International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.9, (2016), pp.253-264
http://dx.doi.org/10.14257/ijsip.2016.9.9.24

Hexagonal Descriptor Particle Swarm Optimization with Knowledge-Crowding for Face Recognition

Taqdir[1] and Renu Dhir[2]

[1]Computer Science and Engineering Department, G.N.D.U.,Regional Campus, Gurdaspur-143521, Punjab,India
[2]Computer Science and Engineering Department, NIT, Jalandhar–144001 Punjab, India
[1]taqdir_8@rediffmail.com, [2]dhir@nitj.ac.in

Abstract

Automatic identification of various facial expressions with high recognition value is important for human computer interaction as the facial behavior of a human can be treated as an important factor for information representation as well as communication. A number of basic factors such as cluttered background, occlusion, and camera movement and illumination variations degrade the image quality resulting in poor performance for identifying different facial expressions. Moreover the identification of the automatic feature detection in facial behavior requires high degree of correlation between the training and test images. Face recognition is done by minimizing the objective function which leads to selection of optimal set of fiducial points. The method preserves the local information from different facial views for mapping neighboring input to its corresponding output, resulting in low dimensional representation for encoding the relationships of the data. The proposed method Hexagonal Descriptor Particle Swarm Optimization with Knowledge-Crowding (HDPSO-KC) overcomes from local optima and improves global search process and collaborative work. The method also covers the problem of eliminating the particles in denser regions in Pareto front distribution. The proposed methodology is validated with benchmark datasets for analyzing the performance over other methods.

Keywords: *Face recognition, Descriptors, optimization, Facial Views, Local Polynomial Approximation*

1. Introduction

Face recognition has emerged as an area of interest over past few years. In the broader area of application such as surveillance systems, identification system, human computer interaction and many more led the researchers to focus in this domain [1]. The selection of the best features results in proper face recognition and is independent of the arbitrary environmental variations such as pose, scale, illumination and facial expressions. A human face is a dynamic object having variation in appearance, pose and orientation that makes the face identification problem more difficult. Faces have high variability and are dynamic in nature. A number of conditions such as orientation, lighting, scale and expressions [2] variates over period of time under circumstances become problematic in some cases. Face recognition is done in two approaches holistic approach and feature-based approach. Features obtained for holistic approach are global features from faces whereas in feature-based approach the features are obtained from the local features from the face.

ISSN: 2005-4254 IJSIP

1.1 Overview

The extract of local features from the image gives the information of the important features face recognition. The basic descriptor is affected by the noise and more affected when square image coordinates are considered that eliminates the low contrast features. A hexagonal descriptor results in sharp and prominent edges. From the obtained set of features, the optimal features are computed with the application of MOPSO using shared learning and crowding distance. The proposed method Hexagonal Descriptor Particle Swarm Optimization with Knowledge-Crowding (HDPSO-KC) overcomes from local optima and improves global search process and collaborative work. The method also covers the problem of eliminating the particles in denser regions in Pareto front distribution. We formulate a local neighbor structure (LNS) which is a transformation of training set for compressed representation. A dissimilarity matrix is computed for identifying the most correlated face image with the obtained LNS. Finally the most recognized face is obtained based on higher similarity value. We have extensively evaluated the performance of our method with different face images under variability such as lightening conditions, orientations, different facial expression and occlusions. Experiments have been done for comparing the performance of the proposed method with existing methods based on various measures such precision, recall, confusion matrix and recognition rate. Different datasets namely, Yale Face Database, Extended Yale Face Database, ORL database, has been considered for our experiments. Results show that our method performs significantly well with a better recognition accuracy.

1.2. Organization of Paper

The remaining part of the paper is organized as follows. Section 2 presents a review state-of-art literature. In Section 3 the proposed approach is detailed with algorithmic steps and flowchart. Section 4 deals with the experimental results and discusses the outcomes considering the performance of the proposed method and Section 5 draws the conclusion.

2. Related Work

The analytical technique Active Shape Models (ASM) [3] has massively contributed for the face recognition. The presence of artifacts such as noise and illumination variations distorts the edges. Despite of better landmark location accuracy the exact facial feature is challenging. The obtained feature are high dimensional that are prone to over fitting making the classification task difficult. To address the problem of proper feature extraction, Pareto based MOGA, variation of the classical GA is considered that ranks the individual chromosome according chromosomes are proposed. However, such a method suffers from the problem of diversity preservation. In 2013 Vignolo *et. al.,* [3], propose a method that wraps three methods for face recognition. In the first wrapper, ASM selects the particular feature from the input image. The second wrapper does MOGA with an objective function that combines features and classification accuracy. The third wrapper implements the same MOGA function with another alternative. The KNN classifier is deployed for recognition the face from the given class of face images.

In 2014 Gaidhane *et. al.,* [4] proposed a simple method for face recognition. The method illuminates the computation of eigen values and eigenvectors which reduces the complexity. For set of training images, a polynomial characteristic is computed that obtains the features for each image. These features of training image are stored in a matrix form known as companion matrix. Similar procedure is repeated for a test image. A sequence of vectors for both the test image matrix and training image matrix is computed. The determinant of the symmetric matrix is calculated for finding the singularity. If the matrix is singular the test image is similar to training set else no common features are detected. The maximum value of nullity of matrix shows recognition of unknown face.

The method reduces the computation complexity and the problem of high dimensionality of data is resolved but requires more number of steps for computations.

Mehta *et. al.,* [5] proposed an Optimal Directional Face (ODF) recognition method that extracts the directional information using the image directive. An image derivative enhances the facial features that have characteristics information. The method performs image derivative with Local Polynomial Approximation (LPA) which is a directional filter at multiple scales. As the directional features captures increases the dimensionality, hence the Intersection of Confidence Interval (ICI) is applied for selection of scale at each pixel. The textural features are extracted by applying modified Local Binary Pattern followed by partitioning the feature image and finally the histogram of each feature values are computed and concatenated together. After portioning the ODF at different level, the LDA is applied for reducing in dimensionality; finally the SVM is applied for classification.

Seo *et. al.,* [6] developed a method that is robust to partial variations by extracting the features through Scale Invariant Feature Transform (SIFT). The method is test on different bench mark datasets with facial image expression and occlusions. Facial image is represented with local features using SIFT that uses scale-space difference of Gaussian to detect key points in the image. The SIFT is applied on set of training image and test image and estimation of probability density function is performed. With the density function weight of each descriptor smaller weight value are used to represent occluded area. Once the weight for feature values for test image are computed, the K-NN classifier is applied for finding the similarity and assign the test image to a class. However, the method when applied for occluded images and partial variations results in high dimensionality and improper results with local variations.

In 2011 Sparse Representation gained a lot of attention for representing the training samples with sparse linear combination. The sample selection done for representation helps to find the relation between the query image and training samples. This relation is obtained from the correlation structure. The sparse representation is mainly focused on the sparsity rather than the correlation that results in loss of information. The proposed approach Adaptive Sparse Representation based Classification (ASRC) [7] tries to capture trace norm for representation done in vector form for better correlation. In the training set, the samples are rarely correlated, the algorithm ASRC behaves like Sparse Representation based Classification (SRC) [8] whereas when the samples are highly correlated the Correlation based Classification is performed. The proposed method properly handles pixel corruption along with information loss when applied with correlated samples.

Most of the face recognition methods based on deep neural networks fail to represent the face identities for training. The approach leads to inefficiency and improper classification. The proposed method is based on deep convolution network [9]. A triplet selection is done for representing verification, recognition and clustering. For generating triplet, exemplar mining strategy is performed along with mining techniques for clustering. The face recognition method faces many problems such as pose variation with illumination and expression. The proposed method addresses the occlusion for facial texture. The gallery of images is transformed into a discriminative subspace with a learning method by finding the correlation between the matrices of different poses known as Multi-Task Feature Transformation Learning (MtFL) [10]. This approach has limited memory size and has advantage when training set is limited. MTL is learning technique that captures the correlation among different task improving generalization with increase in sample size. The pose is detected by identifying the occluded facial texture based on the contour of 3D model projected to 2D image plane. The edges are detected by canny operator followed by registration of point sets called Coherent Point Drift (CPD). For feature transformation, the images from different poses are differenced in transformation and find the correlation among different feature transformations.

3. Proposed Work

Hexagonal Scale Invariant Feature Transform (H-SIFT) is applied as an important feature descriptor for object recognition. From the input image, SIFT extracts a large set of local feature vectors which are invariant to translation, scale, geometrical distortions and illumination changes. Due to its spatially localized nature, SIFT features are less affected by noise and corrupted pixels. These feature set mostly applied on square image coordinates provide less amount of features across edges, eliminating low contrast pixels containing discriminate facial features.

On the other hand, hexagonal image coordinates have an inherit advantage of providing sharp edges and makes the discriminate features along edges more prominent. In H-SIFT process, the input square image is converted into hexagonal shaped image pixels, designing SIFT feature set into its hexagonal form. In this process, the input square pixels form a rectangular grid by shifting half the width of pixels *i.e.*, every alternative pixel row is either left or right shifted by half of its width, resulting in enhancement of the image feature set, makes the edges more prominent for identification.

The mathematical representation of hexagonal form is represented as in Equation (1).

$$H(x,y)=\sum_{p=0}^{p-1}\sum_{q=0}^{q-1}S(p,q)h(x-p,y-q) \tag{1}$$

Where, $S(p,q)$ be square image with the pixels $p\times q$, h is symbolize kernel for interpolating the square image into its hexagonal form. The symbols (p,q) and (x,y) are associated with pixels of square and converted hexagonal image respectively. Gaussian function G(x,y,α) is initiated in this process with hexagonal image H(x,y) as a convolution operator. The resultant convolved image is represented by the function C(x,y,α),

$$C(x,y,\alpha)=G(x,y,\alpha)*H(x,y) \tag{2}$$

The symbol α in Equation (2) represents the convolution operation results in regular blurring of the image in different scales and α is the scale factor. The Gaussian function is represented as:

$$G(x,y,\alpha)=\frac{1}{2\pi\alpha^2}e^{-(x^2+y^2)}/2\alpha^2 \quad . \tag{3}$$

This process eliminates noisy pixels contained in image through regular blurring makes advantageous of hexagonal image processing by enabling the enhancement of edges as well as highlighting the areas with low contrast and poor edge response. Figure 1 shows the brief methodology of proposed algorithm.

Particle Swarm Optimization is a stochastic meta-heuristic method based on mutual coordination and information distribution among particles. In the process, the particles rely on the information of local best particle and global best particle, reducing the diversity of the particle swarm through ignoring the other particles information, results in struck to the local optimum of the complex optimization problems. To overcome these shortcomings, the average best position (S) is calculated by the following formula:

$$S(t)=\frac{1}{M}\sum_{i=1}^{M}P_i(t) \tag{4}$$

Where, each particle has a position Xi(t) = [Xi,1(t),Xi,2(t),...,Xi,D(t)] and velocity Vi(t) = [Vi,1(t),Vi,2(t),...,Vi,D(t)] in a D-dimensional search space. During the movement, the local best position for each particle is recorded as Pi(t) = [Pi,1(t),Pi,2(t),...,Pi,D(t)], and the global best particle of the swarm is to be G(t) = [G1(t),G2(t),...,GD(t)].

The velocity and position of the particle are updated in every iteration by the following equations:

$$V_{ij}(t+1) = wV_{ij}(t) + c_1 r_1 (P_{ij}(t) - X_{ij}(t)) + c_2 r_2 (G_j(t) - X_{ij}(t)) + c_3 r_3 (S_j(t) - X_{ij}(t)) \tag{5}$$

$$X_{ij}(t+1) = X_{ij}(t) + V_{ij}(t+1) \tag{6}$$

Where, i represents the ith particle of the swarm, j represents the jth dimension in the search space, t is the number of current iteration, w is the inertia weight controls the impact of the previous velocity on the current velocity, c1, c2, c3 are the acceleration coefficients that represents the dynamically changeable weight of the average best position in order to increase the chance of not getting stuck at the local optimum, r1, r2, r3 ∈ [0-1] are uniformly distributed random variables.

During the iteration process, the Pareto dominance relationship between the current particle and local best particle is compared and updated with the best value. The mutation rate is varying with the increasing iteration number results in escape from not receiving stuck at the local optimum in complex optimization problem. The greedy approach in addition can ensure the fast convergence of the algorithm.

A dynamic crowding distance strategy is maintained in order to overcome the limitation of the traditional approaches which deletes lots of particles in dense regions, damage the distribution of the Pareto front.

In the decision step, from the obtained optimized feature vector, NDD is computed which is a pixel based approach for finding less-dissimilar faces thereby reducing the number of comparisons. It computes the L2 norm between the test images and obtained LNS of each training set as given in Equation 2. The LNS of the training set selects the relevant face images, increasing the chances of obtaining the exact match.

$$d^2 E = \sum_{i,j=1}^{MN} g_{ij} \left(x^i - y^i\right)\left(x^j - y^j\right) \tag{7}$$

M and N are dimensions of images and gij is a metric matrix with orientation

Here we have a Sparse Dissimilarity (SD) computation that groups the similar test images based on the distance value from the LNS of each training set as shown in Figure 2. This grouping is one for identifying similar persons face image that validates the decision of known or unknown face. The dissimilarity is formulates as given in Equation 2.

$$Sparse \quad Dissimlarity \quad T_m = \sum_{i=1}^{m} \sum_{j=1}^{n} dist\left(T_i, LNS_j\right) \tag{8}$$

Where m = number of test images, n = training sets

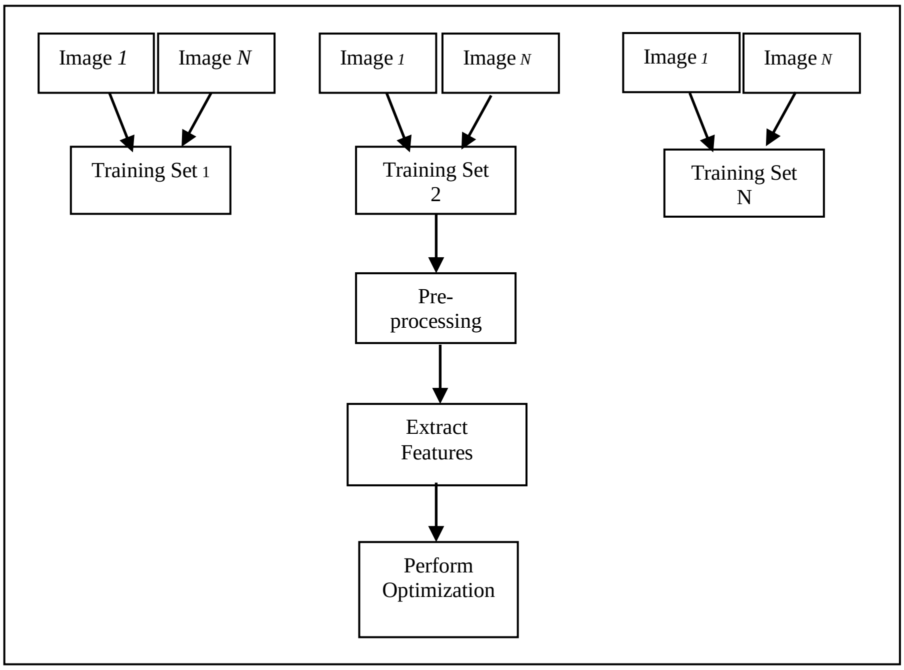

Figure 1. Pre-Processing and Optimization

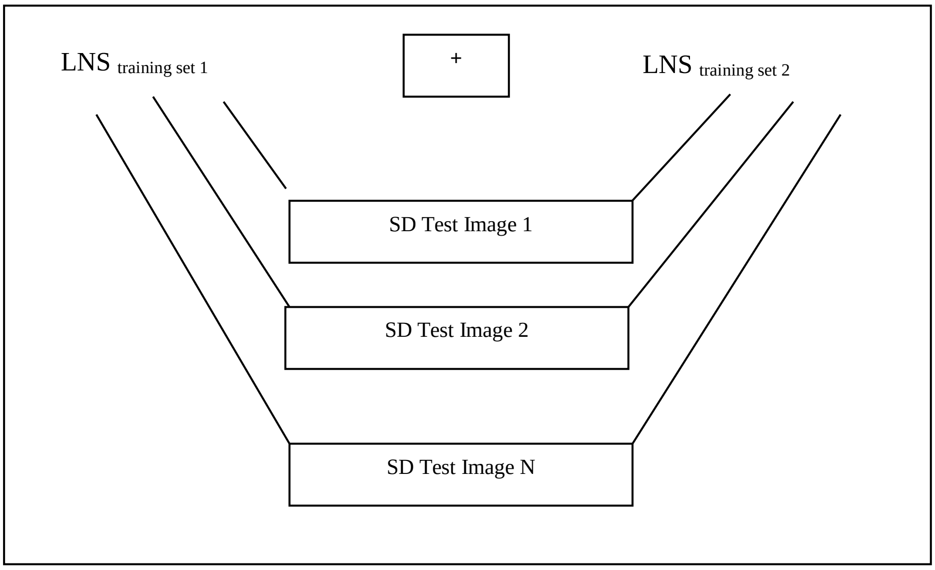

Figure 2. Sparse Dissimilarity for 2 Training Set

After obtaining the SD value of test images a Normalized Max-Min Sparse Dissimilarity is computed for that normalizes the value to a range for finding the correlated face images. The correlation is minimum difference value in the comparison matrix of the test image dissimilarity value. However, the difference value is threshold certain conditions where the test image is an unknown face. The method of SD has a major contribution in grouping and identifying the faces in correct training set, correct training set with same expression given the test image, correct training set with invalid expression given the test image and incorrect training set with invalid expression. The L2 norm value of training set LNS and test image abet identification of the training set for respective test image. The lower is the L2 value of test image, higher is belongingness to the training set.

The final recognition phase is the Decisive Feature based Similarity (DFS), the obtained optimum features of the test set is correlated with the obtained optimum features of the training set corresponding to the respective LNS as given in Figure 3. The recognition result identifies the exact image of the person with valid expression given the expression exists in training set otherwise it outputs a nearby-value face. The method also segregates an unknown face when the person does not exist in the dataset. The recognition rate for different cases is shown in Section 5.

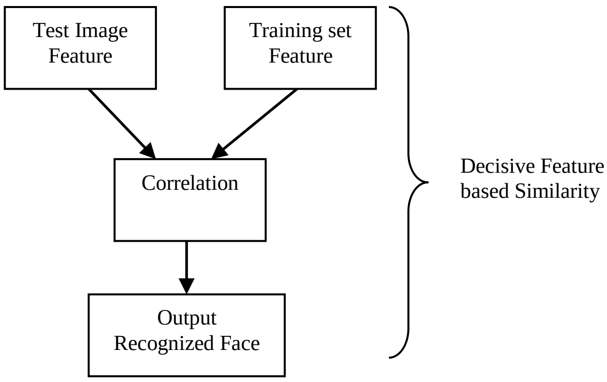

Figure 3. Decisive Feature Based Similarity

Algorithm:

1. Read the Image
2. Pre-process the image for noise removal
3. Perform hexagonal descriptor for finding the feature set

A. Initialization Step

4. Initialize MaxItr- maximum number of iterations
5. Initialize the PSO parameter- c1, c2, r1,r2, w
6. Initialize N – total number of population

B. MOPSO Computation

7. While (i<MaxItr)
8. Do
9. While (j<N)
10. Select the initial particle position
11. Update the velocity and position
12. Evaluate the objective function
13. Update the best particle position-Pbest
14. Perform mutation by crowding
15. Check for boundary encapsulation
16. Update the best particle position-Pbest

End

C. Decision Step

17. LNS Tr = Avg (Tm)
18. For each Ti
19. NDDvec = L2 (Ti, LNSTr)

$$SD = \sum_{i=1}^{m} \sum_{j=1}^{n} dist\left(T_i, LNS_j\right)$$

20. Obtain the group in Ti
21. Find the training set SELTr
22. DFSvec = (Ti (feature set),SELTr)
23. Check Ti for different cases
24. Output the recognized face with recognition rate R

4. Results and Analysis

For analyzing the recognition accuracy and efficiency, the proposed method has been experimented on a number of publicly available face databases such as the Yale Database, Extended Yale Face Database and the ORL database.

The ORL database [26] contains 10 different images of 40 distinct individuals. For some of the individuals, the images were captured at the different intervals of time, varying the poses, lighting, facial expressions (smiling/ non-smiling, open or closed eyes) along with or without glasses. All the images were taken against a dark homogeneous background with the individual upright frontal position. The Yale database contains 165 gray scale images of 15 different persons. There are 11 images of each person having different facial expression under various conditions (center-light, left-light, right-light, with glasses, without glasses, normal, happy, sad, sleepy, surprised and wink). The Extended Yale database [27] contains 5760 single light source images of 10 persons, 9 different poses and 64 illumination conditions of each person. The face images vary with facial expressions and illumination. The images have normal, sleepy, sad and surprising expressions. There are some images with or without glasses. These images are captured by varying the position of light source at the center, left or right.

The Yale database is a standard face recognition database. For this experiment we have considered images of 3 persons with 9 different expressions as training set. The test set comprises of 6 different person images excluding the training set from Yale database. The experiment initially finds the feature set of matrix dimension 59x2 for training set and 69x2 for test set. The method optimizes the selected set of features that reduces the dimensionality, increasing the accuracy rate thereby reducing the computation cost. The optimization method ranks the solution obtained in the non-domination front which includes those selected features that are prominent for decision. The optimized feature set reduces its dimensionality to 20x2 and 26x2 for training set and test set respectively. The classification step performs LNS that computes the Sparse Dissimilarity for given test image against the training set validating known or unknown face image. The minimum dissimilarity value selects the training set for given test image. The optimized features set of the test image 26x2 is computed against the LNS of the training set. The DFS outputs the image matrix with recognition value showing the true positive, false positive and true negative results.

- The true positive (TP) illustrates the person exists in the database with valid expression present in the training set.
- The false positive (FP) illustrates person with invalid expression with respect to the person present in the training set.
- The true negative (TN) illustrates invalid expression of unknown person not present in the entire database.

The experiment on the ORL database comprises the training set of 4 persons with different expressions and various poses. The test set comprises of 9 images of 6 different persons. The algorithm selects the set of features that are optimized for finding optimal feature set in the training as well as test set. The LNS for training set is computed for finding the grouping among the test set images and finding the relevant training set for given test image. The DFS is computed for finding recognition rate from the selected training set and the optimum features in test image.

The comparison of objective functions with proposed technique has been shown in Figure 4. The overall performance of proposed technique on the basis of recognition rate has been shown in Figure 5.

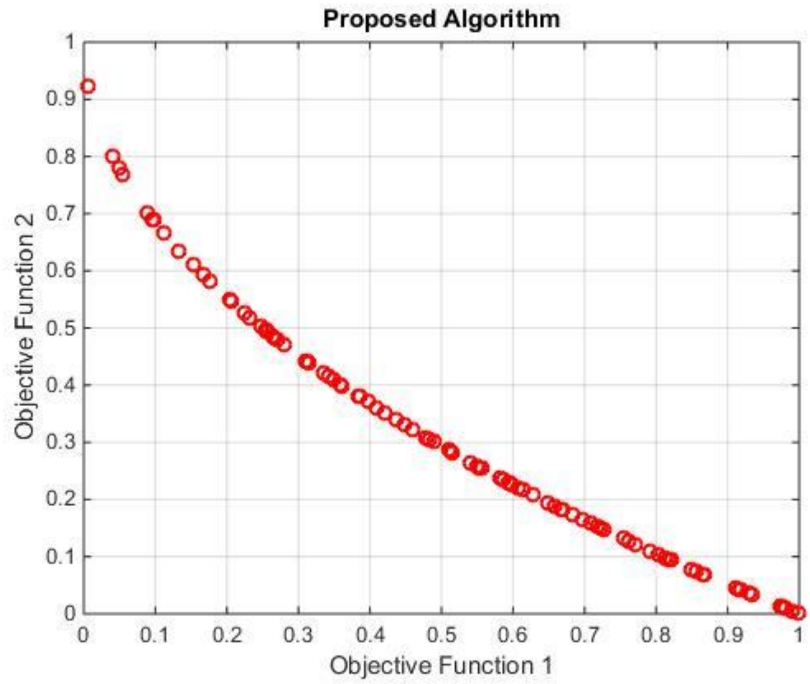

Figure 4. Objective Function 1 Vs Objective Function 2 of Proposed Algorithm

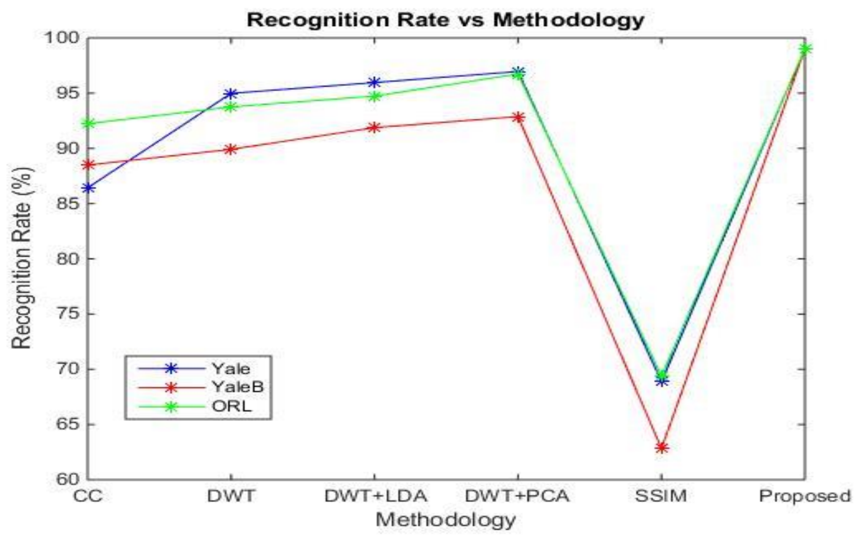

Figure 5. Recognition Rate Vs Methodology

5. Conclusion

In this paper we have studied a number of methods for face recognition. The proposed method tries to address the problem of illumination variation and pose for increasing recognition accuracy. The MOPSO approach with knowledge based crowding updates the value from the previous iteration and mutates the population based on the greedy strategy that results in selection optimal set of features. These features are classified with decision based classification based on LNS of the features of the training image. The performance measures shows that the proposed method obtains better accuracy than the conventional methods.

Acknowledgments

The authors are greatly thankful to the Department of Computer Science & Engineering, G.N.D.U., Regional Campus, Gurdaspur, India and Department of Computer Science & Engineering, in NIT, Jalandhar, India for providing excellent lab facilities that make this work possible.

References

[1] W. R. Schwartz, H. Guo, J. Choi and L. S. Davis, "Face Identification Using Large Feature Sets", IEEE Trans. On Image Processing, vol. 21, no. 4, **(2012)**, pp. 2245-2255.
[2] X. Luan, B. Fang, L. Liu, W. Yang and J. Qian, "Extracting sparse error of robust PCA for face recognition in presence of varying illumination and occlusion", Pattern Recognition, vol. 47, **(2014)**, pp. 495-508.
[3] L. D. Vignolo, D. H. Milone and J. Scharcanski, "Feature Selection for Face Recognition Based on Multi-Objective Evolutionary Wrappers", Expert Systems with Applications, vol. 40, no. 13, **(2013)**, pp. 5077-5084.
[4] V. H. Gaidhane, Y. V. Hote and V. Singh, "An efficient approach for face recognition based on common eigenvalues", Pattern Recognition, vol. 47, **(2014)**, pp. 1869-1879.
[5] R. Mehta, J. Yuan and K. Egiazarian, "Face recognition using scale-adaptive directional and textural features", Pattern Recognition, vol. 47, **(2014)**, pp. 1846-1858.
[6] J. Seo and H. Park, "Robust recognition of face with partial variations using local features and statistical learning", Neurocomputing, vol. 129, **(2014)**, pp. 41-49.
[7] J. Wang, C. Lu, M. Wang, S. Yan and X. Hu, "Robust Face Recognition via Adaptive Sparse Representation", Computer Vision and Pattern Recognition, **(2014)**, pp. 1-11.
[8] R. Rigamonti, M. Brown and V. Lepetit, "Are sparse representations really relevant for image classification?", In Proceeding of IEEE Conference on Computer Vision and Pattern Recognition, **(2011)**, pp. 1545–1552.
[9] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", Computer Vision and Pattern Recognition, **(2015)**, pp. 815-823.
[10] C. Ding, C. Xu and D. Tao, "Multi-task Pose-Invariant Face Recognition", IEEE Transactions On Image Processing, vol. 24, no. 3, **(2015)**, pp. 980-993.
[11] A. Azeem, M. Sharif, J. H. Shah and M. Raza, "Hexagonal scale invariant feature transform (H-SIFT) for facial feature extraction", Journal of Applied Research and Technology, **(2015)**, pp. 402-408.
[12] B. Xue, M. Zhang and W. N. Browne, "Multi-Objective Particle Swarm Optimization (PSO) for Feature Selection", In Proceedings of the 14th annual conference on Genetic and evolutionary computation, **(2012)**, pp. 81-88.
[13] B. Xue, A. K. Qin and M. Zhang, "An Archive Based Particle Swarm Optimization for Feature Selection in Classification", IEEE Conference on Evolutionary Computation, **(2014)**, pp. 3119 - 3126.
[14] B. Xue, M. Zhang and W. N. Browne, "Particle Swarm Optimization for Feature Selection in Classification: A Multi-Objective Approach", IEEE Transactions on Cybernetics, **(2012)**, pp. 1-16.
[15] The ORL Database of Faces, AT&T Research Laboratories, Cambridge, [Online]. Available: <http://www.cam.ac.uk/Research/DTG/attarchive/pub/data/attfaces>.tar.Z.
[16] The Yale Face Database B [Online]. Available http://cvc.yallefacesB.html.

Authors

Taqdir, she received her B.Tech degree in Computer Science & Engineering from Sant Longowal Institute of Engg. and Technology, Longowal, India in 1999. She received M.Tech degree in 2005. Currently she is working as Assistant Professor in Department of Computer Science & Engineering, G.N.D.U., Regional Campus, Gurdaspur, India and also perusing PhD in the area of Image Processing.

Dr. Renu Dhir, **s**he is working as a Associate Professor in Department of Computer Science & Engineering, in NIT, Jalandhar, India. She received Ph.D. degree from Punjabi University Patiala, India. She has expertise in Machine Learning, Natural Language Processing and Digital Image Processing. She has more than 30 publications in international journals and conferences.