

## **IROLS Based Radial Basis Function Neural Network for Face Recognition**

Venkata Siva Kumar K<sup>1</sup>, Sofia Priya Dharshini J<sup>2</sup>

*1 (M.Tech [DSCE], RGM CET, )*

*2 (Associate prof., ECE, RGM CET)*

*India*

**ABSTRACT**--Face representation (FR) plays a typically important role in face recognition and methods such as principal component analysis (PCA) and linear discriminant analysis (LDA) have been received wide attention recently. These FR methods will inevitably lead to poor classification performance in case of great facial variations such as expression, lighting, occlusion and so on, due to the fact that the image gray value matrices on which they manipulate are very sensitive to these facial variations. The recognition of faces is very important because of its potential commercial applications, such as in the area of video surveillance, access control systems, retrieval of an identity from a data base for criminal investigations and user authentication. The recognition performance of the face recognition system deteriorates when the system is exposed to the real world scenario. This problem happens because we do not have a complete set of training samples that consists of all types of visual variations. Furthermore, the extendibility of the system to recognize more new people who join the existing groups in the future may cause a problem to the system. In this work, a radial basis function (RBF) neural network with a new incremental learning method based on the regularized orthogonal least square (ROLS) algorithm is proposed for face recognition. It is designed to accommodate new information without retraining the initial network. In addition, it accumulates previous experience and learns updated new knowledge of the existing groups to increase the robustness of the system. The proposed work is to be developed on Matlab platform for its realization.

**Index Terms** —Face recognition, incremental learning, neural network, orthogonal least square, radial basis function (RBF), visual variation.

### **I. INTRODUCTION**

The recognition performance of the face recognition system deteriorates when the system is exposed to the real- world scenario. This problem happens because we do not have a complete set of training samples that consists of all types of visual variations. Furthermore, the extendibility of the system to recognize more new people who join the existing groups in the future may cause a problem to the system. A conventional way to solve this problem is to retrain the system to learn the new information. However, this will cause high computational complexity. A more practical and faster way to solve this problem is by embedding incremental learning in feature selection and classification. In recent years, several incremental learning methods have been developed independently. As for incremental learning in feature selection, principal component analysis (PCA) was extended to incremental PCA (IPCA) and was applied in the face recognition systems. If the labels of the data are available, the incremental linear discriminant analysis (ILDA) which maximizes the between-class scatter and minimizes the within class scatter can be used to optimize the class separability incrementally. Since the feature space of these methods is updated every time a new training sample is available, the dimension of the feature space might be increased. This will cause a problem in most neural network classifiers as the networks should also be updated to adapt to the dimension extension of the input data.

In this paper, we propose a new incremental learning method for the regularized orthogonal least square (ROLS)-based radial basis function (RBF) neural network. This proposed algorithm is named as the incremental ROLS (IROLS) algorithm. The IROLS algorithm is designed to accommodate a new class and updated new data while avoiding retraining the network. The IROLS algorithm combines the zero-order regularization with the orthogonal least square to construct a parsimonious RBF network to improve the generalization ability of the system. This algorithm is capable of constructing small RBF networks which generalize well and requires low computational complexity. For the conventional ROLS algorithm, it first selects the basis vector that provides the most significant error reduction among all the basis vectors and orthogonalizes all the remaining basis vectors into a Euclidean space formed by the selected basis vector. This process will be

repeated until the error reduction ratio is lower than a predetermined threshold. Unlike the conventional ROLS algorithm, the proposed IROLS reduces the computational complexity of the ROLS algorithm by selecting the basis vector which corresponds to the new class of training data locally instead of globally. For updating data, the incremental learning of the updated training sample is achieved by reinforcing the positions of hidden neurons to the new information, and this avoids the retraining process. The proposed IROLS-algorithm-based RBF neural network is then applied to address the face recognition problems: the ability to add a new class to the existing classifier and to accumulate new knowledge without retraining the system.

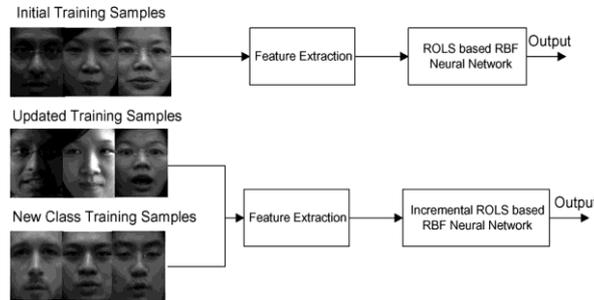


Fig.1.IROLS framework for face recognition system

As shown in Fig. 1, we allow the new information to appear in two forms: the new classes that need to be included in the existing groups and the updated training samples of the existing groups in the system. The proposed learning scheme not only allows the addition of new classes but also allows the addition of new updating samples to existing groups to accumulate knowledge to help future recognition processes. During the training stage, each of the training samples will be first transformed into feature representation. The training samples which belong to the initial classes will be first trained by the ROLS-based neural network. When new class samples appear, the IROLS-based neural network will add the new class data to the existing network without retraining the whole network. For new updated samples, the IROLS will learn the new information in the updated samples, together with the previous knowledge in the network. In addition, we investigate how the updated new knowledge that contains different visual variations from the training samples can help to improve the recognition performance of the system. To the best of our knowledge, this is the first study that specifically focuses on the influence of incremental learning of various visual variations to the face recognition performance. The recognition performance of the proposed algorithm will be tested on three face databases and compared with that of other state-of-the-art methods.

This paper is organized as follows. In Section II, the details of the proposed IROLS-based neural network will be presented. Then, the implementation of the proposed neural network in the face recognition system will be discussed in Section III. The experimental results of the proposed IROLS-based neural network will be given in Section IV. Section V describes the conclusion of this paper.

## II. IROLS-ALGORITHM-BASED NEURAL NETWORK

The ROLS algorithm is used to build a global scheme where a new class or updated data can be added after the initial set of classes is trained. The proposed IROLS comprises two major steps: First, the ROLS algorithm is run to create the initial model which contains the knowledge of the initial classes, and second, the training samples which can be the training samples for the new classes or updated data are presented to the IROLS for incremental learning.

### A. ROLS Algorithm for RBF Neural Network

The RBF neural network is formulated as a linear regression model as follows:

$$f_r(\mathbf{x}) = \sum_{i=1}^{ns} \theta_i \phi(\|x - c_i\|)$$

where  $\mathbf{x} = \{x_1 \dots x_m\}$  is the training data of the initial training set with the desired outputs  $y = \{y_1 \dots y_m\}$

$\}$ ,  $m$  is the total number of training samples,  $\theta_i$  represents the weights,  $c = \{c_1 \dots c_{n_S}\}$  represents the RBF centers,  $\cdot$  refers to the Euclidean norm,  $\varphi(\cdot)$  is the nonlinearity of the hidden neurons, and  $n_S$  is the total number of selected hidden neurons. At the first stage of the selection procedure  $l = 1$  of the algorithm,  $n_S = m$ . The ROLS algorithm for the RBF neural network is discussed hereinafter.

Assuming that we have every training sample  $x = \{x_1 \dots x_m\}$  as the center  $c_j = x_j$ , for  $1 \leq j \leq m$ , we compute the regressor matrix  $\Phi$  as follows:

$$\Phi_j = \exp\left(-\frac{\|x - c_j\|^2}{2\sigma^2}\right),$$

Where  $\sigma$  is the width of the Gaussian function.

1) For  $1 \leq j \leq m$ : Test: Conditioning number check. If  $(\Phi_j^{(l-1)})^T \Phi_j^{(l-1)} < T_z$ , the  $j$ th candidate is not considered, where  $T_z$  is a very small positive value. This checking is to avoid illconditioning or singular situations. The  $T_z$  term is set as 0.02 in this paper. Compute

$$g_l^{(j)} = ((\Phi_j^{(l-1)})^T y^{(l-1)}) / ((\Phi_j^{(l-1)})^T \Phi_j^{(l-1)} + \lambda), \quad 1 \leq s \leq n_T$$

$$[rerr]_l^{(j)} = (g_l^{(j)})^2 ((\Phi_j^{(l-1)})^T \Phi_j^{(l-1)} + \lambda) / y^T y$$

where  $\lambda$  is the regularization parameter,  $g$  is the orthogonalized weight,  $y$  is the desired output matrix, and  $\Phi^{(l-1)}$  is the  $j$ th column of the regressor matrix of selection stage  $l-1$ .

2) Select the significant regressors by using the regularized error reduction ratio

$$[rerr]_l = \max\{[rerr]_l^{(j)}, 1 \leq j \leq m, j \text{ passes Test}\}$$

The above equation selects the significant regressors that have the maximum error reduction ratio among the regressors that pass the test at the selection stage  $l$ . The selected regressor will be assigned as the center of the  $l$ th selection stage. For  $l = 1$ , the  $j$ th column of  $\Phi^{(l-1)}$  is interchanged with the  $l$ th column of  $\Phi^{(l-1)}$ .

3) Perform the orthogonalization as follows.

For  $l = 1, 2, \dots, m-1$

$$w_l = \Phi^{(l-1)}$$

$$a_{l,j} = w_l^T \Phi^{(l-1)} / (w_l^T w_l), \quad l+1 \leq j \leq m$$

$$\Phi^{(l)} = \Phi^{(l-1)} - a_{l,j} w_l, \quad l+1 \leq j \leq m$$

where  $a_{l,j}$  is substituted to the  $l$ th row of  $A$ .  $A$  is an  $m \times m$  triangular matrix with 1's on the diagonal and 0's below the diagonal.  $A$  and  $W$  are as follow:

$$A = \begin{pmatrix} 1 & a_{1,2} & \dots & a_{1,m} \\ 0 & 1 & & \dots \\ \vdots & & \ddots & \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

$$W = [w_1 \dots w_m]$$

with orthogonal columns that satisfy

$$w_i^T w_j = 0, \quad \text{if } i \neq j.$$

Then, calculate  $g_l$  and update  $y^{(l-1)}$  into  $y_l$  in the way shown hereinafter.

For  $1 \leq l \leq m$

$$g_l = (w_l^T y^{(l-1)}) / (w_l^T w_l + \lambda)$$

$$y^{(l)} = y^{(l-1)} - g_l w_l$$

When  $l > 1$ , the selection procedure repeats, the  $j$ th column of  $\Phi^{(l-1)}$  is interchanged with the  $l$ th column of  $\Phi^{(l-1)}$ , and the  $j$ th column of  $A$  is interchanged up to the  $(l - 1)$ th row with the  $l$ th column of  $A$ . This selects the  $j$ th candidate as the  $l$ th regressor in the subset model.

The selection is terminated at the  $n_s$  stage when

$$1 - \sum_{k=1}^{n_s} [rerr]_k < \xi$$

Is satisfied, where  $0 \leq \xi \leq 1$  is a chosen tolerance. Alternatively, the selection will be terminated when there are no more candidates which would not cause an illconditioning or singular problem .

**B. IROLS Algorithm for RBF Neural Network:**

After  $n_s$  stages, an  $n_s$ -dimensional Euclidean space  $E^{n_s}$  is established which has the orthogonal basis vectors  $\{w_1, w_2, \dots, w_{n_s}\}$ . The original data that correspond to the selected basis vectors can be referred to as the centers  $C = \{x_1, x_2, \dots, x_n\} = \{c_1, c_2, \dots, c_n\}$  of the initial network. Assuming that the new training samples are  $x_{NEW} = \{x_1, \dots, x_n\}$ , where  $n_M$  is the number of new training samples, we perform the following procedure. The new input training samples are first determined manually whether the samples are an additional new class or samples that belong to the existing class (updated sample). If the input training samples are from a new class, go to Step 1). Otherwise, go to Step 7)

Step 1) Let the new training samples of that class be  $x_{NEW} = \{x_1, \dots, x_{n_M}\}$ , where  $n_M$  is the total number of new training samples which belong to a new class. Calculate the mean  $x_{NEW}$ , and find the three centers from  $C_j \in \{c_1, \dots, c_{n_s}\}$  which are closest to  $x_{NEW}$  as follows:

$$d_j = \|x_{NEW} - c_j\|, \quad 1 \leq j \leq n_s$$

Then, the first three  $j$ th centers that are closest to  $x_{NEW}$  will be selected. The three closest centers are chosen based on an experimental result which tested on the influence of the number of centers to the recognition accuracy. The results show that the proposed algorithm achieves the highest recognition accuracy when the number of centers is three.

Step 2) Assuming that the corresponding training samples which belong to the three centers are  $x_c = \{x_1, \dots, x_{n_O}\}$  and concatenating them with the new training samples from  $x = \{x_c, x_{NEW}\} = \{x_1, \dots, x_{n_T}\}$ . Since  $x = c_s$ , then the regressors of the input  $x$  are as follows:

$$\Phi_s = \exp\left(-\frac{\|x - c_i\|^2}{2\sigma^2}\right), \quad 1 \leq s \leq n_T$$

Compute  $g_s = ((\Phi_s)^T) / ((\Phi_s)^T \Phi_{s+1} \lambda)$ ,  $1 \leq s \leq n_T$

$$[rerr]_s = (g_s)^2 ((\Phi_s)^T \Phi_{s+1} \lambda) / y^T y$$

where  $y$  is the local desired output of the training samples. This step is important to compute the error reduction ratio of the new training samples with the three closest classes.

Step 3) The regressor for the new class is selected according to the scalar measure of  $[rerr]$  where only the re- gressors that correspond to the new class are chosen to be the candidates

$$[rerr]_p = \max \{[rerr]_p, n_o + 1 \leq p \leq n_T\} .$$

This is to choose the regressor  $\Phi_p$  belonging to the new class locally instead of reselecting  $\{\Phi_1, \Phi_2, \dots, \Phi_{n_s}, \Phi_p\} = \{\Phi_1, \Phi_2, \dots, \Phi_{nk}\}$ , Where  $nk = n_s + 1$ .

Step 4) Compute the regressors of input training set  $x = \{x, x^{NEW}\}$  with  $c = x$  as follows:

$$\Phi_j^{New} = \exp\left(-\frac{\|x' - c_j\|^2}{2\sigma^2}\right), \quad 1 \leq s \leq n_H,$$

where  $n_H = m + n_M$

Step 5) Perform the orthogonalization with the regressors  $\{\Phi_1, \Phi_2, \dots, \Phi_{n_S}, \Phi_p\}$  in which the first  $n_S$  th regressor correspond to the regressors chosen during the training of the initial network. For  $l=1, 2, \dots, n_k$

$$w_l = \Phi_l^{(l-1)}$$

$$a_{l,j} = W_l^T \Phi_l^{(l-1)} / (W_l^T W_l), \quad l+1 \leq j \leq n_H$$

$$\Phi_l^{(l)} = \Phi_l^{(l-1)} - a_{l,j} w_l, \quad l+1 \leq j \leq n_H.$$

Then, the elements of  $g$  and  $y$  are computed as

Follows:

$$g_l = w_l^l y^{(l-1)} / w_l^l w_l + \lambda$$

$$y^{(l)} = y^{(l-1)} - g_l w_l.$$

The  $j$ th column of  $\Phi_l^{(l-1)}$  is interchanged with

the  $l$ th column of  $\Phi_l^{(l-1)}$ , and the  $j$ th column of  $A$  is interchanged up to the  $(l-1)$ th row with the  $l$ th column of  $A$ .

Step 6) Calculate the output of the network  $y = wg$ . If the  $j$  th output is not equal to the desired output, then Steps 1)-3) are repeated to select the three centers which are nearest to the center of a class that  $x_j$  belongs to. Note that the “new class” and “new training sample” referred to in Steps 1)-3) should now be referred to as the class of the misclassified training sample  $x_j$ . Then, add a hidden unit (i.e.,  $n_k \leftarrow n_k + 1$ ) to the network if the selected regressor has not been selected before  $\Phi_{n_k}^{(n_j \in n_k)}$ . Then, Steps 4)–6) are computed. This step is repeated once to reduce the training error. Otherwise, the following procedure is carried out.  $x_{NEW} = x_{up}$  and the center correspond to the class of the updated data that are selected from  $C_{up} \in \{C\}$  or  $C_{up} \in \{C_{NEW}\}$ . If there is more than one center which belongs to the updated class, the center which has the closest distance to the updated sample will be selected as the  $C_{up}$ . The  $C_{up}$  is updated as follows:

$$C_{up}^{(New)} = C_{up}^{(old)} + \eta (X_{up} - C_{up}^{(old)})$$

where  $\eta$  is a positive learning rate. Then, repeat Steps 4)–6)

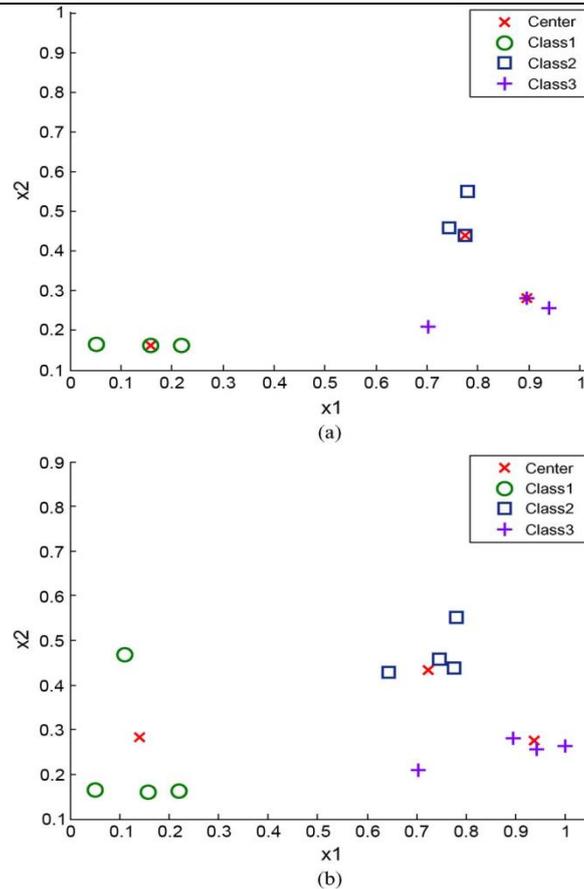


Fig. 2. Data scatter diagram for (a) the initial training samples and selected centers and (b) the updated samples and the adjustment of the old centers to the new updated samples.

The proposed algorithm avoids the retraining of the network by using the initial network architecture. The regressor which corresponds to the new training data will be selected locally with the data that have the minimum distance to the new training data. The incremental learning of the updated training sample is achieved by reinforcing the old center to the updated training samples. Fig. 2 shows how works. The  $\eta$  parameter determines how much the old centers have moved toward the updated samples.

### III. PROPOSED IROLS ALGORITHM BASED NEURAL NETWORK FOR FACE RECOGNITION:

In the previous section, the details of the proposed IROLS-based RBF neural network are presented. In this section, the proposed IROLS will be implemented in the face recognition system. In this paper, we employ the dual optimal multiband feature (DOMF) fusion method as the feature extraction module. This method is suitable for incremental learning because it provides a fixed input dimension to the neural network, and it extracts the optimal sets of wavelet subbands that are invariant to facial expression and illumination.

It uses an adaptive fusion method to avoid the adverse effect caused by combining the optimal feature sets. Fig. 3 shows the block diagram of the DOMF in the face recognition system. The wavelet packet transform decomposes the image into frequency subbands to represent the facial features of the face image.

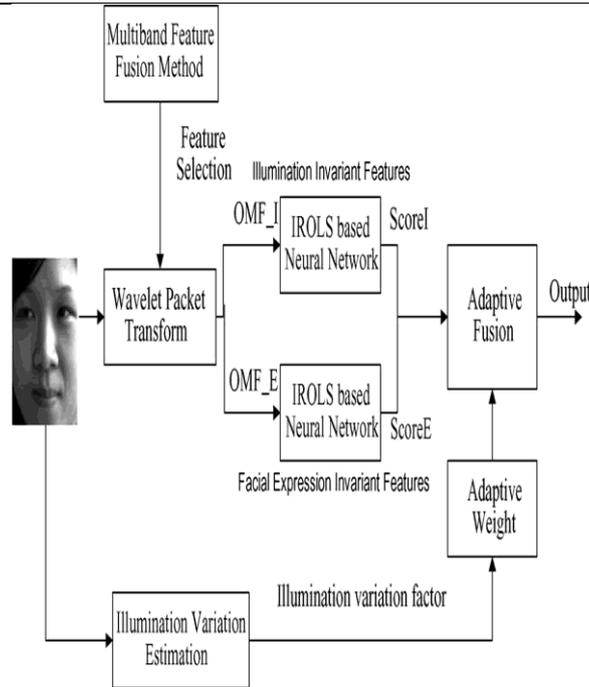


Fig.3. Block diagram of the DOMF for face recognition.

The multiband feature fusion technique is incorporated to search for subbands that are invariant to illumination and facial expression variations separately. The optimal multiband features that are invariant to illumination and facial expression are used and named as optimal multiband feature for illumination (OMF\_I) and optimal multiband feature for facial expression (OMF\_E). The proposed IROLS-based RBF neural networks are then used to classify the OMF\_I and OMF\_E features to generate scoreI and scoreE, respectively. During the testing stage, the decision scores are linearly combined through a set of fusion weights. In the DOMF method, the weights are determined by the illumination variation estimator where the illumination variation factor will be assigned based on the illumination variation level of the input testing image. For example, the weight assigned to scoreI is higher than that assigned to scoreE if the input image is influenced by high illumination variation. The illumination variation estimator uses morphological opening to remove the facial features and then estimates the level of illumination variation of the image. The level of illumination variation can be described as the illumination variation factor  $k$ . Assuming that the face image that contains illumination variation has one side of the image brighter than the other, the  $k$  can be determined as the difference between the mean pixel values at the left and the right sides of the image. The weight of the system can then be determined adaptively based on the  $k$  value of the testing image. The sum rule is incorporated to combine the scores. The sum rule computes the final score from

$$S = \sum_{i=1}^J f_{weight_i} S_i$$

Where  $J$  is the number of modalities (which is two in this case),  $f_{weight}$  is the fusion weight, and  $s_i$  represents the scores obtained from the  $J$  modalities (scoreI or scoreE). The fusion weights are adaptive in the sense that the weights assigned to the modalities are based on the illumination variation factor of the testing image. The weight  $f_{weight_i}$  for each image is determined with the following definition:

$$F_{weight_i} = \begin{cases} F_{weight_i}, & k \geq T \\ 1 - F_{weight_i}, & k < T \end{cases}$$

$T$  denotes the threshold of the illumination variation factor where it is determined as the mean value of the illumination variation factors of the training images. The value of  $f_{weight}$  is fixed and is obtained experimentally. During the training stage, the illumination variation factor of the updated training sample will

be computed to help the incremental learning in the proposed IROLS algorithm. As we have seen in the previous section, the learning of the updated training sample is achieved by reinforcing the old center to the updated training samples as describe. Only the old centers will be moved toward the updated samples based on the following criteria:

$$\eta_{OMF\_I_i} = \begin{cases} \eta, & k \geq T \\ 0, & k < T \end{cases}$$

$$\eta_{OMF\_E_i} = \begin{cases} 0, & k \geq T \\ \eta, & k < T \end{cases}$$

Where  $\eta_{OMF\_I_i}$  and  $\eta_{OMF\_E_i}$  refer to the learning rates for the neural network for OMF\_I and OMF\_E features, respectively. The first criterion explains that, if the updated sample  $i$  has an illumination variation level that is higher than the  $T$  threshold value, the old center which belongs to the class of updated sample  $i$  will be moved toward the updated sample in the network that corresponds to the OMF\_I features. Otherwise, the old center remains stagnant. The same theory will be applied to the network that corresponds to the OMF\_E feature which is shown on the second criterion: The updated samples that do not contain illumination variation only will be learned by the network that corresponds to OMF\_E features.

#### IV. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed IROLS-based RBF neural network with different face databases. The experiments included in this section are as follows.

- 1) Evaluate the recognition performance of the IROLS- based RBF neural network. The recognition accuracy of the IROLS-based RBF neural network is compared with that of the conventional ROLS-based RBF neural network and other incremental learning methods.
- 2) Evaluate the effect of knowledge accumulation of the IROLS-based RBF neural network on the recognition accuracy in various visual variations. The recognition accuracy of the IROLS is compared with that of other face recognition methods

#### Recognition Performance of IROLS-Based RBF Neural Network:

In this experiment, the AR face database is used. The AR face database consists of images that are taken under various visual variations such as occlusion, illumination, facial expression, and aging. Fig. 4 shows some of the example images from the AR database.



Fig. 4. Examples of images from the AR database.

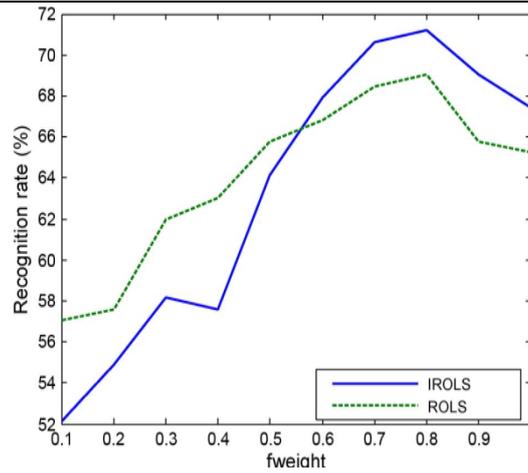


Fig. 5. Recognition rate against fweight for IROLS- and ROLS-based RBF neural network in AR database

Fig. 5 shows the recognition rates of IROLS and ROLS against fweight values. The result shows that both the algorithms achieve the optimal fweight value at 0.8

Table 1 RECOGNITION ACCURACY OF OF IROLSLAMBDAWITH AR DATABASE

| $\lambda$ | Recognition rate (%) |
|-----------|----------------------|
| 0         | 70.9                 |
| 0.01      | 71.6                 |
| 0.1       | 66.0                 |
| 1         | 70.9                 |

The recognition accuracies versus regularization parameter  $\lambda$  with the values of 0, 0.01, 0.1, and 1 are shown in above Table . The  $\lambda = 0.01$  is chosen to be the optimal value as the highest recognition rate is achieved at this point of the  $\lambda$  value. The experiment has been carried out to find the optimal  $\eta$  value in the AR database. We found that the most suitable  $\eta$  value for all the testing samples that we have tested is 0.4.

Table 2 RECOGNITION ACCURACY OF INCREMENTAL LEARNING METHODS WITH AR DATABASE (THE RESULTS ARE OBTAINED BASED ON 12 SIMULATIONS

| Methods         | Recognition rate (%) | Standard Deviation |
|-----------------|----------------------|--------------------|
| IPCA            | 60.5                 | -                  |
| ILDA            | 67.9                 | -                  |
| Online Boosting | 75.2                 | -                  |
| I-ELM           | 56.5                 | 8.7                |
| EI-ELM          | 60.8                 | 9.0                |
| ROLS            | 71.8                 | 7.8                |
| IROLS           | 75.5                 | 5.7                |

The experiment protocol is as follows: 1) Ten classes are randomly selected as the initial model of the neural network, and 2) one hundred thirteen classes are progressively added to the initial model by the IROLS algorithm. The widths of the neural network neurons for OMF\_E and OMF\_I features are respectively . The regularization parameter is 0.01 for both the neural networks. The total number of hidden neurons selected for the neural networks that correspond to OMF\_E and OMF\_I features are 124. The weight fweight for the DOMF is set as 0.8.

Table 2 shows the comparison of recognition rates and standard deviations among the incremental learning methods in different face recognition methods. The recognition rates of the algorithms in [1], [2], and

[4] shown in the table are extracted from [1] directly. The average recognition rates of the ROLS, IROLS, incremental extreme learning machine (I-ELM), and enhanced I-ELM (EI-ELM) are computed in 12 simulations using 12 different training and testing data sets. In Table III, it can be seen that the IROLS achieves 75.5% which is 3.7% higher than the ROLS. This indicates that the recognition performance of the IROLS is comparable to that of the ROLS. The IROLS and online boosting achieve a comparable recognition rate, which is 75.5% and 75.2%, respectively. The third best performing method is the ROLS followed by the ILDA, EI-ELM, IPCA, and I-ELM.

To test the recognition performance of the video based face recognition, the equal-weight fusion approach is used to combine the scores obtained by each frame. In the equal-weight fusion approach, the mean score is used for decision making, which is defined as

$$\hat{S} = (1/F)\sum_{f=1}^F S_f$$

Where  $F$  is the total number of frames which is three in this experiment,  $f$  is the frame number, and  $S_f$  refers to the resultant score for that particular frame. the recognition accuracy of the IROLS- based RBF neural network and eigen face method without up- dated samples. The recognition performance is clearly degraded when the system is tested with testing samples that contain illumination and expression + illumination variations. This is because there is more than one visual variation that appears in the same testing image. On the other hand, the overall recognition performance of both the methods is improved when updated samples are added. The most obvious improvement in the recognition performance can be observed when the system is tested with testing samples that contain illumination and expression + illumination variations. It can be seen that the proposed IROLS-based RBF neural network out- performs Eigen faces in terms of recognition accuracy. Hence, we can conclude that, with the additional information that exists in the updated samples, the proposed IROLS improves the robustness of the face recognition system to testing images that contain more than one visual variation.

## V. CONCLUSION

In this paper, we have proposed the IROLS algorithm for RBF neural networks to solve the problems in the face recognition. The conventional ROLS involves retraining the whole neural network when new training data are added. In our proposed algorithm, the selection of the regressors for the new data is done locally, hence avoiding the expensive reselecting process. In addition, it accumulates previous experience and learns updated new knowledge of the existing groups to in- crease the robustness of the system. The proposed algorithm achieves comparable recognition accuracy and requires lesser training time and hidden neurons compared to the conventional ROLS-based RBF neural network. The experimental results have shown that the proposed method achieves higher recognition accuracy as compared to the IPCA, ILDA, I-ELM, and EI-ELM. It also achieves a comparable recognition rate to the online boosting in the AR database. Moreover, the recognition performance of the proposed method in visual variations is tested on the Yale and UNMC-VIER databases. The results have shown that the proposed method outperforms most of the state-of-the-art face recognition methods in terms of recognition accuracy. We have also shown that improvement in the robustness of the face recognition system to real-world face recognition challenges is achieved by the proposed algorithm.

## REFERENCES

- [1] D. Masip, A. Lapedriza, and J. Vitria, "Boosted online learning for face recognition," *IEEE Trans. Syst., Man, Clyburn. B, Cybern.*, vol. 39, no. 2, pp. 530-538, Apr. 2009.
- [2] M. Artac, M. Jogan, and A. Leonardis, "Incremental PCA for online visual learning and recognition, in *Proc. ICPR*, 2002, vol. 3, pp. 781784.
- [3] S.Ozawa, S. L. Toh, S. Abe, S. Pang, and N. Kasabov, "Incremental learning of feature space and classifier for face recognition", *Neural Netw.*, vol. 18, no. 5/6, pp. 575-584, Jul./Aug. 2005.
- [4] H. Zhao and P. C. Yuen, "Incremental linear discriminant analysis for face recognition," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 38, no. 1, pp. 210-221, Feb. 2008.
- [5] S. Chen, E. S. Chng, and K. Alkadhimi, "ROLS algorithm for constructing radial basis function networks," *Int. J. Control*, vol. 64, no. 5, pp. 829-837, Jul. 1996.
- [6] Y. W. Wong, K. P. Seng, and L.-M. Ang, "Dual optimal multiband feature for face recognition", *Expert Syst. Appl.*, vol. 37, no. 4, pp. 2957-2962, Apr. 2010.