IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 24, 2020, accepted February 1, 2021, date of publication February 11, 2021, date of current version February 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3058761

Application of Locally Invariant Robust PCA for Underwater Image Recognition

PENGFEI BI[1] AND XUE DU[2]

[1]School of Artificial Intelligence, Nanjing University of Information Science and Technology, Jiangsu 210044, China
[2]College of Intelligent Systems Science and Engineering, Harbin Engineering University, Heilongjiang 150001, China

Corresponding author: Pengfei Bi (lchenxi@outlook.com)

This work was supported by the National Natural Science Foundation of China under Grant 51709062.

ABSTRACT Recently, many PCA with robust low-dimensional representation models have been applied in imaging. However, most models ignore the manifold geometry of the data and fail to minimize the reconstruction error. Here, a novel robust PCA structure, locally invariant robust principal component analysis (LIRPCA), is proposed for underwater image recognition. The contributions of LIRPCA are as follows: (1) LIRPCA selects the l_2-norm as distance metric criterion to describe the global geometry and the intrinsic geometry, which ensures the robustness of the overall model structure. (2) LIRPCA constructs a close relationship between the reconstruction error of the projected data and the input data in the cost function to minimize the reconstruction error. (3) To solve the challenging optimization function of LIRPCA, we design an iterative algorithm with fast convergence to obtain the desired solution. The proposed model is applied to feature extraction and recognition tasks from several underwater image datasets and performs better than other models.

INDEX TERMS Manifold geometry, principal component analysis, l_2-norm, underwater image recognition.

I. INTRODUCTION

In recent years, with the rapid development of artificial intelligence and machine learning, increased attention has been paid to underwater optical image recognition technology [1]. However, an underwater image in a real environment has unique attributes. First, because the underwater environment is complex and changeable, fewer samples are available for underwater environments compared with the atmosphere. Second, the random movement of underwater organisms usually leads to the formation of bubbles and the disturbance of suspended solids. This is the salt-and-pepper noise that underwater images often pick up. Moreover, the underwater optical propagation property involves diffuse reflection. Therefore, the light incident into the camera will become uneven based on different types of objects. In other words, underwater images can easily contain Gaussian noise [2], [3]. Third, because there are many uncertainties in underwater conditions, the probability of underwater image affected by occlusion is greater than that of other media, such as plankton and rocks [4]. Finally, similar to most optical images, available underwater image data are of high dimension [5], which greatly inhibits the efficacy of the model learning tasks. Thus, determining how to construct a simple and efficient recognition model for image classification according to the characteristics of underwater images has become an important issue in the field of pattern recognition.

After continuous exploration by researchers, an increasing number of subspace learning methods have been successfully applied in image recognition and representation [6], [7]. Among many methods, principal component analysis (PCA) [8] and linear discriminant analysis (LDA) [9] are the two most widely employed models. PCA technology extracts the most important features of image data by maximizing the variance of projected data. LDA is a supervised learning algorithm that yields the optimal solution through maximizing interclass variance after projecting and minimizing intraclass variance.

In image representation and recognition based on the aforementioned methods, the starting point is to protect the global geometry of the data while ignoring the manifold structure of the data. To address this issue, He Niyogi *et al.* [10] put forward the LPP method, which ensures that the structural relationship of data samples in low-dimensional space is consistent with that in high-dimensional space. Inspired by LPP, many related feature extraction algorithms have been widely developed, such as local graph embedding based on maximum margin criterion [11], fuzzy 2D discriminant

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

LPP (F2DDLPP) [12] and sparse 2DDLPP (S2DDLPP) [13]. The method proposed in [11] effectively solves the influence of some physical changes on image recognition, such as illumination, posture and expression. F2DDLPP and S2DDLPP are the representations of LPP based on a 2D matrix [14]. F2DDLPP combines 2DDLPP with fuzzy set theory to extract the discriminative features more accurately. S2DDLPP obtains the optimal sparse projection matrix by combining 2DDLPP with elastic net regression, which effectively improves the performance of the model in capturing key feature information.

The above-mentioned models adopt large distance as the metric criterion, which affects the robustness of the model to noise data. To avoid this problem, some methods employ different criterion functions to enhance the robustness of the algorithm, mainly including robust low-dimensional representation [15], [16] and robust low-rank representation [17], [18], and studies employing similar methods are popular in the field of computer vision. This paper focuses on the feature extraction of image data rather than denoising. Therefore, models related to robust low-dimensional representation will be mainly introduced. In the process of image recognition based on this type of methods, l_1-norm as a distance metric to suppress noise interference is a common strategy [19]–[22]. For instance, Ke and Kanade [20] proposed an l_1-norm-based method called L1-PCA to determine reconstruction errors in the objective function. However, solving L1-PCA problems is computationally expensive and requires difficult calculations. Kwak [22] maximized the variance based on the l_1-norm in a method to obtain principal components (PCs) named PCA-L1, which uses a greedy algorithm that does not consider the rationality of function optimization. To address this issue, Nie *et al.* [23] solved the optimal solution of PCA-L1 using an effective nongreedy iterative algorithm. Markopoulos *et al.* [24] solved PCA-L1 by using a global optimal iterative algorithm. Inspired by the l_1-norm PCA technique, l_1-norm LDA [25] and l_1-norm LPP [26] were developed for feature extraction. To further this research to include the robustness of subspace learning methods, the l_1-norm was expanded to the l_p-norm; thus, based on the l_p-norm PCA [27] and LDA [28], further methods were developed.

Feature extraction methods based on l_1-norm methods are robust to noise, but in these methods, the solutions typically are not associated with the weighted covariance matrix of the image [29]–[33], which is important for revealing the global geometric information of the data space. Moreover, these methods do not accomplish the essential purpose of the PCA model, which is to minimize reconstruction errors. To address these problems, Ding *et al.* [34] presented a novel unsupervised feature selection algorithm called R_1-PCA, which employs an l_1-norm with rotation invariance (R_1-norm) to minimize the reconstruction error. Motivated by the R_1-PCA method, a series of more effective recognition methods [35]–[38] have been developed. For example, Gao *et al.* [35] extended R_1-PCA and proposed R_1-2DPCA, which effectively protected the spatial structure of data and selected a nucleus norm as the distance metric to improve the classification effect in the classification stage. Wang *et al.* [37] proposed an Angle PCA model for image analysis, which considers the relationship between the variance of projection data and its reconstruction error, and used l_2-norm as a robust metric. Bi *et al.* [38] extended the l_2-norm to a more flexible $l_{2,p}$-norm and proposed a generalized PCA model for image analysis, which can extract the important features of the image more accurately.

In real application problems, only considering global geometry or manifold geometry in isolation has certain limitations, which make the model unable to accurately transfer the structure information in the original data space. To better balance the advantages of the two types of geometry mentioned above, Jiang *et al.* [39] proposed a robust version of the classical graph-Laplacian PCA (GLPCA) model, called robust GLPCA (RGLPCA), which employs $l_{2,1}$-norm in the cost function to suppress the influence of noise data on the model. Based on the gradual accumulation of contributions made in [39], many GLPCA models with different robust representations have been proposed [40]–[42]. Among them, the p-norm GLPCA (PGLPCA) model proposed by Feng *et al.* [40] is the most influential; this model further improves the robustness to outliers by introducing more possibilities of p-norm into the error function. However, the application scope of the robust metric norm and its inherent properties in PGLPCA are not optimal.

According to the existing problems of underwater imaging and the analysis of related research methods, we need to find a more effective feature extraction scheme to accurately capture the important information of the image and then complete the image recognition. To address this issue, in this paper, a new formulation for PCA, called locally invariant robust principal component analysis (LIRPCA), is presented. LIRPCA employs the l_2-norm as the distance metric in the optimization function, which not only improves the robustness of the model but also preserves the rotational invariance of the algorithm. Moreover, LIRPCA considers the relationship between the reconstruction error of each projected data and original input data in the cost function, which further realizes the demand of minimizing the data reconstruction error. Finally, to improve modeling performance and achieve part-based representation for the underwater optical image data, a graph-Laplacian regularization term with robust characteristics is used in the LIRPCA, which can effectively protect the manifold geometric information of the original data space. Experimental results based on three underwater optical image databases fully demonstrate the excellent performance of the proposed model.

II. RELATED WORKS

A. PCA

Consider a group of M vectorized training sample images $\mathbf{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) \in \Re^{m \times M}$, where $\boldsymbol{x}_k \in \Re^m$ $(k = 1, \ldots, M)$ represents the matrix of the kth training

image. We usually suppose that the training samples have been standardized, i.e., $\sum_{k=1}^{M} \boldsymbol{x}_k = 0$. The core demand of PCA [8] is to obtain a semi-orthogonal projection matrix $\mathbf{V} = \left[v_1, \ldots, v_q\right] \in \Re^{m \times q}$ that minimizes the reconstruction error. Generally, q is less than m in feature extraction. Therefore, the optimal projection matrix $\mathbf{V}$ can be obtained after finding the solution of the following constrained optimization problem

$$\min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} \sum_{k=1}^{\mathbf{M}} \left\| \boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T \boldsymbol{x}_k \right\|_2^2, \tag{1}$$

where $\|*\|_2$ represents the l_2-norm operations and $\mathbf{I}_q \in \Re^{q \times q}$ is an identity matrix. By using matrix trace operation, the objective function (1) can be transformed into the following set of equation

$$\max_{\mathbf{V}^T\mathbf{V}=I_q} \sum_{k=1}^{M} \left\| \mathbf{V}\mathbf{V}^T \boldsymbol{x}_k \right\|_2^2 = \max_{\mathbf{V}^T\mathbf{V}=I_q} tr\left(\mathbf{V}^T G_t \mathbf{V}\right), \tag{2}$$

where matrix $\mathbf{G}_t = \sum_{k=1}^{M} \boldsymbol{x}_k \left(\boldsymbol{x}_k\right)^T$ is called the image covariance (scatter) matrix and operator $tr\,(*)$ indicates the trace operation of a matrix. The projection matrix of Eq. (1) consists of the eigenvectors of $\mathbf{G}_t$ corresponding to the q largest eigenvalues.

B. PCA-L1

As discussed in the above analysis, the squared l_2-norm has an influence on the robustness of the traditional algorithm, which is sensitive to noise, and dominates the optimal solution of PCA [34]. To improve robustness, a large number of PCA methods based on the l_1-norm have been presented. The most classic PCA-L1 [22] technology involves finding the principal directions by solving the following formula

$$\min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} \sum_{k=1}^{M} \left\| \boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T \boldsymbol{x}_k \right\|_1, \tag{3}$$

where $\|*\|_1$ represents the l_1-norm of a vector. Although l_1-norm PCA is robust to noise, it is not flexible and does not effectively preserve the geometric information of the original data space. Moreover, the optimization process of criterion function (3) is complex and time consuming, which affects the computational efficiency of the algorithm. Thus, the solution of equation (3) is limited. To overcome above difficulties, PGLPCA is proposed [40], in which the optimization function will be introduced below.

C. PGLPCA

PGLPCA obtains the desired projection direction $\mathbf{V}$ by solving the following optimization function

$$\begin{aligned} &\min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} \sum_{i=1}^{N} \left\| \boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T \boldsymbol{x}_k \right\|_p + \xi \sum_{i,j=1}^{N} \left\| \mathbf{V}^T \boldsymbol{x}_i - \mathbf{V}^T \boldsymbol{x}_j \right\|_2^2 \mathbf{W}_{kg} \\ &= \min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} \sum_{i=1}^{N} \left\| \boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T \boldsymbol{x}_k \right\|_p + \xi tr\left(\mathbf{V}^T \mathbf{X}\mathbf{L}\mathbf{X}^T \mathbf{V}\right), \end{aligned} \tag{4}$$

where $\|\cdot\|_p$ calculates the p-norm of a vector and the range of p can vary randomly from 0 to 1.

Because equation (3) uses the flexible robust norm as the distance metric in the cost function and considers the manifold geometry, the ability of PGLPCA to suppress noise is greatly improved. However, PGLPCA is not the most ideal robust model for the following reasons: (1) PGLPCA does not consider the relationship between reconstruction error and input data in the cost function, which affects the essential goal of PCA-based models mentioned above. (2) The selection of p-norm value makes it difficult for the model to retain the rotational invariance of traditional GLPCA. (3) PGLPCA ignores the reasonable selection of the distance metric strategy in the regular term, which causes the projection direction obtained to deviate from the expected projection direction.

To improve the robustness of PCA-based models and make up for the defects in the overall structure, we propose a new conception of PCA in the following section.

III. LOCALLY INVARIANT ROBUST PCA

A. MOTIVATION AND OBJECTIVE FUNCTION

Through the analysis in Section II, we know that the performance of (3) is better than that of (1) and (2) in terms of robustness and structure description. However, the optimization function of (3) does not consider the relationship between the reconstruction results of each projected data and the original image, which causes the extracted features to be unable to reflect the key information of the original data space, thus affecting the robustness of the model. In other words, ideally, the reconstruction error of each projected data is infinitely far away from the input data, which can truly realize the demand for minimizing the reconstruction error based on the PCA model. Moreover, the robustness condition of (3) is established by abandoning the important properties of the traditional PCA model, such as rotational invariance, and the solution is related to the covariance matrix. Therefore, from the perspective of PCA robustness and its valuable properties, we hope to find a suitable robust norm as the distance measurement criterion of the model to minimize the deviation between the real projection direction and the expected projection direction. Finally, (3) only employs the robust distance metric in the cost function and ignores the regular term to protect local geometry ability, which causes the noise data to have an obvious impact on the overall performance of the model. Therefore, we aim to achieve the robustness of the model in the overall structure and improve its performance in image recognition tasks. Based on the above analysis, a more appropriate robust metric norm should be employed for the optimization function in (3), which can retain the important attributes of PCA. Compared with l_p-norm, it is feasible to introduce l_2-norm as a distance metric in the cost function and regular term. Moreover, the need to minimize the reconstruction error to the greatest extent should be considered. Therefore, a better performing PCA model is proposed for

underwater image recognition, which can be redefined as

$$\min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} \sum_{k=1}^{M} \frac{\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2}{\|\boldsymbol{x}_k\|_2} + {}^1\!/_2\xi \times \sum_{k,g=1}^{M} \left\|\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right\|_2 \mathbf{W}_{kg}, \quad (5)$$

where $\xi > 0$ is a nonnegative parameter that controls the smoothness of the low-dimensional subspace. $\mathbf{W}_{kg}$ is a heart kernel weighting and can be expressed as

$$\mathbf{W}_{kg} = \begin{cases} \exp\left(-\dfrac{\left\|\boldsymbol{x}_k - \boldsymbol{x}_g\right\|_2^2}{2\sigma^2}\right), & \text{if } \boldsymbol{x}_k \in M_h\left(\boldsymbol{x}_g\right) \\ & \text{or } \boldsymbol{x}_g \in M_h\left(\boldsymbol{x}_k\right), \\ 0, & \text{otherwise}, \end{cases} \quad (6)$$

where $\sigma > 0$ and h nearest neighbor set of $\boldsymbol{x}_k$ is $M_h\left(\boldsymbol{x}_g\right)$.

B. OPTIMIZATION

To make the algorithm optimization process simple and clear, the cost function in (5) is optimized first. Using a mathematical derivation, we have

$$\begin{aligned} &\sum_{k=1}^{M} \frac{\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2^2}{\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2 \|\boldsymbol{x}_k\|_2} \\ &= \sum_{k=1}^{M} tr\left[\left(\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right)^T\left(\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right)\right]\boldsymbol{d}_k \\ &= \sum_{k=1}^{M} tr\left[\left(\boldsymbol{x}_k\right)^T\boldsymbol{x}_k - \left(\boldsymbol{x}_k\right)^T\mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right]\boldsymbol{d}_k \\ &= \sum^{M} tr\cdot\left[\left(\boldsymbol{x}_k\right)^T\boldsymbol{x}_k\right]\boldsymbol{d}_k - \sum^{M} tr\cdot\mathbf{V}^T\boldsymbol{x}_k\left(\boldsymbol{x}_k\right)^T\mathbf{V}\Bigg]\boldsymbol{d}_k \\ &= tr\left(\mathbf{X}\mathbf{D}\mathbf{X}^T\right) - tr\left(\mathbf{V}^T\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{V}\right), \end{aligned} \quad (7)$$

where $\boldsymbol{d}_k = \frac{1}{\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2\|\boldsymbol{x}_k\|_2}$ and $\mathbf{D}$ is a diagonal matrix composed of the elements $\boldsymbol{d}_k$ on the diagonal.

Then, the regular term in (5) can be optimized using an algebraic operation; we obtain

$$\begin{aligned} &1/2\sum_{k,g=1}^{M} \frac{\left\|\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right\|_2^2}{\left\|\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right\|_2}\mathbf{W}_{kg} \\ &= 1/2\sum_{k,g=1}^{M}\left(\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right)^T\left(\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right)\tilde{\mathbf{W}}_{kg} \\ &= \sum_{k=1}^{M}\left(\mathbf{V}^T\boldsymbol{x}_k\right)^T\mathbf{V}^T\boldsymbol{x}_k\tilde{\mathbf{D}}_{kk} - \sum_{k,g=1}^{M}\left(\mathbf{V}^T\boldsymbol{x}_k\right)^T\mathbf{V}^T\boldsymbol{x}_g\tilde{\mathbf{W}}_{kg} \\ &= tr\left(\left(\mathbf{V}^T\mathbf{X}\right)^T\tilde{\mathbf{D}}\mathbf{V}^T\mathbf{X}\right) - tr\left(\left(\mathbf{V}^T\mathbf{X}\right)^T\tilde{\mathbf{W}}\mathbf{V}^T\mathbf{X}\right) \\ &= tr\left(\mathbf{V}^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}\right), \end{aligned} \quad (8)$$

where $\tilde{\mathbf{W}}_{kg} = \frac{\mathbf{W}_{kg}}{\left\|\mathbf{V}^T\left(\boldsymbol{x}_k - \boldsymbol{x}_g\right)\right\|_2}$, $\tilde{\mathbf{L}} = \tilde{\mathbf{D}} - \tilde{\mathbf{W}}$ denotes the graph-Laplacian matrix with robust characteristics; $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{W}}$ are diagonal matrix and symmetric matrix, respectively; and $\tilde{\boldsymbol{d}}_k = \sum_g \tilde{W}_{kg}$.

Finally, (7) and (8) are combined, and the merged results are substituted into (5). Now, (5) can be written as

$$\min_{\mathbf{V}^T\mathbf{V}=\mathbf{I}_q} tr\left(\mathbf{X}\mathbf{D}\mathbf{X}^T\right) - tr\left(\mathbf{V}^T\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{V}\right) + \xi tr\left(\mathbf{V}^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}\right), \quad (9)$$

Now, we consider a method of finding a projection matrix $\mathbf{V}$ that minimizes the value of the objective function (9). In (9), unknown variables $\mathbf{V}$, $\mathbf{D}$ and $\tilde{\mathbf{L}}$ are related to $\mathbf{V}$. Therefore, it is difficult to directly solve the optimal projection matrix $\mathbf{V}$ without a closed form solution. In this case, $\mathbf{V}$, $\mathbf{D}$ and $\tilde{\mathbf{L}}$ can be obtained through a nongreedy iterative algorithm that we propose. Next, the Lagrangian technique is employed to solve (9). The Lagrangian function of (9) is

$$L\left(\mathbf{V},\ \lambda\right) = tr\left(\mathbf{X}\mathbf{D}\mathbf{X}^T\right) - tr\left(\mathbf{V}^T\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{V}\right) + \xi tr\left(\mathbf{V}^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}\right) + tr\left(\lambda\left(\mathbf{V}^T\mathbf{V} - \mathbf{I}\right)\right), \quad (10)$$

where $\lambda \in \Re^{q\times q}$ is a symmetric matrix, which controls the importance of the orthonormality constraint of $\mathbf{V}$. The Karush-Kuhn-Tucker (KKT) condition is applied to find the optimal solution, we take $\frac{\partial L(\mathbf{V},\lambda)}{\partial \mathbf{V}} = 0$; then,

$$\frac{\partial L\left(\mathbf{V},\lambda\right)}{\partial \mathbf{V}} = -2\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{V} + 2\xi\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V} + 2\mathbf{V}\lambda = \mathbf{0}. \quad (11)$$

By mathematical derivation, equation (11) becomes

$$\left(\mathbf{X}\mathbf{D}\mathbf{X}^T - \xi\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\right)\mathbf{V} = \mathbf{V}\lambda. \quad (12)$$

Next, setting $\frac{\partial L(\mathbf{V},\lambda)}{\partial \lambda} = 0$, we have

$$\mathbf{V}^T\mathbf{V} = \mathbf{I}. \quad (13)$$

By substituting (14) and (13) into the objective function in (9), we can easily obtain the projection matrix $\mathbf{V}$ that satisfies the objective function. This matrix is composed of the first q eigenvectors of $\mathbf{X}\mathbf{D}\mathbf{X}^T - \xi\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T$ corresponding to the q largest eigenvalues, and this method minimizes the objective function. Then, $\boldsymbol{d}_k$ is updated. The iterative procedure is repeated until it converges, as shown in Section III C. The calculations for $\mathbf{V}$ and $\boldsymbol{d}_k$ are summarized in **Algorithm 1**.

C. CONVERGENCE ANALYSIS

In this subsection, we first theoretically prove that **Algorithm 1** monotonically decreases the value of objective function (5) in each iteration; then, we show that the proposed method converges to a local optimal solution. Before completing the convergence analysis, we first introduce the following theorem.

Algorithm 1 LIRPCA Procedure

Input: The sample matrix $\mathbf{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M) \in \Re^{m\times M}$, with parameters ξ, q and σ, where $\mathbf{X}$ is centralized.
Initialize: Set $\mathbf{V}^{(t)} \in \Re^{m\times q}$ and $t = 1$, where $\mathbf{V}$ satisfies $\mathbf{V}^T\mathbf{V} = \mathbf{I}_q$.
while not converge **do**
1. Compute the diagonal matrix $\mathbf{D} \in \Re^{M\times M}$ by each diagonal element $\boldsymbol{d}_k = 1/\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2 \|\boldsymbol{x}_k\|_2$.
2. Compute $\tilde{\mathbf{W}} \in \Re^{M\times M}$, $\tilde{\mathbf{D}} \in \Re^{M\times M}$ and $\tilde{\mathbf{L}} \in \Re^{M\times M}$, where $\tilde{\mathbf{L}} = \tilde{\mathbf{D}} - \tilde{\mathbf{W}}$.
3. Compute the covariance matrix $\mathbf{X}\mathbf{D}\mathbf{X}^T - \xi\boldsymbol{X}\tilde{\mathbf{L}}\mathbf{X}^T$.
4. Solve matrix $\mathbf{V}^{(t+1)}$ by equation (11). $\mathbf{V}^{(t+1)}$ is called the optimal projection matrix, and its column vectors are composed of the eigenvectors corresponding to the q largest eigenvalues of weighted covariance matrix.
5. If the convergence condition $J\left(\mathbf{V}^{(t)}\right) - J\left(\mathbf{V}^{(t+1)}\right) \leq \delta$ is satisfied, end the while process; otherwise, go to step 6, where $J(\mathbf{V}) = tr\left(\mathbf{X}\mathbf{D}\mathbf{X}^T\right) - tr\left(\mathbf{V}^T\mathbf{X}\mathbf{D}\mathbf{X}^T\mathbf{V}\right) + \xi tr\left(\mathbf{V}^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}\right)$.
6. $t \leftarrow t + 1$.
end while
Output: $\mathbf{V}^{(t+1)} \in \Re^{m\times q}$.

Theorem 1: Suppose any two vectors $\boldsymbol{e}^{(t)} \in \Re^m$ and $\boldsymbol{e}^{(t+1)} \in \Re^m$, we can obtain the following inequality [36]

$$\frac{\left\|\boldsymbol{e}^{(t+1)}\right\|_2}{\left\|\boldsymbol{e}^{(t)}\right\|_2} - \frac{p}{2}\frac{\left\|\boldsymbol{e}^{(t+1)}\right\|_2^2}{\left\|\boldsymbol{e}^{(t)}\right\|_2^2} - 1 + \frac{1}{2} \leq 0, \tag{14}$$

where $\boldsymbol{e}^{(t)}$ must be a non-zero vector; otherwise, the denominator is zero, and the inequality is meaningless.

Theorem 2: The value of objective function (5) monotonically decreases in each iteration until it eventually converges to the local optimality using the iterative method depicted in **Algorithm 1**.

Proof: As shown in **Algorithm 1**, in the $(t+1)$th iteration, we have

$$\begin{aligned}
&\sum_{k=1}^{M} tr\left((\boldsymbol{x}_k)^T\boldsymbol{x}_k\boldsymbol{d}_k^{(t)}\right) \\
&\quad - \sum_{k=1}^{M} tr\left(\left(\mathbf{V}^{(t+1)}\right)^T\boldsymbol{x}_k\left(\boldsymbol{x}_k\right)^T\mathbf{V}^{(t+1)}\boldsymbol{d}_k^{(t)}\right) \\
&\quad + \xi tr\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}^{(t+1)}\right) \\
&\leq \sum_{k=1}^{M} tr\left((\boldsymbol{x}_k)^T\boldsymbol{x}_k\boldsymbol{d}_k^{(t)}\right) - \sum_{k=1}^{M} tr\left(\left(\mathbf{V}^{(t)}\right)^T\boldsymbol{x}_k\left(\boldsymbol{x}_k\right)^T\mathbf{V}^{(t)}\boldsymbol{d}_k^{(t)}\right) \\
&\quad + \xi m\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T\mathbf{V}^{(t)}\right).
\end{aligned} \tag{15}$$

Because the equation $\left\|\boldsymbol{x}_k - \mathbf{V}\mathbf{V}^T\boldsymbol{x}_k\right\|_2^2 = tr\left(\boldsymbol{x}_k\left(\boldsymbol{x}_k\right)^T\right) - tr\left(\mathbf{V}^T\boldsymbol{x}_k\left(\boldsymbol{x}_k\right)^T\mathbf{V}\right)$ holds for each k, (15) can be transformed to the following inequality

$$\begin{aligned}
&\sum_{k=1}^{M}\left\|\boldsymbol{x}_k - \mathbf{V}^{(t+1)}\left(\mathbf{V}^{(t+1)}\right)^T\boldsymbol{x}_k\right\|_2^2\boldsymbol{d}_k^{(t)} + \xi tr \\
&\quad\times\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t+1)}\right) \\
&\leq \sum_{k=1}^{M}\left\|\boldsymbol{x}_k - \mathbf{V}^{(t)}\left(\mathbf{V}^{(t)}\right)^T\boldsymbol{x}_k\right\|_2^2\boldsymbol{d}_k^{(t)} + \xi tr \\
&\quad\times\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t)}\right).
\end{aligned} \tag{16}$$

By denoting $\boldsymbol{e}_k^{(t+1)} = \boldsymbol{x}_k - \mathbf{V}^{(t+1)}\left(\mathbf{V}^{(t+1)}\right)^T\boldsymbol{x}_k$, $\boldsymbol{e}_k^{(t)} = \boldsymbol{x}_k - \mathbf{V}^{(t)}\left(\mathbf{V}^{(t)}\right)^T\boldsymbol{x}_k$ and $\boldsymbol{u}_k^{(t)} = \boldsymbol{x}_k$, according to the definition of $\boldsymbol{d}_k$, (16) can be reformulated as

$$\begin{aligned}
&\sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t+1)}\right\|_2^2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2\left\|\boldsymbol{e}_k^{(t)}\right\|_2} + \xi tr\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t+1)}\right) \\
&\leq \sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2} + \xi tr\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t)}\right),
\end{aligned} \tag{17}$$

By transposition inequality (17), we obtain

$$\begin{aligned}
&\sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t+1)}\right\|_2^2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2\left\|\boldsymbol{e}_k^{(t)}\right\|_2} \leq \sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2} \\
&\quad + \xi tr\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t)}\right) - \xi^s tr \\
&\quad\times\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t+1)}\right),
\end{aligned} \tag{18}$$

Based on the property in *Theorem 1* and multiplying $1\Big/\left\|\boldsymbol{u}_k^{(t)}\right\|_2 > 0$ on both sides of the inequality, for each index k, we have

$$\begin{aligned}
&\sum_{k=1}^{M}\frac{1}{2}\frac{\left\|\boldsymbol{e}_k^{(t+1)}\right\|_2^2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2\left\|\boldsymbol{e}_k^{(t)}\right\|_2} \\
&\quad\geq \sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t+1)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2} - \sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2} + \sum_{k=1}^{M}\frac{1}{2}\frac{\left\|\boldsymbol{e}_k^{(t)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2}.
\end{aligned} \tag{19}$$

Combining (18) with (19), we obtain

$$\begin{aligned}
&\sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t+1)}\right\|_2}{\left\|\boldsymbol{u}_k^{(t)}\right\|_2} + \xi tr\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t+1)}\right) \\
&\quad\leq \sum_{k=1}^{M}\frac{\left\|\boldsymbol{e}_k^{(t)}\right\|_2}{\|\boldsymbol{u}^{(t)}\|} + \xi tr\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t)}\right).
\end{aligned} \tag{20}$$

According to the manifestation of $\boldsymbol{e}_k^{(t+1)}$, $\boldsymbol{e}_k^{(t)}$ and $\boldsymbol{u}_k^{(t)}$, (20) can be further expressed as

$$\begin{aligned}&\sum_{k=1}^{M}\frac{\left\|\boldsymbol{x}_k-\mathbf{V}^{(t+1)}\left(\mathbf{V}^{(t+1)}\right)^T\boldsymbol{x}_k\right\|_2}{\|\boldsymbol{x}_k\|_2}\\&\quad+\xi tr\left(\left(\mathbf{V}^{(t+1)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t+1)}\right)\\&\le\sum_{k=1}^{M}\frac{\left\|\boldsymbol{x}_k-\mathbf{V}^{(t)}\left(\mathbf{V}^{(t)}\right)^T\boldsymbol{x}_k\right\|_2}{\|\boldsymbol{x}_k\|_2}\\&\quad+\xi tr\left(\left(\mathbf{V}^{(t)}\right)^T\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}\mathbf{V}^{(t)}\right).\end{aligned}\tag{21}$$

Based on the existence of $\xi > 0$, by combining (21) with (8), we obtain

$$\begin{aligned}&\sum_{k=1}^{M}\frac{\left\|\boldsymbol{x}_k-\mathbf{V}^{(t+1)}\left(\mathbf{V}^{(t+1)}\right)^T\boldsymbol{x}_k\right\|_2}{\|\boldsymbol{x}_k\|_2}\\&\quad+1/2\xi\sum_{k,g=1}^{M}\left\|\left(\mathbf{V}^{(t+1)}\right)^T\left(\boldsymbol{x}_k-\boldsymbol{x}_g\right)\right\|_2\mathbf{W}_{kg}\\&\le\sum_{k=1}^{M}\frac{\left\|\boldsymbol{x}_k-\mathbf{V}^{(t)}\left(\mathbf{V}^{(t)}\right)^T\boldsymbol{x}_k\right\|_2}{\|\boldsymbol{x}_k\|_2}\\&\quad+1/2\xi\sum_{k,g=1}^{M}\left\|\left(\mathbf{V}^{(t)}\right)^T\left(\boldsymbol{x}_k-\boldsymbol{x}_g\right)\right\|_2\mathbf{W}_{kg}.\end{aligned}\tag{22}$$

Equation (22) indicates that the objective function of (5) monotonically decreases in each iteration. Combined with the convergence conditions provided in **Algorithm 1**, it can be determined that objective function (5) has a lower boundary and finally converges to the local optimal solution. Moreover, we also verify that the objective function (5) is non-incremental in the following experimental section. It is worth noting that the objective function (5) employs the Euclidean norm as the distance measurement criteria. Therefore, it is easy to prove that the proposed method has rotational invariance. Finally, note that the execution process of **Algorithm 1** involves only one iterative update process, and the most important step is to obtain projection matrix $\mathbf{V}$ by decomposing weighted covariance matrix $\mathbf{X}\mathbf{D}\mathbf{X}^T-\xi\mathbf{X}\tilde{\mathbf{L}}\mathbf{X}^T$. Based on this important information, it is not difficult to analyze that the time complexity of the algorithm designed for the optimization problem of LIRPCA can be defined as $O\left(tm^3\right)$, where t is the number of iterations until **Algorithm 1** converges to the optimal solution. Among many PCA models with robust low-dimensional representation, LIRPCA has relatively low time complexity.

In conclusion, compared with many PCA models using a robust norm metric, LIRPCA is highly robust, directly minimizes the data reconstruction error and characterizes the real geometric information of the data space, which has an important impact on the recognition results. Moreover, the proposed method protects desirable properties of PCA such as the rotational invariance, which is one of the basic properties mentioned many times in the context. Finally, LIRPCA displays fast and effective convergence.

IV. EXPERIMENTAL RESULTS

In this section, we validate the effectiveness of the proposed method using the FDT-UT [43], JEDI-UT [44] and EPIDHEU-UT [38] databases and compare it with state-of-the-art image recognition algorithms (PCA [8], PCA-L1 greedy [22], R_1-PCA [34], Angle PCA [37], $l_{2,p}$-PCA ($p = 0.5$) [15] and PgLPCA [40]. FDT-UT and JEDI-UT contain relatively well-known underwater image information. In addition, considering the diversity of the experiment, the EPIDHEU-UT database was established in the experimental site provided by Harbin Engineering University. Fig. 1 shows the real collection area of underwater image samples. In the experiments, we employ the 1NN classifier and reconstruction error to estimate the performance of different methods for underwater image recognition, in which the reconstruction error is obtained by

$$\mathbf{E}=\frac{1}{n}\sum_{k=1}^{n}\left\|\boldsymbol{x}_k^{clean}-\mathbf{V}\mathbf{Y}_k^{clean}\right\|_2,\tag{23}$$

where $\mathbf{Y}_k^{clean}=\mathbf{V}^T\boldsymbol{x}_k^{clean}$ denotes the feature vector of the k-th clean training sample.

To better verify the influence of different training samples on the model and ensure the consistency of experimental conditions, we need to fix the selection of feature dimension q in advance. The value of q can be empirically determined by the cumulative percent variance (CPV), which is defined as

$$CPV=\left(\sum_{i=1}^{q}\lambda_i\Big/\sum_{j=1}^{m}\lambda_j\right)\times 100\%\to 90\%,\tag{24}$$

where q represents the number of important feature information extracted [45]. In all experiments, we chose the value of q corresponding to CPV $\to$ 90% by considering the number of PCs. Moreover, to avoid the contingency of the experimental results, the experimental process of each database was repeated 10 times.

A. EXPERIMENTS USING THE FDT-UT DATABASE

In the FDT-UT database, there are more than 28500 images from 991 fish under multiple perspectives; from these, we selected 400 side-view grayscales of 40 fish as experimental data. Each fish was captured in 10 images with different poses. In the experiments, each image was resized to 50 $\times$ 30 pixels. We randomly selected 3 images for each fish and added occlusion in the form of black and white dots. The occlusion was randomly distributed, occupying 2% to 15% of the image region. Some clean images and occluded images of five fish are shown in Fig. 2. We randomly selected TN $= (2, 3, \ldots, 8)$ images per fish as the training data, and the rest (10 $-$ TN) of the images were used as testing data.

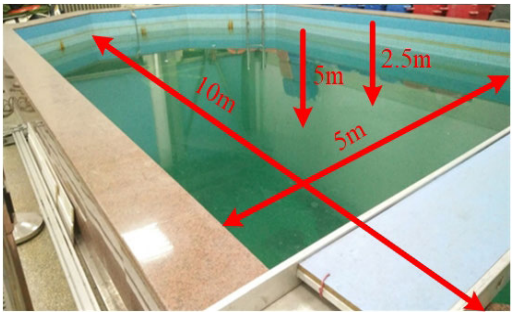

FIGURE 1. Real collection area of underwater image samples. (the length and width of the pool are 10 m and 5 m, respectively; additionally, the depth ranges from 2.5 m to 5 m).

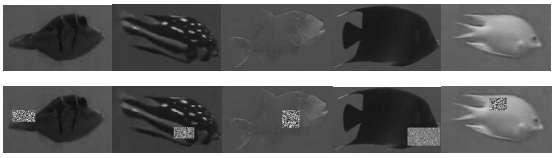

FIGURE 2. Some sample data with or without occlusion based on the FDT-UT database.

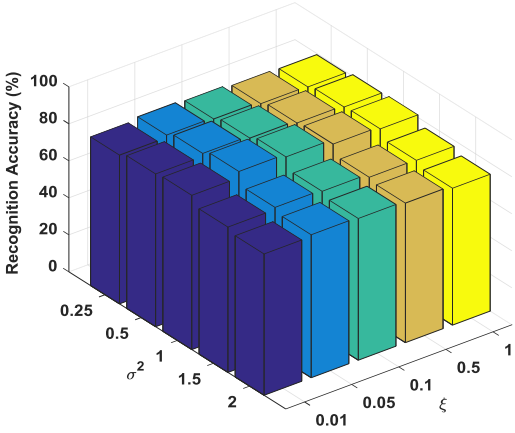

FIGURE 3. Recognition accuracy of LIRPCA versus parameters ξ and σ using the FDT-UT database.

It is worth noting that the number of occluded pictures should be controlled within 60% of the training data and that it is also not possible to involve merely clean images; otherwise, the experimental conditions will not be established, and we will need to reselect the training data. Our method and the seven aforementioned methods were used for image recognition.

Fig. 3 describes the influence of two different parameters ξ and σ on the recognition accuracy of LIRPCA in FDT-UT database. As seen in Fig. 3, when the values of ξ and σ are moderate, the performance of LIRPCA is excellent. However, when the values of ξ and σ tend to the extreme, the recognition ability of LIRPCA is relatively poor. The main reason is that if the values of ξ and σ are too small, the effect of the cost function will be more prominent in (5), which weakens the consideration of the regular term, leading to the destruction

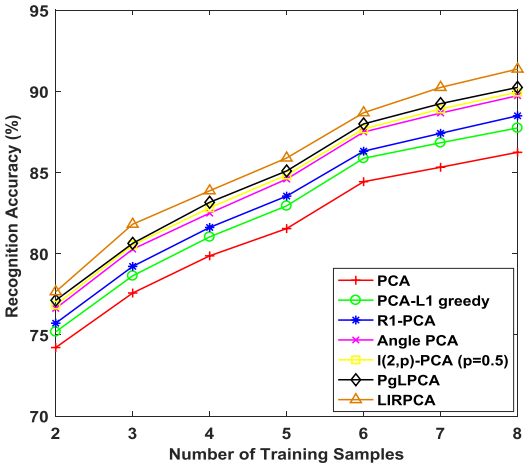

FIGURE 4. Average recognition accuracy using the FDT-UT database.

the manifold geometry of the data. Moreover, LIRPCA is prone to over-fitting. In contrast, if the values of ξ and σ are too large, the proposed model cannot effectively protect the global geometry of the data, and it is easy to cause the situation of under-fitting. Therefore, reasonable parameter selection is very important for the experimental results. The average result of each method is shown in Table 1 and plotted in Fig. 4. Fig. 4 and Table 1 illustrate that the performance for each algorithm is improved when the number of training samples increases, which is likely because each method can learn more feature information as the number of training samples increases, which is likely because each method can learn more feature information as the number of training samples increases. The accuracies of the proposed algorithm, Angle PCA, $l_{2,p}$-PCA ($p = 0.5$) and PgLPCA are better than other methods in comparative experiments, likely because the distance metric based on good robustness can effectively suppress outliers and reveals the manifold geometric information that is possibly embedded between image pixels. Notably, in the experiments, LIRPCA obtains the best recognition accuracy, likely because the low-dimensional representations of our proposed model are rotationally invariant and consider the intrinsic geometry which may be hidden between the sample data. Fig. 5 illustrates the average reconstruction error of these eight methods when TN=8. As shown in Fig. 5, the proposed model displays the minimum overall reconstruction error and is remarkably better than the other comparative models. The result is obtained because the other methods do not fully consider the relationship between the reconstruction error of each projected data and the input data, which is important for image analysis. Additionally, the solution of the proposed algorithm consists of the PCs of a weighted covariance matrix, which further reduces the deviation from the expected projection direction. To observe the reconstruction performance of the model more intuitively, we select three representative models from all the comparison methods, and complete the image reconstruction experiment under different TN (see Fig. 6). The experimental results show that the

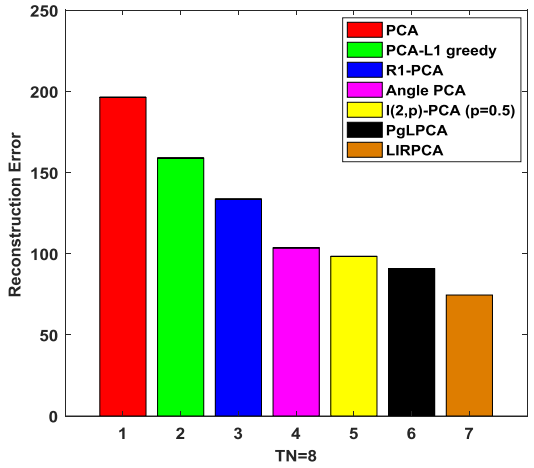

FIGURE 5. A verage reconstruction error for the eight methods based on TN = 8 using the FDT-UT database.

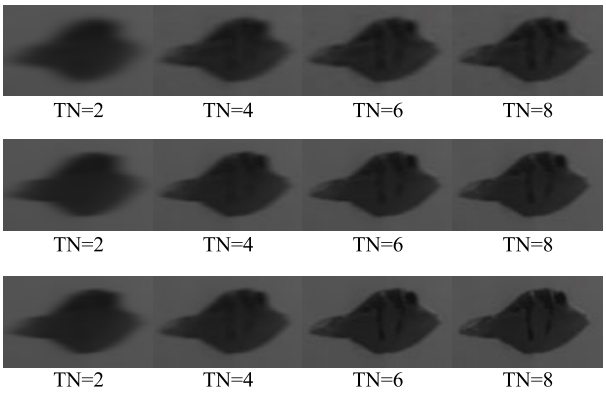

FIGURE 6. Some reconstructed sample data based on PCA (upper), PgLPCA (middle) and LIRPCA (lower) under different TN. The first clean sample data in Fig. 2 is the original data of the reconstructed image.

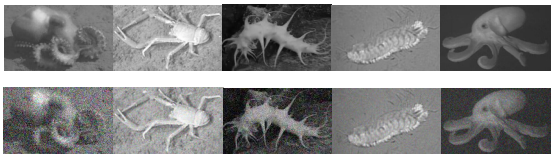

FIGURE 7. Some sample data with or without Gaussian noise based on the JEDI-UT database.

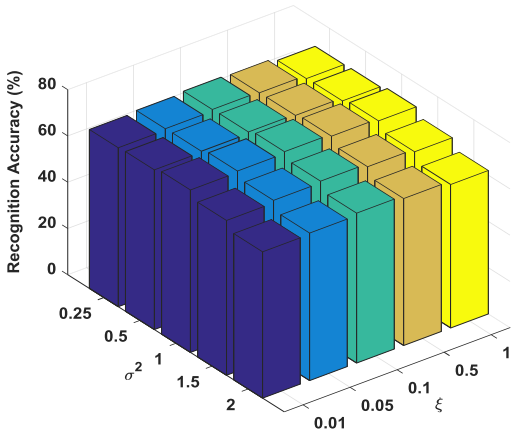

FIGURE 8. Recognition accuracy of LIRPCA versus parameters ξ and σ using the JEDI-UT database.

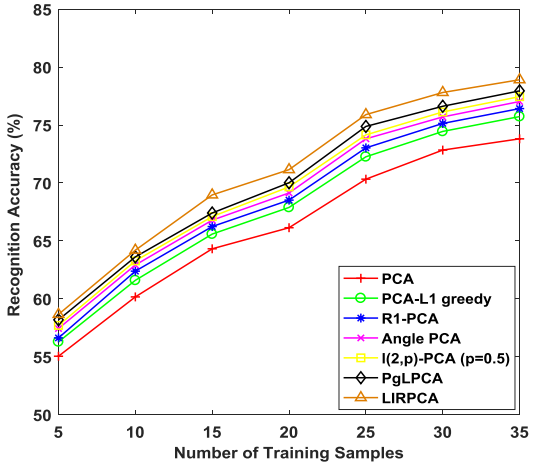

FIGURE 9. Average recognition accuracy using the JEDI-UT database.

proposed model has the best reconstruction effect. In other words, compared with other contrast models, LIRPCA can extract the important feature information of the original image more accurately. The reconstruction performance of the model is more intuitively. Therefore, LIRPCA obtains the optimal recognition result in the FDT-UT database.

B. EXPERIMENTS USING THE JEDI-UT DATABASE

This database contains more than 250 underwater species videos. We selected 30 video data samples whose background was changed; then 42 images of each video were randomly selected for experimental analysis. In the images, the body poses varied. The resolution per image was cropped to 50×30 pixels, and then 11 images per species were randomly selected to add Gaussian noise. We required randomly distributed noise in any area of the underwater image, and the fluctuation range of its intensity was set between 0.005 and 0.050. Some noise-free images and noised images of three species are shown in Fig. 7. We randomly selected $\mathrm{TN} = (5, 10, \ldots, 35)$ images per species as the training data, and the selection scheme was similar to that used with the FDT-UT database. The rest $(42 - \mathrm{TN})$ of the images of each species were used as the testing data. PCA, PCA-L1 greedy, R1-PCA, Angle PCA, $l_{2,p}$-PCA ($p = 0.5$), PgLPCA and the proposed method were used for underwater image recognition.

Fig. 8 shows the influence of model parameters ξ and σ on the recognition result of LIRPCA in JEDI-UT database. The conclusion is consistent with the FDT-UT database, i.e., moderate parameter selection can better protect the geometric structure of the model and improve the overall performance. The recognition precision is listed in Table 2 and plotted in Fig. 9. Table 2 and Fig. 9 illustrate that when TN increases from 5 to 35, the recognition results of LIRPCA, Angle PCA,

TABLE 1. Average Recognition accuracy (Standard Deviation) (%) and corresponding TN using the FDT-UT database.

TN	2	3	4	5	6	7	8
PCA	74.22	77.58	79.87	81.55	84.44	85.33	86.25
	(0.49)	(0.53)	(0.52)	(0.50)	(0.55)	(0.58)	(0.59)
PCA-L1 greedy	75.19	78.64	81.04	82.95	85.88	86.84	87.75
	(0.42)	(0.44)	(0.45)	(0.44)	(0.44)	(0.53)	(0.53)
R_1-PCA	75.72	79.22	81.62	83.55	86.32	87.42	88.50
	(0.42)	(0.44)	(0.46)	(0.44)	(0.46)	(0.47)	(0.53)
Angle PCA	76.63	80.28	82.50	84.60	87.50	88.67	89.74
	(0.36)	(0.40)	(0.44)	(0.39)	(0.42)	(0.43)	(0.53)
$l(2,p)$-PCA ($p = 0.5$)	76.84	80.54	82.83	84.85	87.69	88.92	89.96
	(0.38)	(0.38)	(0.43)	(0.41)	(0.42)	(0.41)	(0.50)
PgLPCA	77.13	80.64	83.17	85.10	88.00	89.25	90.25
	(0.38)	(0.37)	(0.40)	(0.39)	(0.40)	(0.47)	(0.53)
LIRPCA	**77.66**	**81.06**	**83.88**	**85.90**	**88.69**	**90.25**	**91.38**
	(0.27)	**(0.31)**	**(0.35)**	**(0.32)**	**(0.35)**	**(0.40)**	**(0.40)**

TABLE 2. Average recognition accuracy (Standard Deviation) (%) and corresponding TN using the JEDI-UT database.

TN	5	10	15	20	25	30	35
PCA	55.05	60.17	64.32	66.14	70.33	72.84	73.81
	(0.58)	(0.63)	(0.64)	(0.55)	(0.51)	(0.56)	(0.61)
PCA-L1 greedy	56.29	61.60	65.59	67.89	72.26	74.45	75.73
	(0.52)	(0.58)	(0.57)	(0.49)	(0.46)	(0.50)	(0.55)
R_1-PCA	56.63	62.39	66.24	68.52	73.03	75.14	76.42
	(0.50)	(0.56)	(0.55)	(0.51)	(0.46)	(0.49)	(0.57)
Angle PCA	57.44	62.92	66.79	69.14	73.81	75.70	77.03
	(0.46)	(0.51)	(0.52)	(0.42)	(0.44)	(0.46)	(0.50)
$l(2,p)$-PCA ($p = 0.5$)	57.71	63.28	67.10	69.66	74.15	76.13	77.45
	(0.43)	(0.49)	(0.51)	(0.39)	(0.40)	(0.45)	(0.49)
PgLPCA	58.19	63.62	67.41	70.02	74.88	76.62	77.95
	(0.49)	(0.56)	(0.56)	(0.45)	(0.43)	(0.48)	(0.55)
LIRPCA	**58.66**	**64.20**	**68.97**	**71.15**	**75.90**	**77.81**	**78.91**
	(0.41)	**(0.44)**	**(0.42)**	**(0.35)**	**(0.30)**	**(0.33)**	**(0.45)**

TABLE 3. Average recognition accuracy (Standard Deviation) (%) and corresponding TN using the EPIDHEU-UT database.

TN	4	6	8	10	12	14	16
PCA	63.37	67.24	71.22	76.27	79.83	82.22	83.67
	(0.66)	(0.70)	(0.73)	(0.78)	(0.76)	(0.91)	(1.05)
PCA-L1 greedy	64.92	68.86	73.00	78.20	81.92	84.44	85.67
	(0.61)	(0.64)	(0.65)	(0.71)	(0.69)	(0.74)	(0.86)
R_1-PCA	65.17	69.19	73.28	78.47	82.08	84.67	86.17
	(0.56)	(0.64)	(0.66)	(0.71)	(0.70)	(0.71)	(0.80)
Angle PCA	65.58	69.57	73.80	78.73	82.51	84.88	86.67
	(0.54)	(0.52)	(0.62)	(0.68)	(0.67)	(0.70)	(0.85)
$l(2,p)$-PCA ($p = 0.5$)	65.83	69.79	74.06	78.93	82.84	85.11	87.17
	(0.52)	(0.52)	(0.59)	(0.66)	(0.65)	(0.70)	(0.85)
PgLPCA	66.08	70.14	74.39	79.27	83.08	85.45	87.33
	(0.49)	(0.51)	(0.55)	(0.58)	(0.56)	(0.63)	(0.86)
LIRPCA	**66.33**	**70.38**	**74.83**	**79.67**	**83.50**	**86.00**	**87.83**
	(0.43)	**(0.49)**	**(0.53)**	**(0.57)**	**(0.53)**	**(0.57)**	**(0.80)**

$l_{2,p}$-PCA ($p = 0.5$) and PgLPCA are obviously better than other image recognition models. Notably, this result shows that the PCA technology with high robustness is not significantly affected by noise and outliers. Moreover, Table 2 and Fig. 9 illustrate that LIRPCA maintains an overall recognition advantage for the changing training data. This may have occurred because LIRPCA preserves the global geometric information and intrinsic geometric information of the data well, is highly robust, and therefore is ideal for recognition and representation. Fig. 10 plots the average reconstruction error for each method when TN=35. As shown in Fig. 10, LIRPCA performs reconstruction more accurately for clean images than other feature extraction methods in the experiments. The primary reason is that LIRPCA is characterized by rotational invariance and explicitly considers the reconstruction error, which is the fundamental purpose of PCA. Moreover, LIRPCA employs a robust metric scheme called the l_2-norm and reasonably associates the reconstruction error of each projected data with the input image to enhance the robustness of the algorithm, which can suppress underwater noise. Finally, LIRPCA effectively reflects the geometric information of the data space based on manifold learning, and

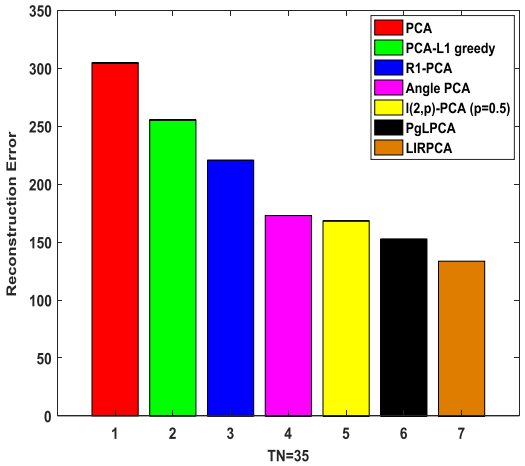

FIGURE 10. A verage reconstruction error for the eight methods based on TN = 35 using the JEDI-UT database.

FIGURE 11. Some sample data with or without Salt-and-pepper noise based on the EPIDHUE-UT database.

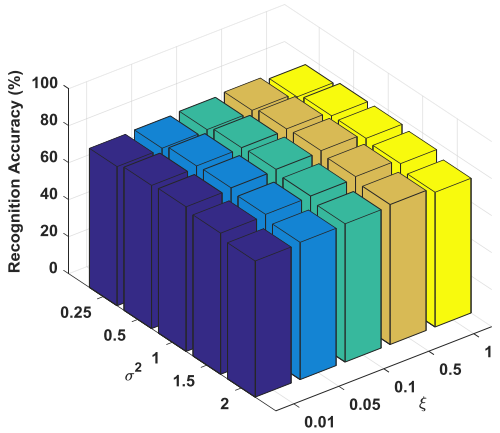

FIGURE 12. Recognition accuracy of LIRPCA versus parameters ξ and σ using the EPIDHEU-UT database.

this method is important for an accurate data representation. Overall, LIRPCA provides the optimal underwater recognition capability. These conclusions are in accordance with the FDT-UT database.

C. EXPERIMENTS USING THE EPIDHEU-UT DATABASE

The EPIDHEU-UT database is built from 300 pictures of 15 underwater targets. Each target includes 20 pictures under different views which are resized to 44 × 36 pixels. From this database, we randomly selected 5 pictures based on each target and add salt-and-pepper noise. We stipulate that the noise intensity and distribution range rules are similar to the previous two databases. Fig. 11 shows some noise-free images of underwater targets and relevant noised pictures. Based on this database, we arbitrarily selected TN = (4, 6, . . . , 16) pictures per target as the training sample, and the remaining (20 − TN) pictures per individual were used

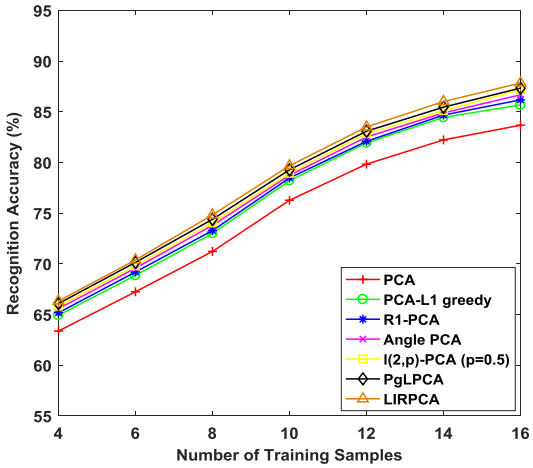

FIGURE 13. Average recognition accuracy using the EPIDHEU-UT database.

TABLE 4. Average running time (Standard Deviation) (%) and corresponding TN using the three databases.

Methods	FDT-UT (TN=8)	JEDI-UT (TN=35)	EPIDHEU-UT (TN=16)
PCA	0.07 (0.04)	0.28 (0.11)	0.24 (0.16)
PCA-L1 greedy	20.86 (0.31)	41.32 (0.23)	39.68 (0.35)
R_1-PCA	1.81 (0.08)	9.07 (0.14)	8.93 (0.19)
Angle PCA	2.03 (0.10)	10.52 (0.18)	10.26 (0.15)
$l(2,p)$-PCA (p = 0.5)	1.96 (0.09)	10.16 (0.16)	9.97 (0.15)
PgLPCA	4.65 (0.16)	17.72 (0.29)	16.54 (0.29)
LIRPCA	**2.28 (0.06)**	**10.87 (0.15)**	**10.40 (0.13)**

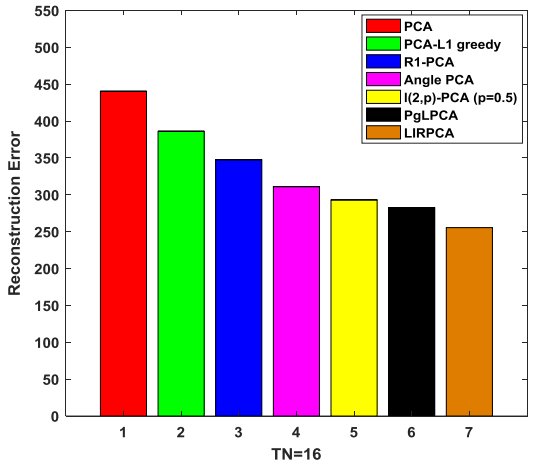

FIGURE 14. A verage reconstruction error for the eight methods based on TN = 16 using the EPIDHEU-UT database.

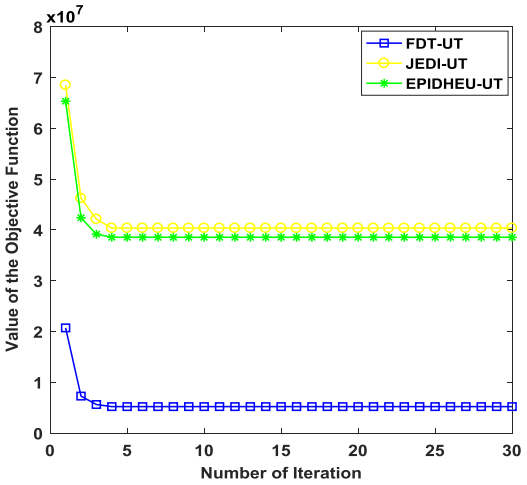

FIGURE 15. Convergence curves of the proposed method over FDT-UT, JEDI-UT and EPIDHEU-UT databases.

as the testing sample. The acquisition criteria for the training data were consistent with the standards applied in the Fig. 12 shows the recognition performance of the model when the values of parameters ξ and σ change in EPIDHEU-UT database, which is the same as the conclusion from the previous two underwater image databases. Table 3 lists the average recognition results per method in this group of experiments, which are depicted in Fig. 13. The accuracy of the LIRPCA improved from 66.33% with TN=4 to 87.83% with TN=16. Moreover, the performance of PCA, PCA-L1 greedy, R1-PCA, Angle PCA, $l_{2,p}$-PCA ($p = 0.5$) and PgLPCA improved from 63.37%, 64.92%, 65.17%, 65.58%, 65.83% and 66.08% with 4 training samples to 83.67%, 85.67%, 86.17%, 86.67%, 87.17% and 87.33% with 16 training samples, respectively. Table 3 and Fig. 13 illustrate that LIRPCA has the best underwater image recognition ability, possibly because of the high robustness strategy and the idea of preserving the original geometry information, which can effectively suppress the underwater noise. In addition, the recognition result of LIRPCA is better than that of all other methods. This superior performance is mainly achieved because the proposed method not only chooses the l_2-norm to measure in the relation between the reconstruction error of each projected data and input sample data in the cost function but also introduces the graph-Laplacian term with robust characteristics in the regular function to protect the manifold geometry of the data. Fig. 14 shows the average reconstruction error for each method when TN=16. The data reconstruction error for LIRPCA is remarkably smaller than that of other methods in comparative experiments, likely because LIRPCA has the desirable property of PCA repeatedly mentioned in the context. Moreover, the optimal solution of LIRPCA is closely related to the weighted covariance matrix. The reconstruction error is affected by weighted coefficients. Thus, LIRPCA shows the optimal performance for underwater images. This finding is consistent with the results of the previous two databases. As shown in Table 4, the running time of the proposed method is similar to those of R1-PCA, Angle PCA and $l_{2,p}$-PCA ($p = 0.5$). Additionally, the computational efficiency of LIRPCA is better than that of other matrix-based subspace learning algorithms with iterative properties, which illustrates that our proposed algorithm is fast and efficient. Thus, combined with the experimental results of recognition accuracy, our method is the most advantageous. The convergence curves of LIRPCA on the FDT-UT (TN=8), JEDI-UT (TN=35) and EPIDHEU-UT (TN=16) databases are given in Fig. 15, which illustrates that the proposed algorithm becomes more accurate with each iteration and quickly converges to a stationary point of (5). This result is in accordance with our convergence analysis in Section III.

V. CONCLUSION

We propose a new robust unsupervised recognition technique for the underwater images with small sample characteristics, namely, LIRPCA. This method selects a robust norm metric mode, called the l_2-norm, to improve the overall robustness of the optimization function and explicitly takes into account the correlation between the reconstruction error of each projection data and the input image in the cost function. Moreover, LIRPCA not only has a strong ability to protect the global geometric structure but also adequately reveals the intrinsic geometric information. Thus, LIRPCA is robust to outliers and reveals the geometric relationship of the data well, which is important for underwater image recognition. Finally, an efficient iterative algorithm is designed to solve the expectation value of LIRPCA, and a closed-form solution can be achieved in each iteration. The experimental results illustrate the effectiveness and feasibility of the proposed method based on several underwater databases.

REFERENCES

[1] X. Deng, H. Wang, and X. Liu, "Underwater image enhancement based on removing light source color and dehazing," *IEEE Access*, vol. 7, pp. 114297–114309, 2019.
[2] X. Sun, J. Shi, L. Liu, J. Dong, C. Plant, X. Wang, and H. Zhou, "Transferring deep knowledge for object recognition in low-quality underwater videos," *Neurocomputing*, vol. 275, pp. 897–908, Jan. 2018.
[3] D. Akkaynak and T. Treibitz, "A revised underwater image formation model," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake, UT, USA, Jun. 2018, pp. 6723–6732.
[4] S. Bazeille, I. Quidu, and L. Jaulin, "Color-based underwater object recognition using water light attenuation," *Intell. Service Robot.*, vol. 5, no. 2, pp. 109–118, Apr. 2012.
[5] T. Fei, D. Kraus, and A. M. Zoubir, "Contributions to automatic target recognition systems for underwater mine classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 505–518, Jan. 2015.
[6] W. Kim, R. Lee, M. Park, S.-H. Lee, and M.-S. Choi, "Low-light image enhancement using volume-based subspace analysis," *IEEE Access*, vol. 8, pp. 118370–118379, 2020.
[7] T. Napoléon and A. Alfalou, "Pose invariant face recognition: 3D model from single photo," *Opt. Lasers Eng.*, vol. 89, pp. 150–161, Feb. 2017.
[8] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jan. 1991, pp. 586–591.
[9] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[10] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Adv. Neural Inf. Process. Syst.*, 2004, pp. 153–160.

[11] M. Wan, Z. Lai, G. Yang, Z. Yang, F. Zhang, and H. Zheng, "Local graph embedding based on maximum margin criterion via fuzzy set," *Fuzzy Sets Syst.*, vol. 318, pp. 120–131, Jul. 2017.

[12] M. Wan, G. Yang, S. Gai, and Z. Yang, "Two-dimensional discriminant locality preserving projections (2DDLPP) and its application to feature extraction via fuzzy set," *Multimedia Tools Appl.*, vol. 76, no. 1, pp. 355–371, Jan. 2017.

[13] M. Wan, G. Yang, C. Sun, and M. Liu, "Sparse two-dimensional discriminant locality-preserving projection (S2DDLPP) for feature extraction," *Soft Comput.*, vol. 23, no. 14, pp. 5511–5518, Jul. 2019.

[14] M. Wan, M. Li, G. Yang, S. Gai, and Z. Jin, "Feature extraction using two-dimensional maximum embedding difference," *Inf. Sci.*, vol. 274, pp. 55–69, Aug. 2014.

[15] Q. Wang, Q. Gao, X. Gao, and F. Nie, "$\ell_{2,p}$-norm based PCA for image recognition," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1336–1346, Mar. 2018.

[16] P. Bi, J. Xu, X. Du, J. Li, and G. Chen, "$L_{2,p}$-norm sequential bilateral 2DPCA: A novel robust technology for underwater image classification and representation," *Neural Comput. Appl.*, vol. 32, no. 22, pp. 17027–17041, Nov. 2020.

[17] Y. Liu, X. Gao, Q. Gao, L. Shao, and J. Han, "Adaptive robust principal component analysis," *Neural Netw.*, vol. 119, pp. 85–92, Nov. 2019.

[18] Q. Gao, P. Zhang, W. Xia, D. Xie, X. Gao, and D. Tao, "Enhanced tensor RPCA and its application," *IEEE Trans. Pattern Anal. Mach. Intell.*, earl access, Aug. 18, 2020, doi: 10.1109/TPAMI.2020.3017672.

[19] D. Meng, Q. Zhao, and Z. Xu, "Improve robustness of sparse PCA by L_1-norm maximization," *Pattern Recognit.*, vol. 45, no. 1, pp. 487–497, Jan. 2012.

[20] Q. Ke and T. Kanade, "Robust L_1 norm factorization in the presence of outliers and missing data by alternative convex programming," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, no. 11, Jun. 2005, pp. 739–746.

[21] R. He, B.-G. Hu, W.-S. Zheng, and X.-W. Kong, "Robust principal component analysis based on maximum correntropy criterion," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1485–1494, Jun. 2011.

[22] N. Kwak, "Principal component analysis based on L_1-norm maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 9, pp. 1672–1680, Sep. 2008.

[23] F. Nie, H. Huang, C. Ding, D. Luo, and H. Wang, "Robust principal component analysis with non-greedy l1-norm maximization," in *Proc. Int. Joint Conf. Artif. Intell.*, Jul. 2011, pp. 1433–1438.

[24] P. P. Markopoulos, G. N. Karystinos, and D. A. Pados, "Optimal algorithms for L_1-subspace signal processing," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5046–5058, Oct. 2014.

[25] Y. Liu, Q. Gao, S. Miao, X. Gao, F. Nie, and Y. Li, "A non-greedy algorithm for L_1-norm LDA," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 684–695, Feb. 2017.

[26] W. Yu, R. Wang, F. Nie, F. Wang, Q. Yu, and X. Yang, "An improved locality preserving projection with L_1-norm minimization for dimensionality reduction," *Neurocomputing*, vol. 316, pp. 322–331, Nov. 2018.

[27] N. Kwak, "Principal component analysis by L_p-norm maximization," *IEEE Trans. Cybern.*, vol. 44, no. 5, pp. 594–609, Apr. 2014.

[28] J. H. Oh and N. Kwak, "Generalization of linear discriminant analysis using L_p-norm," *Pattern Recognit. Lett.*, vol. 34, no. 6, pp. 679–685, Apr. 2013.

[29] S. Zhou and D. Zhang, "Bilateral angle 2DPCA for face recognition," *IEEE Signal Process. Lett.*, vol. 26, no. 2, pp. 317–321, Feb. 2019.

[30] S. Liao, J. Li, Y. Liu, Q. Gao, and X. Gao, "Robust formulation for PCA: Avoiding mean calculation with L_2,p-norm maximization," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 3604–3610.

[31] Z. Hu, S. Liu, W. Luo, and L. Wu, "Resilient distributed fuzzy load frequency regulation for power systems under cross-layer random denial-of-service attacks," *IEEE Trans. Cybern.*, early access, Jul. 22, 2020, doi: 10.1109/TCYB.2020.3005283.

[32] Q. Wang, Q. Gao, X. Gao, and F. Nie, "Optimal mean two-dimensional principal component analysis with F-norm minimization," *Pattern Recognit.*, vol. 68, pp. 286–294, Aug. 2017.

[33] Q. Gao, L. Ma, Y. Liu, X. Gao, and F. Nie, "Angle 2DPCA: A new formulation for 2DPCA," *IEEE Trans. Cybern.*, vol. 48, no. 5, pp. 1672–1678, May 2018.

[34] C. Ding, D. Zhou, X. He, and H. Zha, "R1-PCA: Rotational invariant L_1-norm principal component analysis for robust subspace factorization," in *Proc. 23rd Int. Conf. Mach. Learn. (ICML)*, 2006, pp. 281–288.

[35] Q. Gao, S. Xu, F. Chen, C. Ding, X. Gao, and Y. Li, "R1-2-DPCA and face recognition," *IEEE Trans. Cybern.*, vol. 49, no. 4, pp. 1212–1223, Apr. 2019.

[36] W. Zhou, C. Wu, Y. Yi, and G. Luo, "Structure preserving non-negative feature self-representation for unsupervised feature selection," *IEEE Access*, vol. 5, pp. 8792–8803, 2017.

[37] Q. Wang, Q. Gao, X. Gao, and F. Nie, "Angle principal component analysis," in *Proc. Int. Joint Conf. Artif. Intell.*, vol. 2, Jul. 2017, pp. 1201–1207.

[38] J. Xu, P. Bi, X. Du, J. Li, and D. Chen, "Generalized robust PCA: A new distance metric method for underwater target recognition," *IEEE Access*, vol. 7, pp. 51952–51964, 2019.

[39] B. Jiang, C. Ding, B. Luo, and J. Tang, "Graph-Laplacian PCA: Closed-form solution and robustness," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 3492–3498.

[40] C.-M. Feng, Y.-L. Gao, J.-X. Liu, C.-H. Zheng, and J. Yu, "PCA based on graph Laplacian regularization and P-Norm for gene selection and clustering," *IEEE Trans. Nanobiosci.*, vol. 16, no. 4, pp. 257–265, Jun. 2017.

[41] J.-X. Liu, C.-M. Feng, X.-Z. Kong, and Y. Xu, "Dual graph-Laplacian PCA: A closed-form solution for bi-clustering to find 'checkerboard' structures on gene expression data," *IEEE Access*, vol. 7, pp. 151329–151338, 2019.

[42] P. Bi, J. Xu, X. Du, and J. Li, "Generalized robust graph-Laplacian PCA and underwater image recognition," *Neural Comput. Appl.*, vol. 32, no. 22, pp. 16993–17010, Nov. 2020.

[43] K.-T. Shao, J. Lin, C.-H. Wu, H.-M. Yeh, and T.-Y. Cheng, "A dataset from bottom trawl survey around taiwan," *ZooKeys*, vol. 198, pp. 103–109, May 2012.

[44] Y. Li, H. Lu, J. Li, X. Li, Y. Li, and S. Serikawa, "Underwater image de-scattering and classification by deep neural network," *Comput. Electr. Eng.*, vol. 54, pp. 68–77, Aug. 2016.

[45] J.-M. Lee, S. J. Qin, and I.-B. Lee, "Fault detection of non-linear processes using kernel independent component analysis," *Can. J. Chem. Eng.*, vol. 85, no. 4, pp. 526–536, May 2008.

PENGFEI BI received the M.S. degree in control theory and control engineering from the Harbin University of Science and Technology, China, in 2013. He is currently pursuing the Ph.D. degree in control science and engineering from Harbin Engineering University, China.

In 2020, he joined the Nanjing University of Information Science and Technology. His research interests include pattern recognition, dimensionality reduction, sparse representation, deep learning, and underwater target detection.

XUE DU was born in Harbin, Heilongjiang, China, in 1987. She received the B.S. degree in measure control technology and instruments, the M.S. degree in control engineering, and the Ph.D. degree in navigation guidance and control from Harbin Engineering University, China, in 2010, 2012, and 2016, respectively.

In 2016, she joined Harbin Engineering University. Her research interests include underwater image processing and autonomous control of unmanned underwater vehicles.