ENHANCING PERFORMANCE OF FACE RECOGNITION SYSTEM BY USING NEAR SET APPROACH FOR SELECTING FACIAL FEATURES

[1]Shivani Gupta, [2]K.S.Patnaik

[1]PG Student, Department of Computer Science and Engineering, BIT, Ranchi, India-835215.

[2]Prof., Department of Computer Science and Engineering, BIT, Ranchi , India-835215.

E-mails:shivani.cse.bit@gmail.com, ktoshri@rediffmail.com

ABSTRACT

The application of Support Vector Machines (SVMs) in face recognition is investigated in this paper. SVM is a classification algorithm recently developed by V. Vapnik and his team. We illustrate the potential of SVMs on the Cambridge ORL face database, which consists of 400 images of 40 individuals, containing quite a high degree of variability in expression, pose, and facial details. Our face recognition systems consist of two major phases. We present an automated facial feature extraction procedure and make use of Near set approach to choose the best feature among the considered one which significantly improves face recognition efficiency of SVM. Near Set approach was introduced by James Peters in 2006, as a result of generalization of rough set theory. One set *X* is near to another set *Y* to the extent that the description of at least one of the objects in set *X* matches the description of at least one of the objects in *Y.* Also we have shown that for face recognition in ORL face database using SVM with feature selection by near set approach has error rate 0.2% which is very less as compared to error rate obtained in the previous work done by other authors on the ORL face database.

Keywords: Face recognition, support vector machines, optimal separating hyperplane, kernel, feature space, Near Set, average near coverage.

1. INTRODUCTION

Face recognition technology can be used in wide range of applications such as identity authentication, access control, and surveillance. Interests and research activities in face recognition have increased significantly over the past few years [1] [2] [3]. Different techniques are developed in face recognition, such as eigen-faces [4][5], hidden Markov models [6], elastic graph-matching (EGM) by Wiskott et al. [7], morphological shared-weight neural network (MSNN) [8], EGM based on multi-scale erosion-dilation [9], support vector machines [10], and so on.

As an important computer vision task, the problem of automatic face recognition is a difficult and challenging problem. Two issues are central; the first is what features are to use to represent a face. As face image subjects to changes in viewpoint, illumination and expression. An effective representation should be able to deal with possible changes. The second is how to classify a new face image with the chosen representation.

To deal with first issue we have developed an automated procedure to extract facial features. A pre-processing of the images is done to get the region of interest of the image and to extract facial features. It is, however, uneconomical to obtain all the facial features because this is expensive and time consuming. And also, all the facial features are not independent and some of them are strongly correlated. Hence, it is not important to use all the facial features as input parameters. Hence, taking the above discussion and the aim of the investigation under consideration, we have used the facial features i.e. nose length, nose width and distance between eyeballs. We also use the concept of Near Set to choose the feature among these three features which gives the best face recognition accuracy.

Near sets were introduced by James Peters in 2006[11] and formally defined in 2007[12]. Near

Sets result from a generalization of rough set theory. Briefly, one set *X* is near to another set *Y* to the extent that the description of at least one of the objects in *X* matches the description of at least one of the objects in *Y*. The hallmark of near set theory is object description and the classification of objects by means of features [13]. Rough sets were introduced by Zdzisław Pawlak during the early 1980s [11, 14] and provide a basis for perception of objects viewed on the level of classes rather than the level of individual objects. A fundamental basis for near set as well as rough set theory is the approximation of one set by another set considered in the context of approximation spaces. It was observed by Ewa Orłowska in 1982 that approximation spaces serve as a formal counterpart of perception, or observation [15].

For second issue, we use Support Vector Machines (SVMs). Which acts as face recognizer.SVMs have been recently proposed by Vapnik and his co-workers [16] as a very effective method for general purpose pattern recognition. Intuitively, given a set of points belonging to two classes, an SVM finds the hyperplane that separates the largest possible fraction of points of the same class on the same side, while maximizing the distance from either class to the hyperplane. According to Vapnik [16], this hyperplane is called Optimal Separating Hyperplane (OSH) which minimizes the risk of misclassifying not only the examples of the training set but also the unseen examples of the test set. Support vector machine approach is largely characterized by the choice of its kernel. Once the kernel is decided, all what need to be chosen by user is the error penalty parameter i.e. C [17]. Some work relative to this problem can be seen in [18] [19]. Another important part of SVM is the solution method. Most popular methods are the sequential minimal optimization algorithm (SMO) [20] [21], fast nearest point algorithm [22] are developed recently.

After this section this paper is organized as: Section 2: Basic Concepts of support vector machine, Section 3: Near Sets, Section 4: Face Recognition System, Section 5: Experimental results and Section 6: Conclusions.

2. BASIC CONCEPTS

2.1 Linear Classification

In the linear case for classification, with the training data $\{x_i,y_i\}$,$y_{es}\in \{-1,+1\}$, i=1,…..,l,

the output of an SVM can be written as:

$$x_i.w + b \geq +1 \text{ for } y_i = +1 \quad (1)$$

$$x_i.w + b \leq +1 \text{ for } y_i = -1 \quad (2)$$

These can be combined into one set of inequalities:

$$y_i\,(x_i.w + b) - 1 \geq 0 \quad i=1,2.....,l \quad (3)$$

Here w is a vector normal to the hyperplane $x_i.w + b = 0$ and b is called threshold.

The solution form of a typical 2-dimensional classification is shown in Figure 1, in which, the support vectors are darkened in color. The margin between the two parallel hyper-planes H_1, and H_2, is **2** / || w || .

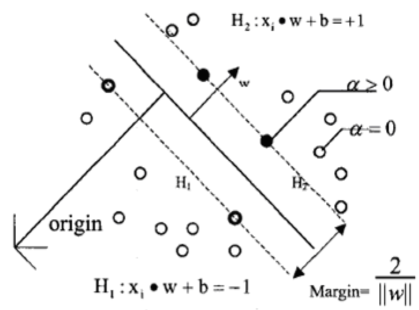

Figure 1: Linear separable case with largest margin

From Figure1 it can be seen that the margin is determined by *w* and *b* for the same training data. To optimize the classification is to make the margin as large as possible. Thus the optimization task is defined to minimize

$$L_p = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{l}\alpha_i y_i (x_i.w + b) + \sum_{i=1}^{l}\alpha_i \quad (4)$$

Where α_i $(i = 1, ...,l)$are non-negative Lagrange multipliers. For inequality constraints in (3), $\alpha_i = 0$ .For each of equality constraints in (3), $\alpha_i \geq 0$.

From $\frac{\partial L_P}{\partial w} = 0$ and $\frac{\partial L_P}{\partial b} = 0$, we have,

$$w = \sum_{i=1}^{l}\alpha_i y_i x_i \quad (5)$$

$$\sum_{i=1}^{l}\alpha_i y_i = 0 \quad (6)$$

Substitute equations (5) and (6) into (4), we get the dual problem of (4),

$$L_D = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j x_i . x_j \qquad (7)$$

Subject to :

$$\alpha_i \geq 0 \, , \; \sum_{i=1}^{l} \alpha_i y_i = 0$$

Then we change the problem from minimizing the primal problem (4), (5), and (6) to maximize the dual problem (7). Using the optimal solution, combined with KKT conditions as convergence criteria, the problem (7) can be solved. Note that there is a weight a_i for every training point (or vector), and for those points whose $a_i > 0$ are called support vectors.

2.2 Nonlinear Classification

In most cases, the training set X=**{ x_i , y_i }** is not linear separable, even with soft margins. But it is still possible to project **{ x_i }** into a feature space F=**{ (x_i)}** such that **{** (x_i), y_i } is linear separable as shown in Figure 2.

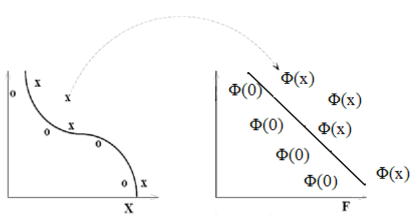

Figure 2: Mapping of training set X into feature space F

After projection, in the test phase, an SVM is used to compute sign of

$$f(x) = \sum_{i=1}^{N_s} \alpha_i y_i \phi(x_i) \phi(x) + b \qquad (8)$$

Where N_s is the number of support vectors.

The problem is that the explicit form of □(x) is difficult to be found, then we define $K(x_i, x_j)$ = □(x_i). □(x_j) as the kernel function.. After such a replacement, Equation (8) becomes,

$$f(x) = \sum_{i=1}^{N_s} \alpha_i y_i K(x_i, x) + b \qquad (9)$$

The selection of the kernel is a keen issue in the training. It is required that the kernel function should satisfy the Mercer's condition, detailed description of which can be found in [9].

3. NEAR SETS

The basic idea in the near set approach to adaptive learning is to compare behavior descriptions. In general, sets *X, X'* are considered near each other if the sets contain objects with at least partial matching descriptions. Let *~B* denote $\{(x, x') | f(x) = f(x') \; \forall f \in B\}$ (called the indiscernibility relation).

Let *X, X'* ⊆ *O*, *B* ⊆ *F*. Set *X* is near *X'* if, and only if there exists x ∈ *X*, *x'* ∈ *X'*, □$_i$ ∈ *B* such that *x* ~{ □$_i$} *x'*.

Object recognition problems, especially in adaptive learning, images and the problem of the nearness of objects have motivated the introduction of near sets.

3.1. Nearness Approximation Spaces

The original generalized approximation space (GAS) model has been extended as a result of recent work on nearness of objects. A nearness approximation space (NAS) is a tuple

$$NAS = (O, F, \sim B_r, N_r, \nu_{Nr}),$$

defined using set of perceived objects *O*, set of probe functions *F* representing object features, indiscernibility relation $\sim B_r$ defined relative to $B_r \subseteq B \subseteq F$, family of neighbourhoods N_r, and neighbourhood overlap function ν_{Nr} . The relation $\sim B_r$ is the usual indiscernibility relation from rough set theory restricted to a subset $B_r \subseteq B$. The subscript *r* denotes the cardinality of the restricted subset B_r, where we consider (|B|, r) *i.e.*, |B| functions □$_i$ ∈ *F* taken *r* at a time to define the relation *~B* . This relation defines a partition of *O* into nonempty, pairwise disjoint subsets that are equivalence classes denoted by $[x]_{Br}$,

Where

$$[x]_{Br} = \{x' \in O \mid x \sim_{Br} x'\}.$$

These classes form a new set called the quotient set $O/\sim B_r$,

Where $O/\sim B_r = \{[x]_{Br} | \; x \in O\}$.

In effect, each choice of probe functions B_r defines a partition ξ_{Br} on a set of objects O, namely,

$$\xi_{Br} = O/{\sim}B_r \,.$$

Every choice of the set B_r leads to a new partition of O. The overlap function ν_{Nr} is defined by

$$\nu_{Nr} : P(O) * P(O) \rightarrow [0, 1],$$

where $P(O)$ is the power set of O. The overlap function ν_{Nr} maps a pair of sets to a number in [0, 1] representing the degree of overlap between sets of objects with features defined by probe functions $B_r \subset B$. For each subset $B_r \subset B$ of probe functions, define the binary relation $\sim B_r$ = {(*x, x'*) ∈ *O***O* : ∀ □$_i$ ∈ B$_r$, □$_i$(x)=, □$_i$(x') } .

Since each $\sim B_r$ is, in fact, the usual indiscernibility relation, let $[x]_{Br}$ denote the equivalence class containing x, *i.e.*, $[x]_{Br} = \{x' \in O \mid \forall f \in B_r, f(x') = f(x)\}$.

If $(x, x') \in \sim B_r$ (also written $x \sim B_r\, x'$), then x and x' are said to be *B-indiscernible* with respect to all feature probe functions in B_r. Then define a collection of partitions $N_r(B)$ (families of neighbourhoods), where

$$N_r(B) = \{\xi_{Br} \mid B_r \in B\}.$$

Families of neighbourhoods are constructed for each combination of probe functions in B using (|B|, r), *i.e.*, |B | probe functions taken r at a time. The family of neighbourhoods $N_r(B)$ contains a set of precepts.

$N_r(B)$ contains a set of precepts. A *percept* is a by product of perception, *i.e.*, something that has been observed. For example, a class in $N_r(B)$ represents *what has been perceived about objects belonging to a neighbourhood, i.e.*, observed objects with matching probe function values.

It is now possible to introduce a near set-based form of coverage that extends the basic rough coverage model. That is, we can formulate a basis for measuring the degree of overlap between each class in $N_r(B)$ and the lower approximation $N_r(B)_*X$ of a set X for each choice of r. The lower approximation $N_r(B)_*X$ defines a standard for classifying perceived objects.

The notation $B_j(x)$ denotes a class in the family of neighbourhoods in $N_r(B)$, where NL ∈ B_r.

Put $\nu_{NL}([x]_{Br}, N_r(B)_*X) = |[x]B_r\, N_r(B)_*X| / |\, Nr(B)_*X \,|$ where ν_j is defined to be 1, if $N_r(B)_*\, X =$□.

Put $B = \{[x]B_r : \mathrm{NL}(x) = j,\ x \in O\}$, a set of equivalence classes that represent nose length NL(x) = j. Let D denote a decision class, *e.g.*, $D = \{x \mid d(x) = 1\}$, a set of object having acceptable behaviours. Define $\bar{\nu}_{\mathrm{NL}}(t)$ (near lower average coverage) with respect to an feature nose length NL(x) = j.

$$\bar{\nu}_{\mathrm{NL}} = \frac{1}{|B|} \sum_{[x]_{B_r} \in B} \nu([x]_{B_r}, N_r(B)_* D) \qquad (10)$$

4. FACE RECOGNITON SYSTEM

The experiment is performed on the Cambridge ORL face database [23], which contains 400 images of 10 distinct persons. Our system does not include a face detection step. The input of our system is a complete face in the image; we assumed that the faces have already been detected. Figure 3 shows the basic diagram of face recognition system which consists of following phases:

PHASE I: Feature Extraction
PHASE II: Face Recognition

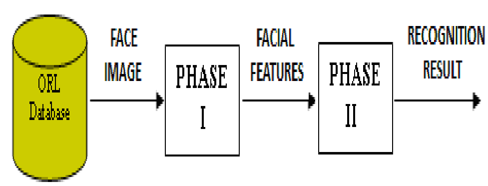

Figure 3: Face Recognition system

4.1. Feature Extraction

The first step towards face recognition is the detection of the facial features. Figure 4 shows the working of this phase. Feature extraction is to obtain features that are fed into a face recognition system (such as lines or fiducially points, or facial features such as eyes, nose, and mouth).The approach for facial feature extraction is not an established procedure but we have developed our own automated procedure to extract facial features which is explained in Algorithm 1 . Also in Algorithm 2 we propose the procedure to select partition for every combination of three features and using

Algorithm 3 we choose the feature i.e. nose length NL which have highest value of $\overline{\nu}$ which gives highest face recognition accuracy.

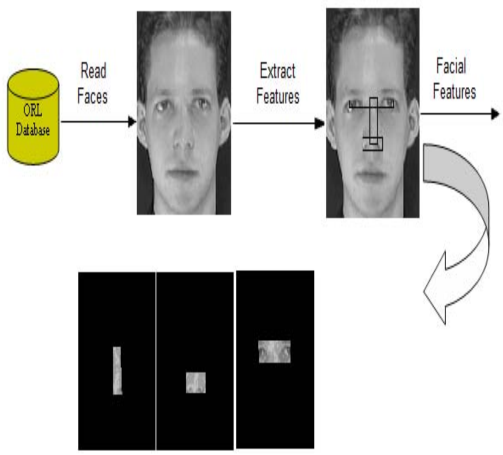

Figure 4: PHASE I

Algorithm 1: Facial Features Extraction Method

Input : Face images, i ∈ I
Output: Facial features: Nose Length (NL), Nose Width (NW) and distance between eyes balls(ED)

while *True* **do**

For (i = 0; i <= I) **do**
Initialize NL (i) =0, NW (i) =0 and ED (i) =0;
Repeat (for number of features):
Retrieve image I;
Extract Region Of Interest (ROI) for each feature;
Calculate (using distance formula):
NL (i) =NL $(i)_b$ - NL $(i)_a$;
NW (i) =NW $(i)_b$-NW $(i)_a$;
ED (i) =ED $(i)_b$-ED (i) a;
Until all futures are extracted.
End

End

Algorithm 2: Partition Selection

Input : $NAS = (O, F, \sim B_r, N_r, \nu_{Nr}), B \in f$, *choice r.*
Output: Partition size list □, where □[i] =number of classes in partition $\xi_{o,Br} \in N_r(B)$.

Initialize i=0;

while (i <= | $N_r(B)$|) **do**

Select ith partition $\xi_{o,Br} \in N_r(B)$;
□[i] = $|\xi_{o,Br} \in N_r(B)|$;
i=i+1;

end

Algorithm 3: Feature Selection

Input : Array □, □[i] =number of classes in $\xi_{o,Br} \in N_r(B)$,threshold th.
Output: Ordered list LIST, where LIST[i] is a winning probe function.

Initialize i=0;
Sort □ in descending order based on the value of average near coverage $\overline{\nu}$;

while (i ≥ th) **do**

Select ith partition $\xi_{o,Br} \in N_r(B)$;
LIST[i]= □[i];
i=i+1;

end

4.2. Face Recognition

Once facial features has been extracted, SVM is used as face recognizer as shown in Fig.5.

Basic method of face recognition system is given in Algorithm 4 .In this algorithm α_i the embedded strength of the pattern x_i i.e. the faces to be recognized which gives the number of times misclassification of x_i has caused the weight to be updated. This value of embedded strength i.e. b is used in the decision function h(x),which gives the recognition result for the test image as shown in Fig. 5.

Algorithm 4: Face Recognition Algorithm using SVM.

Input : Training set S consist of facial features obtained using algo.3,l: Total no of persons to be recognized,

$\alpha \leftarrow 0; b \leftarrow 0$

$R \leftarrow \max_{1 \le i \le l} \|x_i\|$

Output: Recognized face

Repeat

for (i = 0; i< l; i++) **do**

if $y_i \left(\sum_{j=1}^{l} \alpha_j y_j (x_j . x_i) + b\right) \le 0$

$\alpha_i \leftarrow \alpha_i + 1;$

$b \leftarrow b + y_i R^2;$

end if

until no mistakes made within the for loop

return (α,b);

$h(x) = \mathrm{sgn}((w.x) + b);$

$= \mathrm{sgn}\left(\left(\sum_{j=1}^{l} \alpha_j y_j x_j . x\right) + b\right);$

$= \mathrm{sgn}\left(\sum_{j=1}^{l} \alpha_j y_j (x_j . x) + b\right);$

end

end

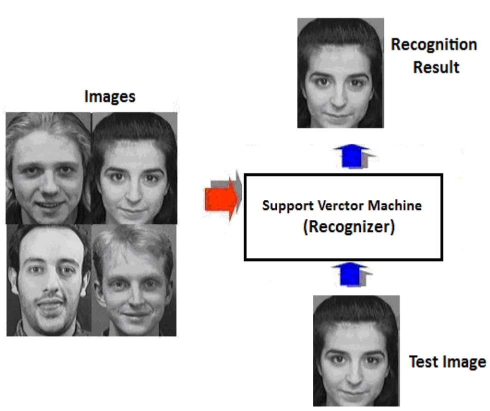

Figure 5: PHASE II

5. EXPERIMENTAL RESULTS

There are several approaches for recognition of the ORL face database. In [4], a hidden Markov model (HMM) based approach is used, and the best model resulted in a 13% error rate. Later, Samaria extends the top-down HMM [4] with pseudo two-dimensional HMMs [24], the error rate reduces to 5% and in our experiment this error rate reduces to 0.2%.To get better results we divide the data into eight part using the concept of V-Fold Cross Validation [25] .We have used RBF kernel with parameter γ equal to 32, and the error penalty parameter C[10] taken to be equal to 100.

Table1.Average Near Coverage ($\bar{\nu}$) Summary Table

Function Name	Average Near Coverage($\bar{\nu}$)
$\bar{\nu}_{NL}$	0.0500
$\bar{\nu}_{NW}$	0.0769
$\bar{\nu}_{ED}$	0.0714
$\bar{\nu}_{NLNW}$	0.0118
$\bar{\nu}_{NLED}$	0.0141
$\bar{\nu}_{NWED}$	0.0208
$\bar{\nu}_{NLNWED}$	0.00591

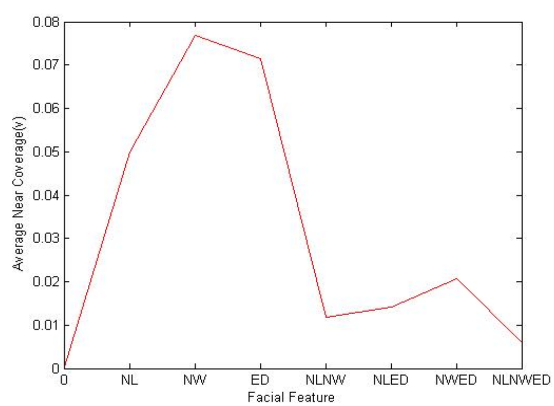

Figure 6: Average Near coverage Vs. Facial Features

Table1 summarized the $\bar{\nu}$ values for all combinations of the chosen three features, also it can be seen from Table 1 and Figure 6 that $\bar{\nu}$ value is maximum for the feature nose width (NW).

Table2. Illustration of %age error rate and correctly recognized samples for different combinations of chosen features

No. of Features	%age Error Rate	No. of correctly recognized samples
1 (NW)	0.2	49
1 (ED)	2.8	38
1 (NL)	3.3	37
2 (NLNW)	2.8	39
2 (NLED)	2.6	39
2 (NWED)	1.7	43
3 (NLNWED)	4.4	32

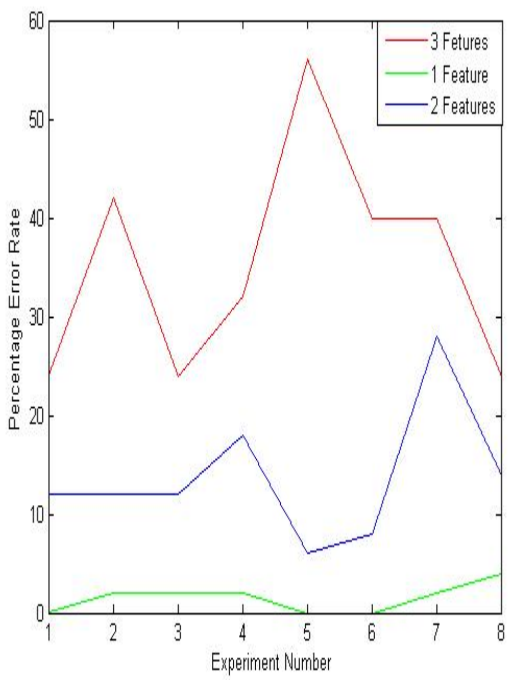

Figure 7. Percentage Error Rate Vs. Experiment Number

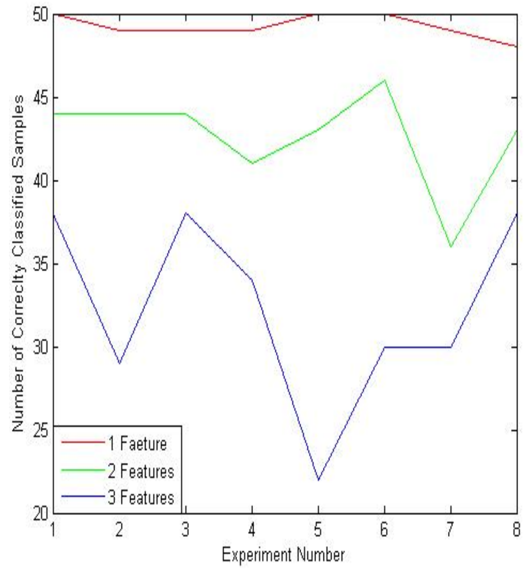

Figure 8. Number of correctly classified samples Vs. Experiment Number

Also Table 2, Figures 7 and 8 illustrates that when feature nose width is chosen the % error rate is least and the number of correctly classified samples is maximum.

Table 3.Face Recognition Result

No. of Features	Objective Function $\|\|W^2\|\|)/2$	Bias (b)
1 (NW)	-1112506.7	-2.5494483
1 (ED)	-247.04	-0.8931243
1 (NL)	-206.99	-0.808815
2 (NLNW)	-127.1746	-5.6
2 (NLED)	-106.244	-0.43
2 (NWED)	-92.3	-20.5
3 (NLNWED)	-207.0	-0.1

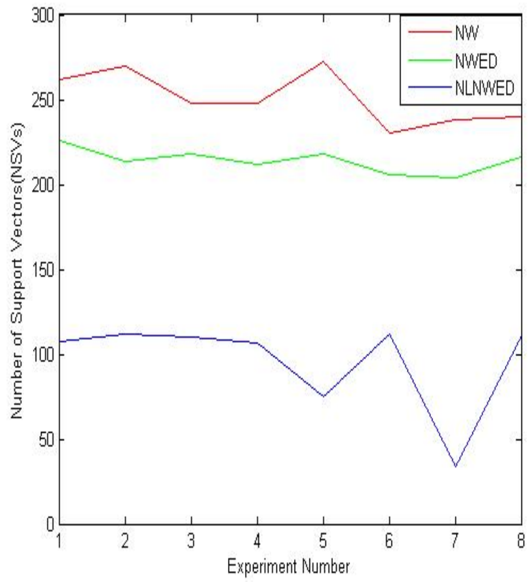

Figure 9. NSVs Vs. Experiment Number

Tables 3 summarizes the face recognition results. From Table 3 it can be seen that value of objective function i.e $||W^2||)/2$ is least when selected feature is nose width, which has to be so as to acquire maximum margin and better recognition accuracy. Also Figure 9 illustrates that NSVs are maximum when the chosen feature is nose width which give in better recognition results.

6. CONCLUSION

This paper present effect of chosen feature on the accuracy of face recognition system.Results shows that number of support vectors and margin are maximum when the feature with largest average near coverage($\bar{\nu}$) is chosen for face recognition.Future work will include Consideration for more features and more faces.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the suggestions and constant support by Er.N.Gupta, Department of Atomic Energy, Bhabha Atomic Reseach Center, Mumbai, India and insights from Er.Pooja Goyal, Bharat Sanchar Nigam limited, Karnal.India

REFERENCES

[1]. A. Samal and P. A. Iyengar, “Automatic recognition and analysis of human faces and facial expressions: A survey”, *Pattern Recognition. Vol* 25, 1992, pp. 65–77.

[2]. D. Valentin, H. Abdi., A. J. O’Toole, and G. W. Cottrell, “Connectionist models of face processing: A survey”, *Pattern Recognition .Vol.* 27, 1994, pp.1209–1230.

[3]. R. Chellappa, C. L.Wilson, and S. Sirohey, “Human and machine recognition of faces: A survey”, *Proc. IEEE.Vol.* 83, May 1995, pp.705– 741.

[4]. M.A. Turk and A.P. Pentland, “Face recognition using eigenfaces”, *In: Proc. Of International Conference on Pattern Recognition.* 1991, pp. 586-591.

[5]. P.S. Penev and L. Sirovich, “The global dimensionality of face space”, *In: 4th IEEE International Conference on Automatic Face and Gesture Recognition.* 2000, pp.264-270.

[6]. F. Samaria, “Face Recognition Using Hidden Markov Models”, *Ph.D. Thesis, University of Cambridge,* 1994.

[7]. L.Wiskott, J.Fellous, N. Kuiger, and C.Malsburg, “Face recognition by elastic bunch graph matching”, *IEEE Transaction on Pattern Analysis and Machine Intelligence.* Vol. 19, no.7, 1997, pp. 775- 779.

[8]. Yonggwan Won, Paul D. Gader, and Patrick C. , “Coffield Morphological haredweight networks with applications to automatic target recognition”, *IEEE Transaction on Neural Network.* Vol. 8, no. 5, 1997.

[9]. V. N. Vapnik. *Statistical learning theory*. John Wiley & Sons, New York,1998.

[10]. E. Osuna, R. Freund and F. Girosi, “Training support vector machines: an application to face detection”, *In IEEE Conference on Computer Vision and Pattern Recognition.* 1997, pp. 130-136.

[11]. Pawlak, Z,” Rough Sets”, Institute for Computer Science, Polish Academy of Sciences, Report 431, March (1981).

[12]. Peters, J.F. Near sets,”eneral theory about nearness of objects”,Applied Mathematical Vol 53,2007,pp 2609-2029.

[13]. Peters, J.F.,”Classification of Perceptual Objects by Means of Features”, International Journal of Information Technology and Intelligent Computing (2007), *in press*.

[14]. Pawlak, Z.,”Classification of Objects by Means of Attributes”, Institute for Computer Science, Polish Academy of Sciences, Report 429, March (1981).Agent Technology, Compieegne University of Technology, France 19-22 Sept 2005.

[15]. Orlowska, E, “Applications of Rough Sets, Semantics of Vague Concepts”, Institute for

Computer Science, Polish Academy of Sciences, Report 469, March (1982). Published in Dorn G, Weingartner P (eds), Foundations of Logic and Linguistics. Problems and Solutions, Plenum Press, London/NY (1985) 465-482.

[16]. A. Tefas, C. Kotropoulos, I. Pitas, "Using support vector machines to enhance the performance of elastic graph matching for frontal face authentication", *IEEE Transaction on Pattern Analysis and Machine Intelligence.* 2001, pp. 735-746.

[17]. Steve R. Gunn, "Support Vector Machines for Classification and Regression", Technical Report, pp.10- 13, 1998.

[18]. O.L. Mangasirian, "Large Scale Kekel Regression via Linear Programming", *Machine Learning.* Vol. 46, 2002, pp. 255-269.

[19]. B. Scholkopf, A. Smola, and K.R. Muller, "Nonlinear component analysis as a kemel eigenvalue problem", *In press Neural Computation,* 1998.

[20]. J. Platt, "Fast training of support vector machines using sequential minimal optimization", In *B.,Sch'olkopf; C.J.C.Burges, and A.J. Smola, eitors, Advances in Kernel Methods-Support Vector Leaning,MTT Press,* 1999, pp 185-208.

[21]. G.W. Flake and S. Lawrence, "Efficient SVM Regression Training with SMO", *Machine Learning. V*o1.46, 2002, pp. 271-290.

[22]. S.S. Keerthi, S.K. Shevade, C.Bhattacharyya, and K.R.K.Murthy, "A fast iterative nearest point algorithm for support vector machine classifier design", *IEEE Transactions on Neural Networks,* Vol. 1I, January 2000.

[23]. http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html.

[24]. Y. Moses, Y. Adini, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction", *European Conference on Computer Vision.* 1994, pp. 286–296.

[25]. Chin-Wei Hsu,Chin-Chung chang and Chin-Jen Lin.,"A practical Guide to Support Vector Classificaion", Department of Computer Science National Taiwan University ,Taipei 106,Taiwan.