Intelligent Face Recognition based on Manifold Learning and Genetic-Chaos Algorithm Optimized Kernel Extreme Learning Machine

Wei He[1], Enjun Wang[2], and Ting Xiong[3]

[1] Transportation Engineering Institute of Minjiang University, Fujian 350108, China
[2] Transportation Research Center, Wuhan Institute of Technology, Wuhan 430073, China
[3] School of Power and Energy Engineering, Wuhan University of Technology, Wuhan 430063, China
Email: hewei1982_mju@163.com; wangenjun@163.com; xiongting163@yeah.net

Abstract—In order to extract sensitive features of face images from high dimensional image data and facilitate the recognition speed, this paper has proposed a novel method based on the manifold learning and genetic-chaos algorithm optimized kernel extreme learning machine (KELM) for the application of face recognition. The locally linear embedding (LLE) algorithm has been employed to extract distinct features by projecting the original high dimensionality of the face image into a low dimensionality space. Then the KELM is introduced to provide quick and accurate pattern recognition on the extracted features. The only parameter need to be determined in KELM is the neuron number of hidden layer. Literature review indicates that very limited work has addressed the optimization of this parameter. Hence, the genetic-chaos algorithm was used for the first time to optimize the KELM parameter in this paper. A robust KELM structure may be attained after the genetic-chaos optimization. In order to evaluate and verify the proposed method, experiment tests have been carried out using standard face expressions. The experimental analysis results indicate that the performance of the proposed LLE- genetic-chaos-KELM method outperforms its rivals in terms of both recognition accuracy and training speed.

Index Terms—face recognition, feature extraction, GA Chaos, KELM.

I. INTRODUCTION

Face expression recognition has found its application in a variant range. It is very useful in the security systems, criminal identification, image and film processing, etc [1]. However, the face image usually has very high dimensionality in digital data and therefore it is difficult to extract sensitive features from such high dimensionality. It will be useful if the face image can be processed in a low dimensional manifold space for distinct feature extraction [2]. By doing so, the precision and speed of the face recognition could be enhanced significantly [3]. This issue has been and is still received considerable attentions in the research and practice of face recognition. How to develop quick and accurate methods for the face recognition

attracts many researchers and institutes. To achieve this goal, a possible solution is to reduce the dimensionality of the image data into a low dimensional space to extract distinct features. Some useful dimension reduction techniques have been proposed for this application in face identification. The most famous one is the principal component analysis (PCA). PCA has been introduced into the face recognition [4] by many scholars and lots of articles have been investigated the functional effect of PCA in the feature extraction of face images. These reports have show that the PCA is very powerful for image pre-processing by extract some distinct features of the original image data into a low dimensional space; however, PCA has its bottlenecks. The main limitation of PCA is the extraction of nonlinear properties of the original data [5]. This is because PCA is built up on the assumption of linear model. The same problems are also found in other linear methods [6], including multi-dimensional scaling (MDS) and linear discriminate analysis (LDA). Fortunately, the manifold learning algorithms provide a new means to deal with the nonlinear dimensionality reduction problems. The Isomap [7] and locally linear embedding (LLE) [6] etc., are able to deal with the underlying nonlinear behavior of the data. Compared with the linear methods, the purpose of manifold learning is to project the original high-dimensional data into a lower dimensional feature space by preserving the local topology of the original data [8]. Thus, useful face expression features can be obtained in the low dimensional space. The remaining problem is how to identify the face images from these features.

Up to date, the artificial intelligent methodologies have been proven to be very useful in the face image recognition. Two common techniques are the artificial neural networks (ANNs) and support vector machines (SVMs). Popular ANNs, such as BP NN, have strong ability to adaptively learn the inherent patterns hidden in the given data; however, they often suffer from local minima and slow convergence speed [9]. For a SVM, it needs to set the kernel function, error control parameters, and penalty coefficient. It is difficult to select these parameters. Hence, although ANNs and SVMs are very useful and effective in machine version and intelligent computation [9], they have some bottlenecks involved

with slow learning speed and poor learning scalability. These shortcomings constrain their applications in industry. In order to enhance the application of artificial intelligent methodologies, the kernel extreme learning machine (KELM) has been proposed as en integration of ANN and SVM to provide quick and accurate pattern recognition ability [10]. The KELM has the features of the LS-SVM and the structure of single-hidden layer feed-forward networks (SLFNs) in the same time. Different from classical ANNs and SVMs, who require setting several parameters in the learning processing, the KELM only needs the number of hidden layer nodes of the network to be specified [11]. Siddartha et al [3] employed the extreme learning machine (ELM) to extract features in face recognition. They found that the ELM is more sensitive on the image features than the BP NN and SVM. Zong and Huang [12] introduced the ELM into the face recognition. Experimental tests have been carried out in the study to show high performance of the ELM in multi-label face recognition applications. Further, they [10] have presented the KELM in the face recognition and found that the KELM outperforms LS-SVM in terms of both recognition prediction accuracy and training speed. However, how to optimize the number of hidden layer nodes of the KELM has not mentioned in their work. If the number of hidden layer nodes of the KELM could be optimized, the training speed and pattern recognition accuracy can be improved [13]. It is therefore crucial to develop an optimization mechanism for the KELM.

In view of genetic algorithm (GA) having good global search capability, much work has used GA as an optimization tool for ANNs and other fields to get high efficiency of the system [14]-[16]. However, GA will sometimes suffer from the phenomenon of premature convergence [17]-[19], which decreases the performance of the GA searching.

To address the mentioned issue, this work presents a new method based on the LLE and GA-chaos optimized KELM to enhance the face recognition. The LLE is used to reduce the original high dimensional image data into a low dimensional space to improve the feature extraction efficiency. Then the KELM is introduced to learn important information about the extracted features. The GA herein is adopted to optimize the neuron number of the KELM to make the KELM achieve high generalization performance. In order to eliminate GA precocious defects, the chaotic technology is adopted in this paper to improve GA optimization process and avoid the above early-maturing problem. A series of experimental tests have been carried out to evaluate the performance of the newly proposed method in face recognition. The recognition performance of the proposed KELM based method has been compared with ELM and SVM based method in terms of both recognition prediction accuracy and training speed.

This paper is organized as following. In Section 2 the theories of the proposed method are described.

Experiments and discussions are presented in Section 3. Section 4 draws the conclusions.

II. THE PROPOSED FACE RECOGNITION METHOD

In order to improve the performance of the face recognition, this work proposes the LLE-GA Chaos-KELM method for the face recognition. The feature extraction of the face images is carried out in a low manifold space through the transform of LLE and then the pattern recognition is implemented by the GA Chaos optimized KELM model. The theories about LLE and GA Chaos-KELM are briefly described as follows.

A. Locally Linear Embedding (LLE)

Given a nonlinear high-dimensional dataset $S = [\mathbf{s}_1 \quad \mathbf{s}_2 \quad \cdots \quad \mathbf{s}_l] \in R^p$, where l is the total sample number and p the dimensionality of each sample, the objective of ALLE is to reconstruct a nonlinear mapping to project S into a reduced manifold space $S_r = [\mathbf{s}_{r1} \quad \mathbf{s}_{r2} \quad \cdots \quad \mathbf{s}_{rl}] \in R^q$ $(q \ll p)$. It takes mainly three steps for ALLE to realize the reduction [7].

Step 1: Compute k neighbours of every sample.

Step 2: Compute the local reconstruction weight matrix **W** by minimizing the following cost function:

$$\min \varepsilon(\mathbf{W}) = \sum_{i=1}^{l} \left| \sum_{j=1}^{k} w_j^i (\mathbf{s}_i - \mathbf{s}_{ij}) \right|^2 , \quad (1)$$

where k is the number of nearest neighbours used for reconstructing each data point and w_j^i is the weight values. If $\mathbf{s}_i$ and $\mathbf{s}_j$ are not neighbours, $w_j^i = 0$ and $\sum_{j=1}^{k} w_j^i = 1$. The local covariance matrix $\mathbf{Q}_{j\alpha}^i \in R^{k \times k}$ is introduced to calculate the weight values [7], and

$$\mathbf{Q}_{j\alpha}^i = (\mathbf{s}_i - \mathbf{s}_{ij})^T (\mathbf{s}_i - \mathbf{s}_{i\alpha}) , \quad (2)$$

where $\mathbf{s}_{ij}$ and $\mathbf{s}_{i\alpha}$ are the neighbors of $\mathbf{s}_i$. Hence, by the means of Lagrange multiplier method, the local reconstruction weight matrix can be obtained as:

$$w_j^i = \frac{\sum_{\alpha=1}^{k} (Q^i)_{j\alpha}^{-1}}{\sum_{\beta=1}^{k} \sum_{\gamma=1}^{k} (Q^i)_{\beta\gamma}^{-1}} . \quad (3)$$

Step 3: Map the original dataset to the embedded coordinates. Compute the reconstructed q-dimensional manifold space $\mathbf{s_r}$ by minimizing the following constraint:

$$\min \varepsilon(\mathbf{s_r}) = \sum_{i=1}^{l} \left| \mathbf{s}_{ri} - \sum_{j=1}^{k} w_j^i \mathbf{s}_{rij} \right|^2 , \quad (4)$$

where, $\mathbf{s}_{ri}$ is the projection vector of $\mathbf{s}_{fi}$ in the embedded coordinates, and $\mathbf{s}_{rij}$ are the neighbours of $\mathbf{s}_{ri}$. (4) can be rewritten as

$$\min \varepsilon(\mathbf{s_r}) = \sum_{i=1}^{l} \sum_{j=1}^{k} m_j^i \mathbf{s}_{ri}^T \mathbf{s}_{rj} = tr(\mathbf{s_r} \mathbf{M} \mathbf{s_r}^T), \qquad (5)$$

where the cost matrix M can be expressed as

$$\mathbf{M} = (\mathbf{I}_{l \times l} - \mathbf{W})^T (\mathbf{I}_{l \times l} - \mathbf{W}). \qquad (6)$$

Hence, the minimization of (6) can be reduced to an eigenvalue problem, and $\mathbf{s_r}$ could be determined by the q smallest nonzero eigenvectors of $\mathbf{M}$.

B. KELM

Given samples $\{(x_i, t_i) : i = 1, 2, ..., N; x_i \in R^p, t_i \in R^q\}$, where x is the feature vector and t is the class label vector, the below SLFN is used to identify the sample [8]

$$\sum_{i=1}^{k} \beta_i \, g(\alpha_i^T x_j - b_i) = o_i, \, j=1, 2, ..., N. \qquad (7)$$

where, k is the number of hidden neuron; o_i is the output of jth sample; $g(\cdot)$ is the activation function; b_i is the threshold of the ith hidden neuron; α_i and β_i are the input and output weight vectors, respectively. If the output o can approximate t, we derive

$$\sum_{i=1}^{k} \beta_i \, g(\alpha_i^T x_j - b_i) = o_j = t_j, \, j=1, 2, ..., N. \qquad (8)$$

(8) can be written compactly as

$$\mathbf{G\beta} = \mathbf{T}, \qquad (9)$$

where,

$$\mathbf{G} = \begin{bmatrix} g(\partial_1^T x_1 - b_1) & \cdots & g(\partial_k^T x_1 - b_k) \\ \vdots & \cdots & \vdots \\ g(\partial_1^T x_N - b_1) & \cdots & g(\partial_k^T x_N - b_k) \end{bmatrix},$$

$$\mathbf{\beta} = [\beta_1, \cdots, \beta_k]^T \text{ and } \mathbf{T} = [t_1, \cdots, t_N]^T.$$

To solve (9), the ELM adopts a least squares error to get solution $\hat{\mathbf{\beta}}$:

$$\hat{\mathbf{\beta}} = \mathbf{G}^{\dagger} \mathbf{T}, \qquad (10)$$

where, $\mathbf{G}^{\dagger}$ is the Moore-Penrose generalized inverse of $\mathbf{G}$. Function $g(\cdot)$ is usually unknown, we can incorporate kernel functions in $g(\cdot)$. This is the so called KEML [13]. The kernel matrix $\mathbf{K} = [K(x; x_1) \cdots K(x; x_N)]^T$ ($K(\cdot)$ is the kernel function) is introduced into (9) and (10) to estimate the output of the KELM:

$$\mathbf{o} = \mathbf{KT} \qquad (11)$$

Herein, the Gaussian kernel function (RBF) is adopted

$$K(x_1; x_2) = \exp(-\lambda \|x_1 - x_2\|^2), \qquad (12)$$

where, λ need to be specified.

C. Genetic-Chaos Algorithm

The number of hidden neuron k needs to be optimized. Hence, the GA Chaos has been adopted to optimize

parameter k to improve the generalization ability of the KELM.

A common method is the genetic algorithm (GA) [20]. GA adopts the biological evolution theory and genetic mechanism to form an iterative process in the optimization. GA usually uses the selection, crossover and mutation operators in the searching. The basic workflow is from the coding, producing initial population, designing fitness function and designing operations of genetic to the design of control parameters.

However, GA may fall into the local extreme value point, namely premature convergence, leading to the fitness value of optimal individuals in a population less than population after some iteration. Unreasonable GA operators or inappropriate control parameters are likely to cause premature convergence. GA must have two properties to find the optimum solution of the problems: (l) GA can search the optimal solution in the region which include the optimal solution and (2) GA can search new area in order to jump out from local optimal solution.

In order to enhance the GA to satisfy the above two properties, the chaos optimization technology is introduced to help GA in the optimization process. Chaos optimization technology has advantage in special space-time dynamic and is able to make GA jump out from local optimal solution to effectively restrain precocious phenomena. The conception of Logistic sequence makes the chaos optimization be efficient in the global searching [21]. In the optimization, the chaotic variables is firstly mapped by Logistic from chaotic spaces into solution space; then chaotic variables finish the searching and recover from solution space to chaotic spaces.

The Mapping expression of Logistic is given by

$$x_n = \mu x_{n-1}(1 - x_{n-1}), \qquad (13)$$

where, μ is the control parameters. The Chaos optimization process is as follows.

Given any initial x_0, N chaotic variables, $X = \{x_1, x_2, ..., x_n\}$ can be obtained after N iterations with different path.

Then the ith chaotic variables will be mapped into optimization variables by the carrier

$$y_{in} = c_i + d_i x_{in}, \qquad (14)$$

where, c_i and d_i are constants.

Remember the ith chromosome fitness is f_i. To search iteratively using chaotic variables, set $n = 1$ and calculate the performance indicators f_{in} which is corresponding to y_{in}. Set the currently best point y^*, the optimal value is f^*, and make $y^* = y_0$, $f^* = f_0$. If f^* remains constant pass N iterations, then the second carrier is

$$y_{im} = x_i^* + \alpha_i(x_{im} - 0.5). \qquad (15)$$

where, *m is the* iterative step after the second carrier, α_i is a constant, and x_{im} is the smaller chaotic variables in

traversal area. Until satisfy terminate qualification, the GA-chaos finally gives the optimal solution y^* and the optimal value f^*. Through this process, the chaos algorithm can effectively find reasonable genetic operation parameters, help genetic algorithm jump out of local optimal solution.

Since chaos optimization has the advantages of global convergence and fast speed, it is able to ensure GA to avoid the premature convergence problem. As a consequence, the GA-Chaos algorithm is introduced to optimize the KELM structure parameters to strengthen its learning ability and robustness by their strong global searching optimal ability.
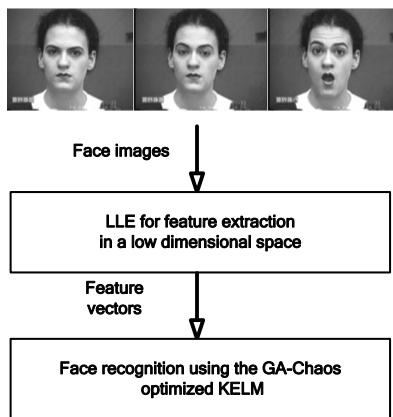

Figure 1. The block diagram of the proposed face recognition method.

The GA-Chaos optimization processes are expressed as follows:

(1) Code the chromosomes of the hidden nodes number k of the KELM.

(2) Set the initial values and the GA parameters.

(3) Calculate the corresponding value of fitness function in the solution space.

(4) Do the crossover and mutation for each chromosome to produce the corresponding fitness value.

(6) Select the optimal individuals as input of the chaos optimization to search global solution of the GA optimization.

(7) End the optimization if the results can satisfy the conditions for the termination.

A diagram block of the proposed face recognition method is shown in Fig. 1.

III. EXPERIMENTS AND RESULTS

The Japanese Female Facial Expression (JAFFEE) Data-base [22] has been used to evaluate and verify the proposed method in this paper. This data base contains 219 images of 7 facial posed expressions by 10 Japanese females. The similar lighting and the hair type of all the photos have been strict controlled. Half of the images are taken as training data and the remaining half images are taken as testing data in this work.

In the experiments, the LLE was adopted to extract useful features of the images in a low manifold space. Then the

GA Chaos-KELM was employed to recognize face images. To highlight the performance of the proposed method, the LLE was compared with Isomap, PCA, MDS, and LDA with respect to the feature extraction efficiency. Meanwhile, the GA Chaos-KELM was compared with GA-KELM, KELM, BP NN, and SVM with respect to the face recognition rate. Fig. 2 and Fig. 3 show the comparison results.
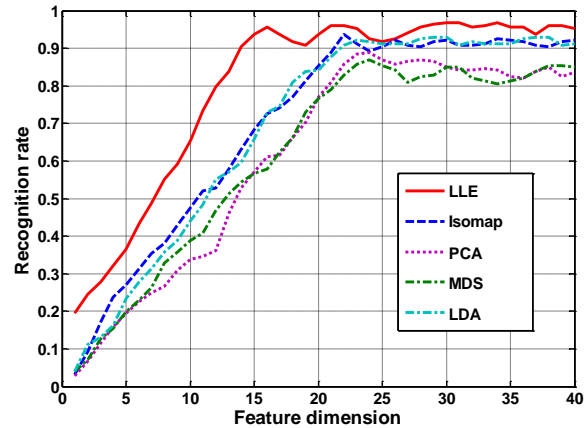

Figure 2. The comparison of the face recognition performance between LLE, Isomap, ICA, PCA and MDS using the same GA Chaos-KELM classifier.
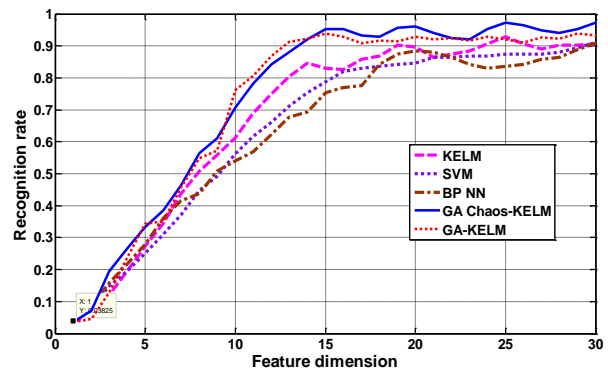

Figure 3. The comparison of the face recognition performance between GA Chaos-KELM, GA-KELM, KELM, BP NN, and SVM using the same LLE feature extraction algorithm.

In Fig. 2, we firstly use the LLE, Isomap, PCA, MDS and LDA to extract the features of the face images. Then the extracted features are input into the GA Chaos-KELM for face recognition according to different extraction methods. It can be seen in Fig. 2 that the LLE based feature extraction can produce the best face recognition rate against the Isomap, PCA, MDS and LDA. The best recognition rate of LLE is 97.2% when the input feature dimension of the GA Chaos-KELM classifier is 22. However, the highest recognition rates of Isomap, PCA, MDS and LDA are 93.8%, 89.1%, 88.2% and 92.9%, respectively. Hence, the LLE extraction improves the face recognition rate by at least 3.4% against the Isomap, PCA, MDS and LDA. This is because the LLE can preserve the nonlinear components hidden in the manifold of the original data. As a result, important features can be extracted by LLE to increase the rate in the face recognition.

In Fig. 3, we firstly use the LLE to extract the features of the face images. Then the extracted features are input into the GA Chaos-KELM, GA-KELM, KELM, BP NN, and SVM, respectively. In Fig. 3, it can be seen that the GA chaos-KELM outperforms its rivals with respect to the face recognition rate. It can be also seen in the figure that the GA-KELM has very high face recognition rate, but its performance is not good as the GA chaos-KELM. This is because the chaos can keep the GA in the global optimum. The KELM takes the third position in the face recognition rate and its performance is better than that of BP NN and SVM. This is because KELM integrates the advantages of both ANN and SVM such that inherent shortcomings of the ANN and SVM can be depressed by KELM [7]. If the only parameter of the KELM could be optimized, the performance of the optimized KELM will achieve to higher level than KELM. This is why the GA chaos algorithm has been proposed to optimize the KELM in this paper. The experimental test results prove the correction and effectiveness of this newly proposed method.

TABLE I: THE COMPARISON RESULTS OF THE FACE RECOGNITION

Method	Recognition rate	Training time
PCA-BP NN	78.9%	0.024 s
PCA-SVM	78.1%	0.019 s
PCA-KELM	81.6%	0.013 s
PCA-GA-KELM	85.5%	0.013 s
PCA-GA chaos-KELM	89.1%	0.013 s
MDS-BP NN	79.2%	0.025 s
MDS-SVM	79.8%	0.021 s
MDS-KELM	82.4%	0.013 s
MDS-GA-KELM	86.1%	0.013 s
MDS-GA chaos-KELM	88.2%	0.013 s
LDA-BP NN	85.7%	0.023 s
LDA -SVM	86.3%	0.020 s
LDA -KELM	88.1%	0.013 s
LDA-GA-KELM	91.7%	0.013 s
LDA-GA chaos-KELM	92.9%	0.013 s
Isomap-BP NN	86.1%	0.023 s
Isomap-SVM	86.6%	0.020 s
Isomap-KELM	89.7%	0.013 s
Isomap-GA-KELM	91.9%	0.013 s
Isomap-GA chaos-KELM	93.8%	0.013 s
LLE-BP NN	87.2%	0.025 s
LLE-SVM	87.8%	0.019 s
LLE-KELM	92.9%	0.013 s
LLE-GA-KELM	95.1%	0.013 s
LLE-GA chaos-KELM	97.2%	0.013 s

To further evaluate the effectiveness of the proposed face recognition method, the comparisons between different feature extraction algorithms and different face recognition classifiers have been carried out in the experiments. Table I gives the comparison results. The hybrid of the LLE, Isomap, PCA, MDS, and LDA extraction algorithms and the GA chaos-KELM, GA-KELM, KELM, BP NN, and SVM classifiers has been investigated in terms of both recognition accuracy and training speed. One can be noticed from the table that both the recognition accuracy and training speed of the

proposed LLE-GA chaos-KELM method is among the best; the optimized KELM based face recognition with the same feature extraction algorithm can get better performance than that of the KELM, BP NN and SVM; the training speeds of the KELM based methods are equal but better than that of the BP NN and SVM. Hence, the comparison results indicate that the KELM has fast training speed owning to its simple structure, and the GA chaos-KELM can enhance the face recognition rate owning to the optimization of the only parameter k.

In order to further verify the proposed method, the face images from the Cohn-Kanade Database [23]-[25] have been employed for the face recognition. The Cohn-Kanade Database contains 500 image sequences from 100 subjects. This database contains only posed expressions. The emotions of the images include joy, surprise, anger, fear, disgust, and sadness. Half of the images are taken as training data and the remaining half images are taken as testing data in this work.
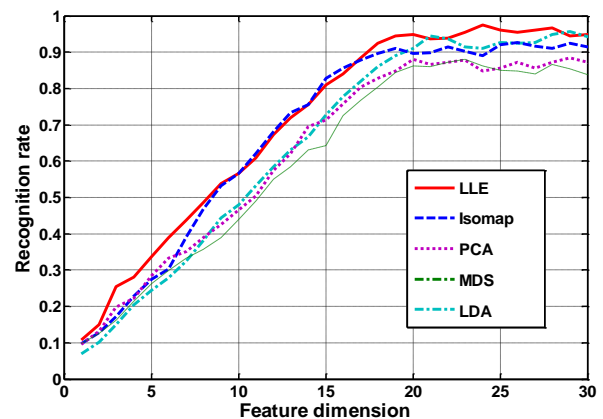

Figure 4. The comparison of the face recognition performance between LLE, Isomap, ICA, PCA and MDS using the same GA Chaos-KELM classifier.
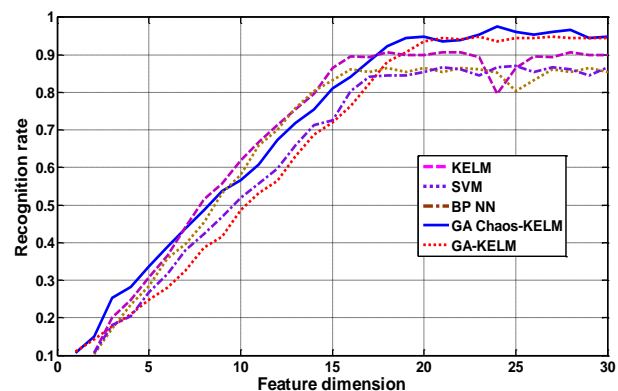

Figure 5. The comparison of the face recognition performance between GA Chaos-KELM, GA-KELM, KELM, BP NN, and SVM using the same LLE feature extraction algorithm.

In the experiments, the LLE was adopted to extract useful features of the images in a low manifold space. Then the GA Chaos-KELM was employed to recognize face images. To highlight the performance of the proposed method, the LLE was compared with Isomap, PCA, MDS,

and LDA with respect to the feature extraction efficiency. Meanwhile, the GA Chaos-KELM was compared with GA-KELM, KELM, BP NN, and SVM with respect to the face recognition rate. Fig. 4 and Fig. 5 show the comparison results.

It can be seen in Fig. 4 that the LLE based feature extraction can produce the best face recognition rate against the Isomap, PCA, MDS and LDA. The best recognition rate of LLE is 97.6%. However, the highest recognition rates of Isomap, PCA, MDS and LDA are 92.7%, 88.4%, 88.1% and 95.7%, respectively. Hence, the results verify that the LLE extraction is better than the Isomap, PCA, MDS and LDA.

In Fig. 5, it can be seen that the GA chaos-KELM outperforms its rivals with respect to the face recognition rate. It can be also seen in the figure that the GA-KELM has very high face recognition rate, but its performance is not good as the GA chaos-KELM. This is because the chaos can keep the GA in the global optimum. The KELM takes the third position in the face recognition rate and its performance is better than that of BP NN and SVM. This is because KELM integrates the advantages of both ANN and SVM such that inherent shortcomings of the ANN and SVM can be depressed by KELM [7]. If the only parameter of the KELM could be optimized, the performance of the optimized KELM will achieve to higher level than KELM. This is why the GA chaos algorithm has been proposed to optimize the KELM in this paper. The experimental test results prove the correction and effectiveness of this newly proposed method. Table II gives the overall performance of different methods.

TABLE II: THE COMPARISON RESULTS OF THE FACE RECOGNITION

Method	Recognition rate	Training time
PCA-BP NN	77.3%	0.024 s
PCA-SVM	78.7%	0.019 s
PCA-KELM	80.9%	0.013 s
PCA-GA-KELM	86.2%	0.013 s
PCA-GA chaos-KELM	88.4%	0.013 s
MDS-BP NN	79.7%	0.025 s
MDS-SVM	79.1%	0.021 s
MDS-KELM	83.3%	0.013 s
MDS-GA-KELM	85.8%	0.013 s
MDS-GA chaos-KELM	88.1%	0.013 s
LDA-BP NN	85.7%	0.023 s
LDA -SVM	86.3%	0.020 s
LDA -KELM	87.6%	0.013 s
LDA-GA-KELM	92.2%	0.013 s
LDA-GA chaos-KELM	95.7%	0.013 s
Isomap-BP NN	86.9%	0.023 s
Isomap-SVM	87.3%	0.020 s
Isomap-KELM	88.4%	0.013 s
Isomap-GA-KELM	91.3%	0.013 s
Isomap-GA chaos-KELM	92.7%	0.013 s
LLE-BP NN	87.7%	0.025 s
LLE-SVM	88.2%	0.019 s
LLE-KELM	91.7%	0.013 s
LLE-GA-KELM	94.5%	0.013 s
LLE-GA chaos-KELM	97.6%	0.013 s

One can be noticed from the table that both the recognition accuracy and training speed of the proposed LLE-GA chaos-KELM method is among the best; the optimized KELM based face recognition with the same feature extraction algorithm can get better performance than that of the KELM, BP NN and SVM; the training speeds of the KELM based methods are equal but better than that of the BP NN and SVM. Hence, the experimental test results prove the correction and effectiveness of this newly proposed method.

IV. CONCLUSIONS

In order to improve the speed and accuracy of the face recognition, a new method based on the integration of LLE, GA Chaos and KELM has been developed in this work. The innovation of this paper is that for the first time the manifold learning has been combined with the KELM to provide precise and quick response to the face recognition; moreover, GA Chaos has been used to optimize the structural parameter of the KELM to increase its generalization ability. Experimental tests have been implemented to evaluate the performance of this newly proposed LLE-GA Chaos -KELM method. The experimental analysis results suggest that the new method can achieve satisfactory performance in the face image identification. In addition, through comparison between different feature extraction algorithms (i.e. LLE, Isomap, PCA, MDS, and LDA) and different face recognition classifiers (i.e. GA Chaos-KELM, GA-KELM, KELM, BP NN, and SVM) the results indicate that the proposed LLE-GA Chaos-KELM approach outperformed its rivals in terms of both recognition accuracy and training speed. Hence the new method has practical importance. Future research will evaluate the practice efficiency of the proposed method.

REFERENCES

[1] A. Tolba, A. El-Baz, and A. El-Harby, "Face recognition: A literature review," *International Journal of Information and Communication Engineering*, vol. 2, pp. 88–103, 2006.

[2] X. Wang, "Supervised manifold learning and kernel independent component analysis applied to the face image recognition," in *Proceedings 2012 Fifth International Conference on Intelligent Computation Technology and Automation*, Zhangjiajie, China, 2012, pp. 600–603.

[3] N. Siddartha, M. Umar, N. Sen, and P. Prasad, "Feature extraction based face recognition using extreme learning machine (ELM)," *International Journal of Computer and Communication Technology*, vol. 3, pp. 10–13, 2012.

[4] M. Turk and A. Pentland, "Face recognition using eigenfaces", in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Maui, HI, 1991, pp. 586–591.

[5] J. Zhang, "New face image recognition method based on Wavelet de-noising and kernel principal component analysis," *Applied Mechanics and Materials*, vol. 235, pp. 74–77, 2012.

[6] J. Tenenbaum, V. Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319-2323, 2000.

[7] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323-2326, 2000.

[8] Z. Li, X. Yan, Z. Tain, C. Yuan, and Z. Peng, "Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis," *Measurement*, vol. 46, pp. 259-271, 2013.

[9] M. Paliwal and U. KumarA, "Neural networks and statistical techniques: A review of applications," *Expert Systems with Applications*, vol. 36, pp. 2–17, 2009.

[10] W. Zong, H. Zhou, G. Huang, and Z. Lin, "Face recognition based on kernelized extreme learning machine," *Lecture Notes in Computer Science*, vol. 6752, pp. 263–272, 201

[11] G. Huang and L. Chen, "Enhanced random search based incremental extreme learning machine," *Neurocomputing*, vol. 71, pp. 16–18, 2008.

[12] W. Zong and G. Huang, and Z. Lin, "Face recognition based on extreme learning machine," *Neurocomputing*, vol. 74, pp. 2541–2551, September 2011.

[13] W. Deng, Q. Zheng, and K. Zhang, "Reduced kernel extreme learning machine," *Advances in Intelligent Systems and Computing,* vol. 226, pp. 63–69, 2013.

[14] Z. Li, X. Yan, Z. Guo, P. Liu, Yuan C., and Z. Peng, "A new intelligent fusion method of multi-dimensional sensors and its application to tribo-system fault diagnosis of marine diesel engines," *Tribology Letters*, vol. 47, pp. 1–15, 2012.

[15] W. Xiong and C.Wang, "Hybrid feature transformation based on modified particle swarm optimization and support vector machine," *Journal of Beijing University of Posts and Telecommunications*, vol. 32, pp. 24–28, 2009.

[16] Z. Li and X. Yan, "Study on data fusion of multi-dimensional sensors for health monitoring of rolling bearings," *Insight*, vol. 55, pp. 147–151, 2013.

[17] Z. Li and X. Yan, "Application of independent component analysis and manifold learning in fault diagnosis for VSC-HVDC systems," *Hsi-An Chiao Tung Ta Hsueh*, vol. 45, pp. 44–49, 2011.

[18] B. Scholkopf, A. Smola, and R. Muller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, pp. 1299–1319, 1998.

[19] Z. Li, X. Yan, C. Yuan, J. Zhao, and Z. Peng, "Fault detection and diagnosis of the gearbox in marine propulsion system based on bispectrum analysis and artificial neural networks," *Journal of Marine Science and Application*, vol. 10, pp. 17–24, 2011.

[20] S. Bassem, T. Safa, G. Eric, and S. Robert, "Securing high resolution grayscale facial captures using a blockwise coevolutionary GA," *Expert Systems with Applications*, vol. 40, pp. 6693–6706, 2013.

[21] M. Hosseinpour, P. Nikdel, M. Badamchizadeh, and M. Poor, "Stabilizing of subspaces based on DPGA and chaos genetic algorithm for optimizing state feedback controller," *Mathematical Problems in Engineering*, vol. 2012, pp. 1–9, 2012.

[22] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, pp. 135–164, 2004.

[23] Z. Ambadar, J. Cohn, and L. Reed, "All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous," *Journal of Nonverbal Behavior*, vol. 33, pp. 17–34, 2009.

[24] S. Pollak, M. Messner, D. Kistler, and J. Cohn, "Development of perceptual expertise in emotion recognition," *Cognition*, vol. 110, pp. 242–247, 2009.

[25] M. J. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with gabor wavelets," in *Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition*, Japan, 1998, pp. 200-205.

Wei He is with Transportation Engineering Institute of Minjiang University, China. He received his PhD in Transportation Engineering from Wuhan University of Technology, China, in 2012. Dr. He is currently the vice Secretary of School of Transportation Engineering, Minjiang University. He has published more than 10 articles in peer-refereed international journals and proceedings. His research interests include prediction of traffic flow, traffic control and management, and artificial intelligence with its application in industry.
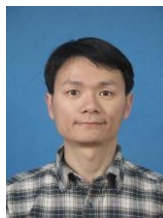
Enjun Wang is with Transportation Research Center, Wuhan Institute of Technology, China. He received his PhD in Transportation Engineering from Wuhan University of Technology, China, in and 2010. Dr. Wang is currently the director of Transportation Research Center, Wuhan Institute of Technology. His research interests include data analysis and mining, artificial intelligence with its application in transportation safety.

Ting Xiong is with School of Power and Energy Engineering, , Wuhan University of Technology, China. He received his B.S., M.S. and PhD in Transportation Engineering from Wuhan University of Technology, China, in 2004, 2007 and 2010, respectively. Dr. Xiong has published more than 10 articles in peer-refereed international journals and proceedings. His research interests include traffic control and management, and maintenance on transportation tools.