frac{\sum_{i=1}^{n} x_i}{n}$

Step 3: Calculate the covariance matrix

Step 4: Calculate the eigenvectors and eigenvalues of the covariance matrix

Step 5: Deriving the new data set by the formula

Final data = Row Feature Vector x Row Data Adjust

3.2 Recognition

Recognition is the process of identifying data from train database for an input data. In pattern recognition system it is the process of identifying pattern according to input pattern. In this thesis experiment three classification techniques are used namely Artificial Neural Network, Genetic Algorithm and Euclidean distance measure technique using PCA. In ANN, BPN learning algorithm is used. Three classifiers give individual output whether they recognize same pattern or different pattern.

3.3 Dynamic Classifier Selection

This paper deals with the performance analysis of multiple classifier selection system. Dynamic classifier selection scheme one of the most effective multiple classifier selection scheme. The dynamic classifier algorithm works on the classifier local accuracy measurement that how a classifier is correct by itself. Two methods of CLA measurement techniques are discussed here. One is CLA- a priory and other is proposed CLA measurement technique.

3.3.1 Classifier Local Accuracy- A priory

Step1: Select pattern from validation database using K-nearest neighbor for the input pattern.

Step2: Identify the weight using Euclidean distance measure of each pattern selected from validation data from input pattern for particular classifier.

Step3: Each pattern selected from validation database identify pattern from test database, the probability of correctness using correlation of each identification is found.

Step4: Measure the classifier local accuracy of each classifier using the formula.

$$CLA_j(X^*) = \frac{\sum_{n-1}^{N} P_j(\omega_i | X_n \in \omega_i).W_n}{\sum_{n-1}^{N} W_n} \quad j = 1, \dots \dots, K \; i = 1, \dots \dots, M \quad \textbf{(Eq. 1)}$$

Where, $P_j(\omega_i | X_n \in \omega_i)$ is the probability of correctly identifying pattern by a classifier.

$W_n = \frac{1}{d_n}$, where d_n is the Euclidean distance between the validation data and input data

K is the total number of classifier and M is the total number of data in train database.

3.3.2 Classifier Local Accuracy using Probability Estimation

Step1: Select pattern from validation database using K-nearest neighbor for the input pattern.

Step2: Measure the distance from input pattern to output pattern of each classifier.

Step3: Find the probability of correctness of selected pattern from test database for the pattern selected from validation database.

Step4: Measure the Classifier Local Accuracy of each classifier using the formula

$$CLA_j(X^*) = P_j * W_j \quad j = 1, \dots \dots, K \quad \textbf{(Eq. 2)}$$

Where, P_j is the probability of correctly identifying pattern by classifier j.

$W_j = \frac{1}{d_j}$, where d_j is the Euclidean distance between the input image and the output image by classifier j and K is the total number of classifier.

3.3.3 Dynamic Classifier selection algorithm

In this section a dynamic classifier selection algorithm is discussed according CLA measurement discussed above.

Input Parameters: Test pattern X^* classification labels of validation data, size of neighborhood, rejection threshold value, and selection threshold value.

Output: Classification of test pattern X^*

Step1: If all the classifiers assign to the same data class, then the pattern is assigned to this class.

Step2: Compute $CLA_j(X^*), j=1,\dots\dots\dots,k$

Step3: If $CLA_j(X^*)$ <rejection-threshold then disregard classifier C_j.

Step4: Identify the classifier exhibiting the maximum value of $CLA_j(X^*)$.

Step5: For each Classifier compute the following differences

$$d_j = [CLA_m(X^*) - CLA_j(X^*)]$$

Step6: If for every j, j≠m, d_j > selection-threshold then select classifier C_m else select randomly one of the classifier for which d_j < selection-threshold. [2]

4. EXPERIMENTAL RESULT

The used database consists of different image of 10 peoples. Here AT&T Cambridge face database is used for this experiment**.** The experiment has two train set consist of 20 and 30 images having each person's 2 and 3 images respectively and two test set consist of 40 and 50 images having each person 4 and 5 images respectively. The performance of each classifier and MCS are shown in percentage given below.

Table 1: Recognition Rate Analysis of Classifier

Train---Test images	Euclidean distance measure using PCA	ANN	GA
20-50	88	96	90
20-40	93.5	91.5	89
30-50	95	91	96
30-40	92	95	94.5

Table 2: Recognition Rate Analysis of Numerous Classifiers

Train---Test images	Voting method	DCS using CLA- a priori	DCS using CLA-proposed
20-50	91	92	96
20-40	94	95	94.5
30-50	98	98	96
30-40	96	95.5	95.5

From the Table 1 and Table 2 it can be conclude that the performance of MCS is better than single classifier and among the MCS systems the proposed system gives stable high performance .The performance is clearly shown in Fig 3.

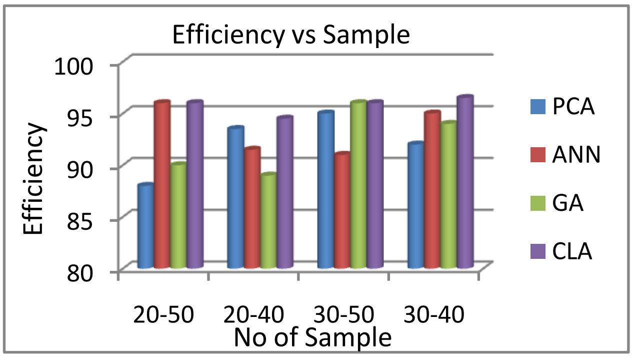


Fig 3: Performance analysis chart between single classifier and proposed multiple classifier

5. CONCLUSION

Applying classifier method the best output is selected from them using different Multiple Classifier Scheme. From the experimental result shown in graph, it is clear that the performance of MCS system is better than single classifier and the proposed system always give stable high performance. If the dimension of resized image is less critical in Principle Component Analysis and Genetic Algorithm takes less time for searching it will be the better output result. In Future it will be overcome.

6. REFERENCES

[1] International Journal of Computer Applications (0975 – 8887),Volume 91 – No.15, April 2014

[2] Giorgio Giacinto, Fabio Roli, "Dynamic Classifier Selection based on Multiple Classifier Behaviour".2001

[3] Giorgio Giacinto, Fabio Roli, "A Theoretical Framework for Dynamic Classifier Selection," icpr, vol. 2, pp.2008, 15th International Conference on Pattern Recognition (ICPR'00) - Volume 2, 2000.

[4] Sentayehu Endeshaw Wolde, Dr. Kumudha Raimond, "Face Recognition using Artificial Neural Network", ESEE and ECE Department of Jimma University, August 17, 2008.

[5] Shubhendu Trivedi, Face Recognition using Eigenfaces and Distance Classifiers: A Tutorial, February 11, 2009.

[6] S. N. Sivanandam , S. N. Deepa, "Introduction to Genetic Algorithms", PublisherSpringer Publishing Company, Incorporated ©2007.

[7] Smits, P.C. , "Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection" Geoscience and Remote Sensing, IEEE Transactions on IEEE, vol.40 pp 801-833 , Apr 2002.

[8] Woods_K. Dept. of Comput. Sci. & Eng., Univ. of South Florida,Tampa,FL Kegelmeyer, W.P. Jr. , Bowyer, K. "Combination of multiple classifiers using local accuracy estimates" Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.19, pp 405-410, Apr 1997.

[9] Nathan J. Leap, Captain, USAF, "An Investigation Of The Effects Of Correlation, Autocorrelation, And Sample Size In Classifier Fusion", Air Force Institute of Technology, Dept. of Operational Science, March 2004.

[10] Jon Shlens," A Tutorial on Principal Component Analysis, Derivation, Discussion and Singular Value Decomposition", jonshelns@ucsd.edu, 25 March 2003.

[11] Philip David Brierley, "Some Practical Applications of neural networks in the electricity industry", engd thesis - 1998, canrnfield university UK.

[12] Lindsay I Smith, "A tutorial on Principal Component Analysis", Feburary 26,2002.

[13] DavidE.Goldberg," Genetic Algorithms in Search, Optimization, and Machine Learning ".

[14] S.RAJASEKARAN and G.A. VIJAYALAKSHMI PAI,"Neural Networks, Fuzzy Logic, and Genetic Algorithms Synthesis and Applications".

[15] Paul Dawkins, "LINEAR ALGEBRA".

[16] Rafael C. Gonzalez and Richard E Woods, "Digital Image Processing".

[17] Matthew A. Turk and Alex P. Pentland, "Eigenfaces for Recognition",Journal of Cognitive Neuroscience '1991.

[18] R. Beale and T Jackson, "Neural Computing: An Introduction".