



# Classification by Principal Component Regression in the Real and Hypercomplex Domains

Moumen T. El-Melegy<sup>1</sup> · Aliaa T. Kamal<sup>2</sup> · Khaled F. Hussain<sup>2</sup> · H. M. El-Hawary<sup>3</sup>

Received: 22 April 2022 / Accepted: 16 October 2022 / Published online: 27 November 2022  
© The Author(s) 2022

## Abstract

Linear regression is a simple and widely used machine learning algorithm. It is a statistical approach for modeling the relationship between a scalar variable and one or more variables. In this paper, a classification by principal component regression (CbPCR) strategy is proposed. This strategy depends on performing regression of each data class in terms of its principal components. This CbPCR formulation leads to a new formulation of the Linear Regression Classification (LRC) problem that preserves the key information of the data classes while providing more compact closed-form solutions. For the sake of image classification, this strategy is also extended to the 4D hypercomplex domains to take into account the color information of the image. Quaternion and reduced biquaternion CbPCR strategies are proposed by representing each channel of the color image as one of the imaginary parts of a quaternion or reduced biquaternion number. Experiments on two color face recognition benchmark databases show that the proposed methods achieve better accuracies by a margin of about 3% over the original LRC and like methods.

**Keywords** Linear regression classification · Principal component analysis · Hypercomplex numbers · Face recognition

## 1 Introduction

Linear regression is a simple and widely used machine learning algorithm that has received a lot of attention in many fields. In the image recognition area, Naseem et al. [1] proposed a Linear Regression Classification (LRC) algorithm that represents each class's training images independently assuming a linear regression relationship. The algorithm depends on applying the least squares method to find the regression coefficient then decides the class label that gives the smallest reconstruction error. To enhance its performance, Huang and Yang [2] and Zhu et al. [3] proposed to apply principal component analysis (PCA) [4] to extract the vital information from images and reduce the feature vector dimensions. Then, the original data are transformed into

a low-dimensional subspace. Finally, LRC is performed on the projected data.

This paper contributes to this literature by proposing a new strategy for classification by performing regression of data in terms of its principal components. A novel formulation of the LRC problem, called Classification by Principal Component Regression (CbPCR), is presented. Moreover, a novel closed-form solution based on this formulation is derived. This classification strategy preserves key data class information and removes redundant and correlated details, yet yielding a more compact solution. Several experiments on public face recognition benchmark databases are reported to provide evidence that the proposed strategy outperforms the original LRC method [1] and its recent variants [2, 3].

The proposed strategy is also extended to color images. PCA techniques [4–6] and the existing methods [1–3] work in principle on grayscale, single-channel images. They may operate on color images after converting them to grayscale images, thus losing the important color information. Some methods (e.g., [7]) apply LRC to every color channel separately, then select the class having the smallest total prediction error over all color channels. Unlike those methods, inspired by several studies [8–14], the current paper proposes to use 4D hypercomplex numbers to represent color images.

✉ Moumen T. El-Melegy  
moumen@aun.edu.eg

<sup>1</sup> Electrical Engineering Department, Faculty of Engineering, Assiut University, Assiut, Egypt

<sup>2</sup> Computer Science Department, Faculty of Computers and Information, Assiut University, Assiut, Egypt

<sup>3</sup> Mathematics Department, Faculty of Science, Assiut University, Assiut, Egypt



This allows treating the color components of each image pixel as one entity thus considering the correlation between color components. Among the studies [8–14], two address the LRC problem. Zou et al. [11] proposes a quaternion LRC (QLRC) method that extends the classical LRC algorithm to quaternion space. QLRC converts the quaternion quantities to real ones to circumvent using quaternion derivatives. The recent paper [14] develops closed-form solutions for QLRC from the principles of quaternion calculus. In addition, the current paper proposes novel solutions based on reduced biquaternions (RBs), another hypercomplex space consisting of one real component and 3 imaginary ones. In addition to having commutative algebra—in contrast to quaternions—RBs may be represented using the so-called  $e_1$ - $e_2$  form [15] that can lead to more time-efficient computation.

The proposed CbPCR formulation is extended to both the quaternion and RB domains to process color images. To that end, the current paper exploits an efficient algorithm derived by the authors in [10] for computing the principal components (eigenvectors) of an RB matrix by casting it as an  $\mathbf{x} + \mathbf{y}$  selection problem [16, 17]. The experimental results on public benchmark databases for *color* face recognition demonstrate the better performances of the new quaternion and RB-based CbPCR algorithms over competing algorithms [7, 11, 14].

The rest of this paper is organized as follows: Sect. 2 gives a brief history of using 4D hypercomplex domains, namely quaternion and reduced biquaternions, in color image processing, with focus on color face recognition. Section 3 gives some notations and formally defines the problem of our concern here. Section 4 briefly reviews the quaternion and RB domains. Section 5 describes the proposed CbPCR method and its extension to the quaternion and the RB domains. The classification results on two benchmark color face databases are reported in Sect. 6. Section 7 concludes the paper.

## 2 Related Work

This section briefly reviews the use of 4D hypercomplex domains, namely quaternion, and reduced biquaternions, in representing color image, with focus on their application to color face recognition.

In 1996, Sangwine [18] introduced the idea of using 4D hypercomplex numbers (Quaternion numbers) in color image processing by encoding the pixel's color components into the three imaginary parts of a quaternion number. Bihan and Sangwine [8] and Pei et al. [19] proposed a Quaternion PCA method, which extracts more informative and robust features from the color image than conventional PCA. In 2011, Sun et al. [20] proposed 2DPCA and bi-dimensional PCA (BDPCA) based on quaternion representation. Also,

Javier et al. [21] proposed an independent component analysis algorithm based on quaternions. Years later, Jia et al. [22] presented a 2DPCA based on a quaternion model (2DQPCA) that depends on reducing feature dimension in a row direction. In addition, Liu et al. [23] proposed a quaternion-based maximum margin criterion algorithm. The methods proposed in [20, 22, 23] mainly targeted color face recognition and showed an enhancement in the recognition rate over their grayscale-based counterpart methods.

Using the other 4D domain of reduced biquaternions for the same purpose started rather late. A reduced biquaternion PCA based on reduced biquaternion representation was proposed [10] to represent color images in the typical PCA framework, which takes full advantage of the color characteristics of the image. Recently, a color occlusion face recognition method [24] based on the quaternion non-convex sparse constraint mechanism was proposed.

On the other hand, LRC was originally designed for grayscale images. Zou et al. [11] proposed an LRC model based on quaternion representation to consider the color image information. In [14], an LRC model was developed based on a reduced biquaternion representation. Both works [11, 14] were also applied to the face recognition problem and achieved better performances than the original LRC.

## 3 Notations and Problem Statement

Let us first give some notations to distinguish between real, quaternion, and reduced biquaternion. Scalars and vectors were denoted using italic and bold lowercase letters, respectively, while matrices are indicated by bold uppercase letters. The number of dots on top of a symbol indicates its intended domain: real ( $\mathbb{R}$ ), complex ( $\mathbb{C}$ ), quaternion ( $\mathbb{Q}$ ), or reduced biquaternion ( $\mathbb{B}$ ). Symbols without any dots on top indicate real or complex quantities, where the intended domain is disambiguated from the context. A quaternion is represented by a symbol with one dot above, while an RB quantity has two dots.

In the context of image classification, the problem of our concern can be defined as follows. Suppose there are  $L$  classes in a training set of images, where the  $l$ -th class consists of  $n_l$  images. Each image is of size  $m \times n$ . The typical goal of LRC is to find the label of a new probe image  $\mathbf{Y}$  from the given training samples and their labels. Assuming grayscale images, each image  $\mathbf{X} \in \mathbb{R}^{m \times n}$  is represented as a 1D vector  $\mathbf{x} \in \mathbb{R}^m$  by stacking the rows, one after another. As such, one can form a matrix  $\mathbf{A}_l = [\mathbf{x}_1^l, \mathbf{x}_2^l, \dots, \mathbf{x}_{n_l}^l] \in \mathbb{R}^{m \times n_l}$  that represents samples from the  $l$ -th class. For our work here, a color image is represented as a 1D pure quaternion vector  $\dot{\mathbf{x}} \in \mathbb{Q}^m$ . Analogously, the color image can be represented as a 1D pure reduced biquaternion vector  $\ddot{\mathbf{x}} \in \mathbb{B}^m$ .

## 4 Hypercomplex Domains

### 4.1 Quaternions

A quaternion number [25] consists of one real and three imaginary parts:  $\dot{q} = q_r + q_i i + q_j j + q_k k$  (a quaternion number with no real part is called a pure quaternion). The three imaginary parts satisfy:

$$\begin{aligned} i^2 = j^2 = k^2 = -1, \quad ij = -ji = k, \\ jk = -kj = i, \quad ki = -ik = j. \end{aligned} \tag{1}$$

The quaternion conjugate is  $\bar{\dot{q}} = q_r - q_i i - q_j j - q_k k$ , and the quaternion norm is  $|\dot{q}| = \sqrt{\dot{q}\bar{\dot{q}}}$ . The Hermitian (conjugate transpose) satisfies  $(\dot{\mathbf{P}}\dot{\mathbf{Q}})^H = \dot{\mathbf{Q}}^H\dot{\mathbf{P}}^H$ , see [26] for more details.

Due to the noncommutativity of quaternion multiplication, a quaternion matrix has left and right eigenvalues that may be different [15, 27]. Any quaternion matrix  $\dot{\mathbf{Q}} \in \mathbb{Q}^{n \times n}$  can be expressed as:  $\dot{\mathbf{Q}} = \mathbf{Q}_1 + \mathbf{Q}_2 j$ , where  $\mathbf{Q}_1 = \mathbf{Q}_r + i\mathbf{Q}_i$  and  $\mathbf{Q}_2 = \mathbf{Q}_j + i\mathbf{Q}_k$ , with  $\mathbf{Q}_r, \mathbf{Q}_i, \mathbf{Q}_j$ , and  $\mathbf{Q}_k$  being the real and the three imaginary parts of the quaternion matrix  $\dot{\mathbf{Q}}$ . The eigenvalues (and eigenvectors) can be calculated from the equivalent complex matrix [27]:

$$\mathcal{H}(\dot{\mathbf{Q}}) := \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ -\mathbf{Q}_2 & \mathbf{Q}_1 \end{bmatrix} \in \mathbb{C}^{2n \times 2n}, \tag{2}$$

where the mapping  $\mathcal{H}(\cdot)$  transforms a quaternion matrix into an equivalent complex matrix. Thus, there are  $2n$  eigenvalues (eigenvectors) for any  $n \times n$  quaternion matrix.

### 4.2 Reduced Biquaternions

A reduced biquaternion number also has one real and three imaginary parts:  $\ddot{p} = p_r + p_i i + p_j j + p_k k$ , where:

$$\begin{aligned} j^2 = 1, \quad i^2 = k^2 = -1, \quad kj = jk = i, \\ ji = ij = k, \quad ki = ik = -j. \end{aligned} \tag{3}$$

In contrast to quaternions, multiplication on the RB domain is commutative. There are two special numbers  $e_1$  and  $e_2$  [15] such that any RB number can be represented as:  $\ddot{p} = p_1 e_1 + p_2 e_2$ , where  $e_1 = (1 + j)/2$ ,  $e_2 = (1 - j)/2$  and  $p_1 = (p_r + p_j) + i(p_i + p_k)$ ,  $p_2 = (p_r - p_j) + i(p_i - p_k)$ .

Expressing many operations in terms of the  $e_1$ - $e_2$  forms reduces their complexity. For example, direct RB multiplication requires 16 real multiplications while applying the  $e_1$ - $e_2$  form requires only 8. The RB norm and Hermitian can be defined in a similar way as quaternion numbers. The RB

conjugate [28] is defined as:

$$\bar{\ddot{p}} = p_r - p_i i + p_j j - p_k k. \tag{4}$$

There are other definitions of conjugate [28–31] while (4) is the only definition satisfying  $\bar{\ddot{p}} = \bar{p}_1 e_1 + \bar{p}_2 e_2$ .

For any  $n \times n$  RB matrix there are  $n^2$  eigenvalues (eigenvectors) (see [10, 15] for proof). The computation of  $n^2$  eigenvalues and their corresponding eigenvectors would increase the computational cost required to find these eigenvalues. The time complexity of finding the  $t$  largest eigenvalues will be  $O(n^2 t)$ . A more efficient algorithm for this purpose was derived in our paper [10] based on the well-known computer science problem  $x + y$  selection [16, 17] with time complexity of  $O(nt + n \log n)$ .

For an  $m \times n$  RB matrix  $\ddot{\mathbf{P}}$ , the Frobenius norm  $\|\ddot{\mathbf{P}}\| = \frac{1}{\sqrt{2}} \|\mathcal{M}(\ddot{\mathbf{P}})\|$  [14] where  $\mathcal{M}(\cdot)$  maps any RB matrix to its complex equivalent matrix:

$$\mathcal{M}(\ddot{\mathbf{P}}) := \begin{bmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{P}_2 \end{bmatrix} \in \mathbb{C}^{2m \times 2n}, \tag{5}$$

where  $\mathbf{P}_1$  and  $\mathbf{P}_2 \in \mathbb{C}^{m \times n}$  and are defined as  $\mathbf{P}_1 = (\mathbf{P}_r + \mathbf{P}_j) + i(\mathbf{P}_i + \mathbf{P}_k)$ ,  $\mathbf{P}_2 = (\mathbf{P}_r - \mathbf{P}_j) + i(\mathbf{P}_i - \mathbf{P}_k)$ .

## 5 Proposed Methods

In this section, we review LRC [1], derive the CbPCR for grayscale images and then extend it for color images using the theory of quaternions and RBs.

### 5.1 Linear Regression Classification (LRC)

The goal of LRC is to infer the correct label of any new probe image  $\mathbf{y} \in \mathbb{R}^{mn}$ . LRC [1] seeks to represent  $\mathbf{y}$  as a linear combination of the training images of each class by setting up the following real-valued linear regression problem:

$$\min_{\mathbf{c}_l \in \mathbb{R}^{n_l}} \|\mathbf{y} - \mathbf{A}_l \mathbf{c}_l\|^2, \quad l = 1, 2, \dots, L, \tag{6}$$

for which a closed-form solution is found via

$$\hat{\mathbf{c}}_l = (\mathbf{A}_l^T \mathbf{A}_l)^{-1} \mathbf{A}_l^T \mathbf{y}. \tag{7}$$

Then, the predicted vector  $\hat{\mathbf{y}}_l$  is given by:

$$\hat{\mathbf{y}}_l = \mathbf{A}_l \hat{\mathbf{c}}_l, \quad l = 1, 2, \dots, L. \tag{8}$$

The test sample  $\mathbf{y}$  is eventually assigned to the class with the minimal distance

$$\min_l \|\mathbf{y} - \hat{\mathbf{y}}_l\|, \quad l = 1, 2, \dots, L. \tag{9}$$

## 5.2 Classification by Principal Component Regression

The idea in our proposed methods is based on the fact that an image (column) in  $\mathbf{A}_l$  can be represented as [4]

$$\hat{\mathbf{x}}_g^l = \boldsymbol{\mu}_l + \sum_{b=1}^t w_g^b \mathbf{U}_l^b, \quad (10)$$

where  $\boldsymbol{\mu}_l$  is the mean vector of the  $l$ -th class, and  $\mathbf{U}_l \in \mathbb{R}^{mn \times t}$  represents the  $t$  largest principal components of the class scatter matrix  $\bar{\mathbf{A}}_l \bar{\mathbf{A}}_l^T$  with  $\bar{\mathbf{A}}_l = [(\mathbf{x}_1^l - \boldsymbol{\mu}_l), (\mathbf{x}_2^l - \boldsymbol{\mu}_l), \dots, (\mathbf{x}_{n_l}^l - \boldsymbol{\mu}_l)] \in \mathbb{R}^{mn \times n_l}$ . That is, an image can be represented as a linear combination of the Eigen-components of the class-specific scatter matrix. The real weights  $\{w_g^b\}_{b=1}^t$  represent the projections of the  $g$ -th image along these components. Huang and Yang [2] and Zhu et al. [3] apply the standard LRC on these weights. Nevertheless, our CbPCR model is defined as:

$$\min_{\mathbf{c}_l \in \mathbb{R}^t} \|\mathbf{y} - \boldsymbol{\mu}_l - \mathbf{U}_l \mathbf{c}_l\|^2, \quad l = 1, 2, \dots, L. \quad (11)$$

Forcing the gradient of the objective function (11) with respect to  $\mathbf{c}_l$  to vanish, the closed form is found as:

$$\hat{\mathbf{c}}_l = (\mathbf{U}_l^T \mathbf{U}_l)^{-1} \mathbf{U}_l^T (\mathbf{y} - \boldsymbol{\mu}_l). \quad (12)$$

Since  $\mathbf{U}_l$  is orthonormal,

$$\hat{\mathbf{c}}_l = \mathbf{U}_l^T (\mathbf{y} - \boldsymbol{\mu}_l). \quad (13)$$

The response vector  $\hat{\mathbf{y}}_l$  is predicted as:

$$\hat{\mathbf{y}}_l = \mathbf{U}_l \hat{\mathbf{c}}_l + \boldsymbol{\mu}_l, \quad l = 1, 2, \dots, L. \quad (14)$$

The distance between  $\mathbf{y}$  and the predicted response vector  $\hat{\mathbf{y}}_l$  is computed as:

$$d_l(\mathbf{y}) = \|\mathbf{y} - \hat{\mathbf{y}}_l\|, \quad l = 1, 2, \dots, L, \quad (15)$$

The test image  $\mathbf{y}$  is decided to belong to the class minimizing (15).

In order to reduce the computation burden of finding the largest principal components of the class scatter matrix, the common practice [4] was followed by finding first the eigenvectors  $\mathbf{V}_l \in \mathbb{R}^{n_l \times t}$  of the matrix  $\bar{\mathbf{A}}_l^T \bar{\mathbf{A}}_l$ , and then the target eigenvectors are computed as:

$$\mathbf{U}_l = \bar{\mathbf{A}}_l \mathbf{V}_l, \quad l = 1, 2, \dots, L. \quad (16)$$

## 5.3 Quaternion-based CbPCR

Q-CbPCR is based on the algebra and calculus of quaternion matrices [32] to identify the class to which a new color image belongs. Here a  $m \times n$  training color image is portrayed as a 1D pure quaternion vector  $\dot{\mathbf{x}} \in \mathbb{Q}^{mn}$ . The columns of matrix  $\dot{\mathbf{A}}_l \in \mathbb{Q}^{mn \times n_l}$  represent samples from the  $l$ -th class where each column represents the difference between the training image and the class mean  $\dot{\boldsymbol{\mu}}_l$ . A test image  $\dot{\mathbf{y}} \in \mathbb{Q}^{mn}$  is represented by setting up the following quaternion regression problem:

$$\min_{\dot{\mathbf{c}}_l \in \mathbb{Q}^t} \|\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l - \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l\|^2, \quad l = 1, 2, \dots, L, \quad (17)$$

where  $\dot{\mathbf{U}}_l$  encompasses the key eigenvectors of the quaternion scatter matrix  $\dot{\mathbf{A}}_l \dot{\mathbf{A}}_l^H$  corresponding to the  $t$  largest eigenvalues in terms of the *quaternion* norm as computed by a QPCA technique [8].

**Proposition 1:** The closed-form solution of (17) is  $\hat{\mathbf{c}}_l = \dot{\mathbf{U}}_l^H (\dot{\mathbf{y}}_l - \dot{\boldsymbol{\mu}}_l)$ .

**Proof:** See Appendix 1.

The reconstructed query color image  $\dot{\mathbf{y}}_l$  is computed by:

$$\hat{\dot{\mathbf{y}}}_l = \dot{\mathbf{U}}_l \hat{\mathbf{c}}_l + \dot{\boldsymbol{\mu}}_l. \quad (18)$$

Eventually,  $\dot{\mathbf{y}}$  is given the label of the class with the minimal *quaternion* norm

$$\min_l \|\dot{\mathbf{y}} - \hat{\dot{\mathbf{y}}}_l\|, \quad l = 1, 2, \dots, L. \quad (19)$$

The proposed quaternion-based representation is depicted in Fig. 1, where a color face image is represented as a linear combination of the mean image and the best  $t$  eigenvectors of its class.

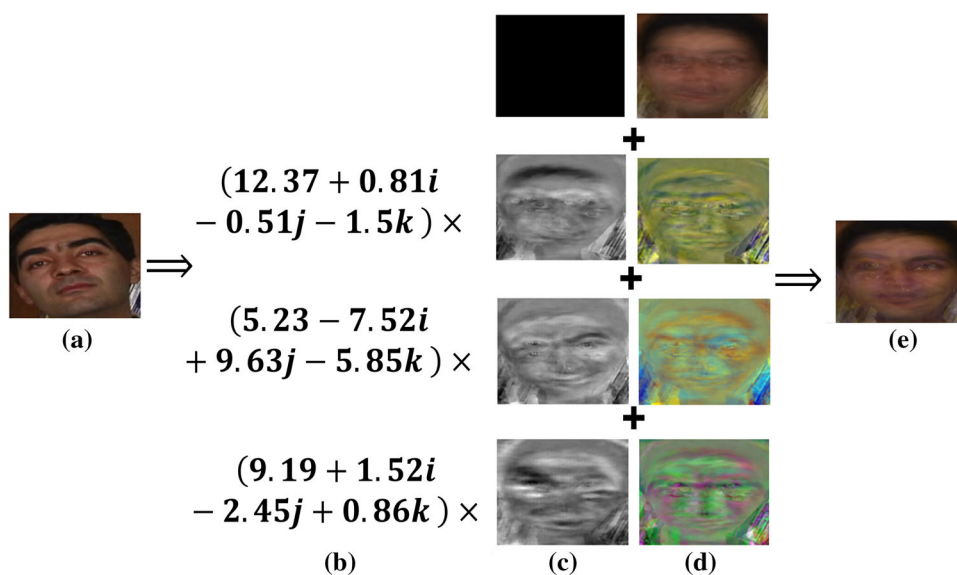
## 5.4 Reduced Biquaternion-based CbPCR

Analogously, RB-CbPCR relies on color image representation using RBs in place of quaternions. Our goal here is to find the correct label of a query color image  $\ddot{\mathbf{y}} \in \mathbb{B}^{mn}$  from the given training data matrices  $\ddot{\mathbf{A}}_l \in \mathbb{B}^{mn \times n_l}$  and their true labels. The proposed RB-CbPCR model is set up as:

$$\min_{\ddot{\mathbf{c}}_l \in \mathbb{B}^t} \|\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}_l - \ddot{\mathbf{U}}_l \ddot{\mathbf{c}}_l\|^2, \quad l = 1, 2, \dots, L. \quad (20)$$

where  $\ddot{\boldsymbol{\mu}}_l$  is the mean of the  $l$ -th class, and  $\ddot{\mathbf{U}}_l$  represents the eigenvectors of the RB scatter matrix  $\ddot{\mathbf{A}}_l \ddot{\mathbf{A}}_l^H$  corresponding to the  $t$  largest eigenvalues in terms of the RB norm as obtained

**Fig. 1** Quaternion-based color representation: Column (a) shows original color image. Column (b) gives the closed-form solution of (17) with  $t=3$ . Column (c) depicts the real parts of the principal components in  $\hat{\mathbf{U}}_l$  as grayscale images. Column (d) depicts the imaginary parts of the principal components as color images. First row depicts the real (zero) and imaginary parts of the mean of the training images. Column (e) represents the reconstructed image from (18)



via our efficient RB-based PCA algorithm [10]. By Lemma 2 in [14], it can be proved that (20) is equivalent to:

$$\min \|\mathcal{M}(\hat{\mathbf{y}}) - \mathcal{M}(\hat{\mu}_l) - \mathcal{M}(\hat{\mathbf{U}}_l)\mathcal{M}(\hat{\mathbf{c}}_l)\|^2, \tag{21}$$

where a closed-form solution is derived in proposition 2.

**Proposition 2:** The closed-form solution of (20) is  $\hat{\mathbf{c}}_l = \hat{\mathbf{U}}_l^H (\hat{\mathbf{y}} - \hat{\mu}_l)$ .

**Proof:** See Appendix 2.

The class-specific reconstructed test image is

$$\hat{\mathbf{y}}_l = \hat{\mathbf{U}}_l \hat{\mathbf{c}}_l + \hat{\mu}_l. \tag{22}$$

The test image  $\hat{\mathbf{y}}$  is finally labeled to the class with the minimal RB norm

$$\min_l \|\hat{\mathbf{y}} - \hat{\mathbf{y}}_l\|, l = 1, 2, \dots, L. \tag{23}$$

## 6 Experiments

In this section, the proposed methods are evaluated on two color face recognition benchmark databases: the GATech database [33] and the FERET database [34]. The proposed CbPCR, Q-CbPCR, and RB-CbPCR methods are compared with LRC [1], the quaternion-based QLRC method [11, 14], the RB-based RBLRC method [14], IPCRC [2, 3], and the CLRC method [7]. Note that LRC, IPCRC, and CbPCR work on grayscale images, while the other methods operate on color images.

All experiments are carried out on a pc with an Intel i7 CPU at 2.5 GHz with 8 GB RAM using MATLAB

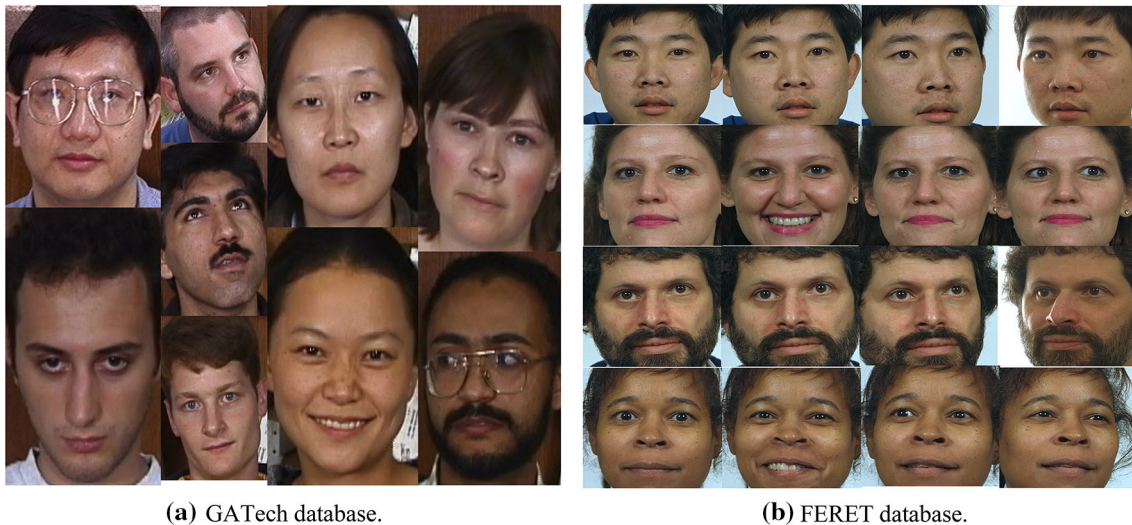
2015. Quaternion computations are done using the quaternion MATLAB toolbox [35], while RB computations are carried out using our own MATLAB toolbox.

### 6.1 Experiments on GATech Database

The GATech database [33] consists of 50 (but only 38 are available to us) subjects with 15 images per subject taken in two or three sessions. It experiences several variations in facial expression, pose, illumination, and scale; see Fig. 2a. Following [36], 10 images from each subject were used for training and the remaining 5 for testing.

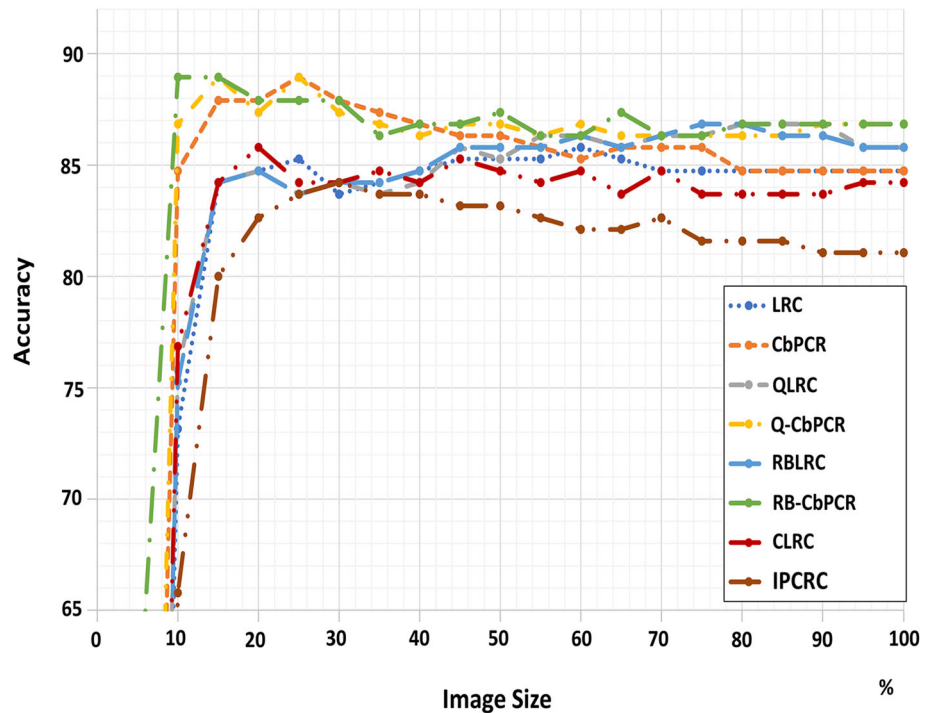
Two factors were studied on the recognition performance: the image size and the number of principal components  $t$ . Figure 3 graphs the rank-1 recognition accuracy by all methods versus the image size ranging from 5 to 100% (size  $54 \times 39$  pixels) in steps of 5%. We fix  $t=7$  per class in our proposed methods and use 266 principal components from the whole training data in the IPCRC method. Expectedly, the performances of the grayscale-based methods are generally worse than those of color-based ones. CbPCR achieves 3% and 4.7% improvement over LRC and IPCRC, respectively. IPCRC is the worst among the three and achieves a peak accuracy of 84.2% compared with respective peak accuracies of 88.9% and 85.8% by CbPCR and LRC.

The color methods offer better performances than the grayscale-based methods except for CLRC which has worse accuracy than CbPCR and a close-performance to LRC with a peak accuracy of 85.8%. The new grayscale-based CbPCR method has a better performance than QLRC and RBLRC till image size 50% then has almost the same performance afterward. The new Q-CbPCR and RB-CbPCR have a close-performance that is better than all other methods. Both have a



**Fig. 2** Sample images from GATech and FERET databases

**Fig. 3** Rank-1 recognition accuracy on the GATech database by all methods for various image sizes. (Better viewed in color)



peak accuracy of 88.95% yielding about 2.28% improvement over QLRC and RBLRC.

Figure 4 shows the rank-1 recognition accuracy of all methods against  $t = 1$  to 9 fixing the image size at  $54 \times 39$  pixels. The number of principal components in IPCRC is taken as the number of classes (38) times  $t$ . The accuracies of CLRC, LRC, QLRC, and RBLRC are not dependent on  $t$ . IPCRC is the worst overall. As more principal components are used, the proposed CbPCR, Q-CbPCR, and RB-CbPCR offer higher accuracies. For  $t \leq 4$ , CbPCR and RB-CbPCR have similar performance, while Q-CbPCR tops

both methods. Afterward, the performances of Q-CbPCR and RB-CbPCR are better than that of CbPCR. Q-CbPCR tends to offer a slightly better accuracy than RB-CbPCR, where the former has a peak accuracy of 88.95% versus 87.37% for the latter. CbPCR has 86.84% peak accuracy.

An additional experiment is carried out to compare the proposed methods (Q-CbPCR and RB-CbPCR) with existing PCA-based methods: QPCA [8], RBPCA [10], 2DQPCA [22], and 2DRBPCA [10]. Figure 5 graphs the rank-1 recognition accuracies of all these methods versus an image size from 5 to 100% (size  $54 \times 39$  pixels) in steps of 5%. Seven

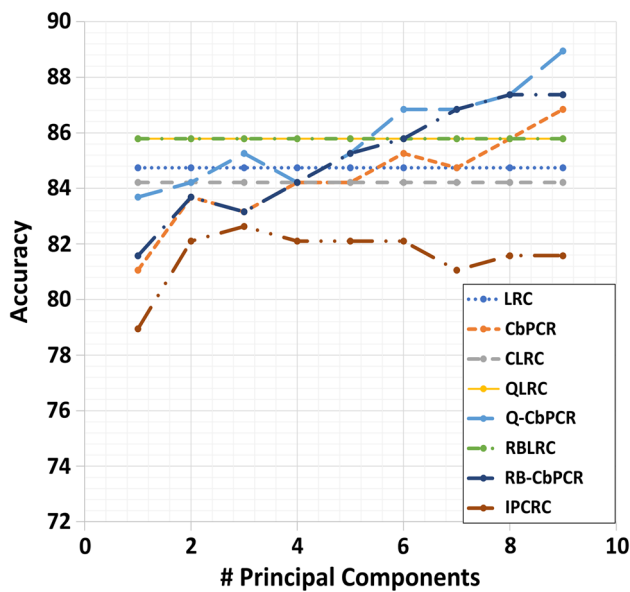


Fig. 4 Rank-1 recognition accuracy on the GATech database by all methods against number of principal components. (Better viewed in color)

principal components per class are used in the proposed methods, while 266 principal components are employed in QPCA, RBPCA, 2DQPCA, and 2DRBPCA. The performance of 1D PCA methods is the worst. QPCA is worse than RBPCA, with a peak accuracy of 82.63%, while RBPCA has a peak accuracy of 85.78%. Q-CbPCR and RB-CbPCR are

Fig. 5 Rank-1 recognition accuracy on the GATech database by PCA color methods and the proposed color methods for various image sizes. (Better viewed in color)

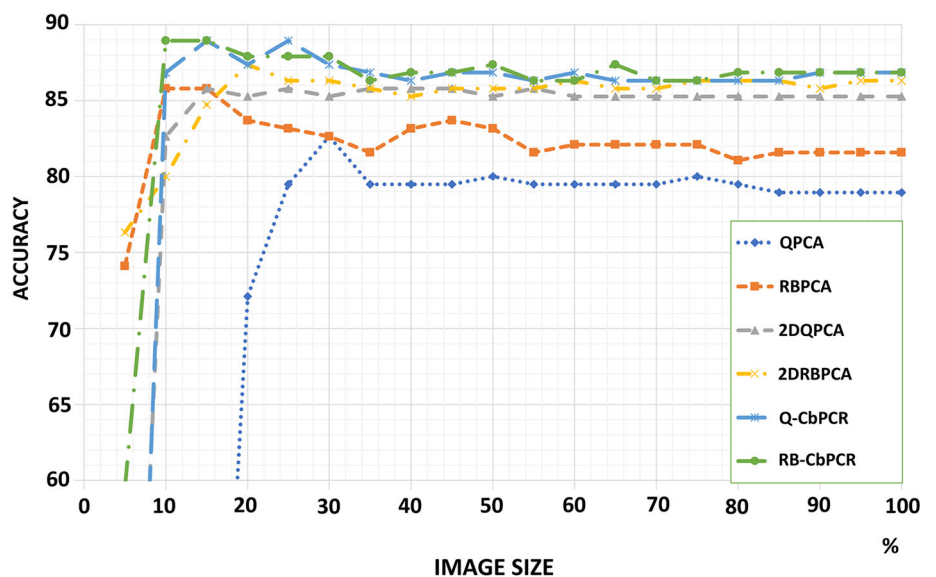


Table 1 Average recognition CPU time (in seconds) over 10 runs of different methods on the GATech database

LRC	CLRC	IPCRC	CbPCR	QLRC	RBLRC	Q-CbPCR	RB-CbPCR	QPCA	RBPCA	2DQPCA	2DRBPCA
0.018	2.860	0.083	0.105	1.631	1.182	2.011	1.121	18.455	6.653	0.475	0.821

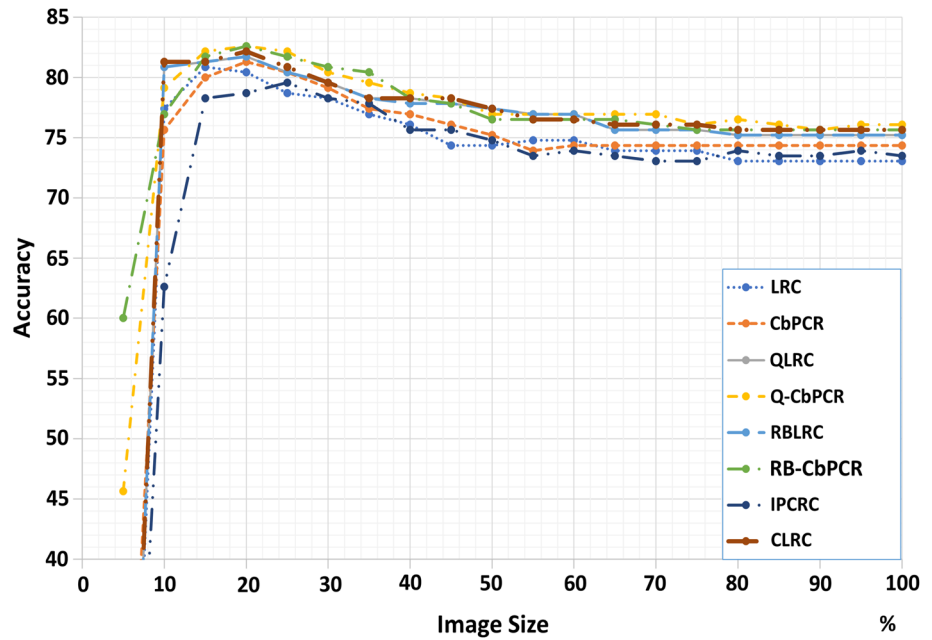
the best overall methods, with a peak accuracy of 88.95% compared to 2DQPCA and 2DRBPCA, which achieve peak accuracies of 85.79% and 87.36%, respectively.

Finally, the recognition time (in seconds) was studied by taking the average of running each algorithm 10 times on image size  $54 \times 39$ , with  $t = 7$  in the proposed methods and 266 principal components in IPCRC, QPCA, RBPCA, 2DQPCA, and 2DRBPCA. As shown in Table 1, QPCA and RBPCA are the slowest while the grayscale-based LRC and IPCRC are the fastest due to their simpler computation. 2DQPCA and 2DRBPCA are faster than Q-CbPCR and RB-CbPCR since the size of the scatter matrix in 2D methods is less than the size of the scatter matrix in the 1D methods. RB-CbPCR is around  $1.8 \times$  faster than Q-CbPCR. This is due to the faster computations of RB operations taking advantage of the  $e_1-e_2$  form and to the faster computation of the RB principal components by our efficient RB-based PCA algorithm [10].

### 6.2 Experiments on FERET database

The FERET database [34] contains more than 14,000 face images having pose and light variations. In this study, we consider a subset consisting of 115 subjects with 4 images from each subject captured in 3 poses, see Fig. 2b. Two faces are captured at  $0^\circ$  while the other two faces are captured at  $15^\circ$  and  $-15^\circ$ . One frontal image and one with the head rotated

**Fig. 6** Rank-1 recognition accuracy on the FERET database by all methods for various image sizes. (Better viewed in color)



**Table 2** Average recognition CPU time (in seconds) over 10 runs of different methods on the FERET database

LRC	CLRC	CbPCR	QLRC	Q-CbPCR	RBLRC	RB-CbPCR	IPCR
0.038	2.642	0.145	1.910	1.782	1.494	1.224	0.06

15° were chosen for training, while testing is performed on the other 2 images.

Figure 6 shows the rank-1 recognition accuracy for various image sizes varying from 5 to 100% ( $48 \times 32$  pixels) in a step of 5% using  $t = 1$  in our proposed methods and 115 principal components in IPCRC. Expectedly, the grayscale-based methods have lower accuracy. IPCRC and LRC are close to each other with respective peak accuracies of 79.6% and 80.9%. CbPCR achieves 1.79% and 1.18% improvements over LRC and IPCRC, respectively. QLRC and RBLRC exhibit nearly the same performance with a top accuracy of 81.74%. CLRC performs better than QLRC and RBLRC with a peak of 82.1%. Both Q-CbPCR and RB-CbPCR offer the best *overall* accuracy of 82.61% at 20% image size. Q-CbPCR shows a slightly better performance than RB-CbPCR for larger image sizes.

Moreover, the average recognition time (in seconds) was assessed by running each algorithm 10 times on image size  $48 \times 32$  using  $t = 1$  in the proposed methods and 115 principal components in IPCRC. As shown in Table 2, LRC and IPCRC are the fastest while CLRC is the slowest. Q-CbPCR and RB-CbPCR take less time than QLRC and RBLRC. Moreover, RB-CbPCR is about  $1.5 \times$  faster than Q-CbPCR.

## 7 Conclusions

In this paper, a novel formulation of LRC based on principal component regression has been proposed. This formulation keeps the key information of the data classes while providing more compact closed-form solutions. This formulation is also extended to the quaternion and RB domains to take into account the color information. The specific contributions of this paper are:

- CbPCR strategy is proposed by performing regression of each data class in terms of its principal components.
- This CbPCR strategy is extended to the hypercomplex domains of quaternions and RBs to consider color images.
- The CbPCR closed-form solutions are developed from the principles of real, quaternion and RB domains.
- Experiments on two color face recognition benchmark databases have showed that the proposed Q-CbPCR and RB-CbPCR have the highest overall accuracy among eight different methods including very recent ones [7, 11, 14]. Moreover, RB-CbPCR is about  $1.8 \times$  faster than Q-CbPCR. The grayscale-based CbPCR algorithm has even outperformed some color-based algorithms in the literature in addition to the original grayscale LRC method [1] and its recent variants [2, 3].



**Acknowledgements** This research is supported by ITIDA, Egypt (Grant CFP130).

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix 1

### Proof of Proposition 1

The gradient of (17) with respect to  $\dot{\mathbf{c}}_l$  is:

$$\begin{aligned} & \frac{\partial}{\partial \dot{\mathbf{c}}_l} \|\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l - \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l\|^2 \\ &= \frac{\partial}{\partial \dot{\mathbf{c}}_l} \text{tr} \left( (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l - \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l)^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l - \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l) \right) \\ &= \frac{\partial}{\partial \dot{\mathbf{c}}_l} \text{tr} \left( (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l)^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l) - \dot{\mathbf{c}}_l^H \dot{\mathbf{U}}_l^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l) \right. \\ & \quad \left. - (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l)^H \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l + \dot{\mathbf{c}}_l^H \dot{\mathbf{U}}_l^H \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l \right) \end{aligned}$$

According to quaternion derivatives [32]:

$$\frac{\partial}{\partial \dot{\mathbf{c}}_l} \|\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l - \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l\|^2 = -\frac{1}{2} \left( \overline{\dot{\mathbf{U}}_l^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l)} \right) + \frac{1}{2} \left( \overline{\dot{\mathbf{U}}_l^H \dot{\mathbf{U}}_l \dot{\mathbf{c}}_l} \right)$$

Nulling the gradient with respect to  $\dot{\mathbf{c}}_l$  at the target solution  $\widehat{\mathbf{c}}_l$  leads to:

$$\begin{aligned} \dot{\mathbf{U}}_l^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l) &= \dot{\mathbf{U}}_l^H \dot{\mathbf{U}}_l \widehat{\mathbf{c}}_l \\ \widehat{\mathbf{c}}_l &= \left( \dot{\mathbf{U}}_l^H \dot{\mathbf{U}}_l \right)^{-1} \dot{\mathbf{U}}_l^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l) \end{aligned}$$

Since  $\dot{\mathbf{U}}_l$  is orthonormal.

$$\Rightarrow \widehat{\mathbf{c}}_l = \dot{\mathbf{U}}_l^H (\dot{\mathbf{y}} - \dot{\boldsymbol{\mu}}_l) \quad \blacksquare$$

## Appendix 2

### Proof of Proposition 2

For notation brevity, let’s first drop the class specific index  $l$ . Assume  $\ddot{\mathbf{y}} = \mathbf{y}_1 e_1 + \mathbf{y}_2 e_2$ ,  $\ddot{\boldsymbol{\mu}} = \boldsymbol{\mu}_1 e_1 + \boldsymbol{\mu}_2 e_2$ ,  $\ddot{\mathbf{U}} = \mathbf{U}_1 e_1 + \mathbf{U}_2 e_2$ , and  $\ddot{\mathbf{c}} = \mathbf{c}_1 e_1 + \mathbf{c}_2 e_2$ . By Lemma 1 and Lemma 2 in [14],

$$\|(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \ddot{\mathbf{U}}\ddot{\mathbf{c}}\|^2 = \frac{1}{2} |\mathcal{M}(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \mathcal{M}(\ddot{\mathbf{U}})\mathcal{M}(\ddot{\mathbf{c}})|^2.$$

According to the properties of calculus on complex domain,

$$\begin{aligned} & \frac{\partial}{\partial \mathbf{c}_1} \|\mathcal{M}(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \mathcal{M}(\ddot{\mathbf{U}})\mathcal{M}(\ddot{\mathbf{c}})\|^2 \\ &= \frac{\partial}{\partial \mathbf{c}_1} \text{tr} \left( (\mathcal{M}(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \mathcal{M}(\ddot{\mathbf{U}})\mathcal{M}(\ddot{\mathbf{c}}))^H (\mathcal{M}(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \mathcal{M}(\ddot{\mathbf{U}})\mathcal{M}(\ddot{\mathbf{c}})) \right) \\ &= 2 \left( \mathbf{U}_1^H \mathbf{U}_1 \mathbf{c}_1 - \mathbf{U}_1^H (\mathbf{y}_1 - \boldsymbol{\mu}_1) \right) \end{aligned}$$

Similarly,

$$\begin{aligned} & \frac{\partial}{\partial \mathbf{c}_2} \|\mathcal{M}(\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) - \mathcal{M}(\ddot{\mathbf{U}})\mathcal{M}(\ddot{\mathbf{c}})\|^2 \\ &= 2 \left( \mathbf{U}_2^H \mathbf{U}_2 \mathbf{c}_2 - \mathbf{U}_2^H (\mathbf{y}_2 - \boldsymbol{\mu}_2) \right) \end{aligned}$$

The solution  $\widehat{\mathbf{c}}_1$  is found by setting the gradient of the objective function with respect to  $\mathbf{c}_1$  to zero

$$\widehat{\mathbf{c}}_1 = \left( \mathbf{U}_1^H \mathbf{U}_1 \right)^{-1} \mathbf{U}_1^H (\mathbf{y}_1 - \boldsymbol{\mu}_1)$$

Similarly, the solution  $\widehat{\mathbf{c}}_2$  is:

$$\widehat{\mathbf{c}}_2 = \left( \mathbf{U}_2^H \mathbf{U}_2 \right)^{-1} \mathbf{U}_2^H (\mathbf{y}_2 - \boldsymbol{\mu}_2).$$

According to (5),

$$\mathcal{M}(\widehat{\mathbf{c}}) = \begin{bmatrix} (\mathbf{U}_1^H \mathbf{U}_1)^{-1} \mathbf{U}_1^H (\mathbf{y}_1 - \boldsymbol{\mu}_1) & 0 \\ 0 & (\mathbf{U}_2^H \mathbf{U}_2)^{-1} \mathbf{U}_2^H (\mathbf{y}_2 - \boldsymbol{\mu}_2) \end{bmatrix}.$$

which is simplified to

$$= \begin{bmatrix} \mathbf{U}_1^H \mathbf{U}_1 & 0 \\ 0 & \mathbf{U}_2^H \mathbf{U}_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{U}_1^H & 0 \\ 0 & \mathbf{U}_2^H \end{bmatrix} \begin{bmatrix} (\mathbf{y}_1 - \boldsymbol{\mu}_1) & 0 \\ 0 & (\mathbf{y}_2 - \boldsymbol{\mu}_2) \end{bmatrix},$$

$$\therefore \widehat{\mathbf{c}} = \left( \ddot{\mathbf{U}}^H \ddot{\mathbf{U}} \right)^{-1} \ddot{\mathbf{U}}^H (\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}})$$

Since  $\ddot{\mathbf{U}}$  is orthonormal.

$$\therefore \widehat{\mathbf{c}} = \ddot{\mathbf{U}}^H (\ddot{\mathbf{y}} - \ddot{\boldsymbol{\mu}}) \quad \blacksquare$$

## References

- Naseem, I.; Togneri, R.; Bennamoun, M.: Linear regression for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(11), 2106–2112 (2010)
- Huang, S.-M.; Yang, J.-F.: Improved principal component regression for face recognition under illumination variations. *IEEE Signal Process. Lett.* **19**(4), 179–182 (2012)
- Zhu, Y.; Zhu, C.; Li, X.: Improved principal component analysis and linear regression classification for face recognition. *Signal Process.* **145**, 175–182 (2018)
- Turk, M.; Pentland, A.: Eigenfaces for recognition. *J. Cogn. Neurosci.* **3**(1), 71–86 (1991)
- Zhao, M.; Jia, Z.; Cai, Y.; Chen, X.; Gong, D.: Advanced variations of two-dimensional principal component analysis for face recognition. *Neurocomputing* **452**, 653–664 (2021)
- Yang, J.; Zhang, D.; Frangi, A.F.; Yang, J.: Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(1), 131–137 (2004)
- Yang, W.-J., Lo, C.-Y., Chung, P.-C., Yang, J. F.: Weighted Module Linear Regression Classifications for Partially-Occluded Face Recognition. *Digit. Image Process. Adv. Appl. IntechOpen* (2021).
- Le Bihan, N., Sangwine, S. J.: Quaternion principal component analysis of color images. In: *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, 2003, vol. 1, pp. I–809.
- Shi, L.: *Exploration in quaternion colour*. Doctoral dissertation, School of Computing Science-Simon Fraser University (2005).
- El-Melegy, M. T., Kamal, A. T.: Color image processing using reduced biquaternions with application to face recognition in a PCA framework. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2017, pp. 3039–3046.
- Zou, C.; Kou, K.I.; Dong, L.; Zheng, X.; Tang, Y.Y.: From grayscale to color: Quaternion linear regression for color face recognition. *IEEE Access* **7**, 154131–154140 (2019)
- Miao, J., Kou, K. I.: Quaternion matrix regression for color face recognition. *arXiv Prepr. arXiv2001.10677* (2020).
- Gai, S.; Huang, X.: Reduced biquaternion convolutional neural network for color image processing. *IEEE Trans. Circuits Syst. Video Technol.* **3**, 1061–1075 (2021)
- El-Melegy, M., Kamal, A.: Linear Regression Classification in the Quaternion and Reduced Biquaternion Domains. *IEEE Signal Process. Lett.*, p. 1, 2022.
- Pei, S.C.; Chang, J.H.; Ding, J.J.; Chen, M.Y.: “Eigenvalues and singular value decompositions of reduced biquaternion matrices. *IEEE Trans Circuits Syst. I Regul. Pap.* **55**(9), 2673–2685 (2008)
- Harper, L.H.; Payne, T.H.; Savage, J.E.; Straus, E.: Sorting  $x+y$ . *Commun. ACM* **18**(6), 347–349 (1975)
- Lambert, J.-L.: Sorting the sums  $(xi+yj)$  in  $O(n^2)$  comparisons. *Theor. Comput. Sci.* **103**(1), 137–141 (1992)
- Sangwine, S.J.: Fourier transforms of colour images using quaternion or hypercomplex, numbers. *Electron. Lett.* **32**(21), 1979–1980 (1996)
- Chang, J.-H., Ding, J.-J.: Quaternion matrix singular value decomposition and its applications for color image processing. In: *Proceedings 2003 International Conference on Image Processing (Cat. No. 03CH37429)*, 2003, vol. 1, pp. I–805.
- Sun, Y.; Chen, S.; Yin, B.: Color face recognition based on quaternion matrix representation. *Pattern Recognit. Lett.* **32**(4), 597–605 (2011)
- Vía, J.; Palomar, D.P.; Vielva, L.; Santamaría, I.: Quaternion ICA from second-order statistics. *IEEE Trans. Signal Process.* **59**(4), 1586–1600 (2010)
- Jia, Z., Ling, S.-T., Zhao, M.-X.: Color Two-Dimensional Principal Component Analysis for Face Recognition Based on Quaternion Model. In: *International Conference on Intelligent Computing*, Springer, 2017, pp. 177–189.
- Liu, Z.; Qiu, Y.; Peng, Y.; Pu, J.; Zhang, X.: Quaternion based maximum margin criterion method for color face recognition. *Neural Process. Lett.* **45**(3), 913–923 (2017)
- Wen, C.; Qiu, Y.: Color occlusion face recognition method based on quaternion non-convex sparse constraint mechanism. *Sensors* **22**(14), 5284 (2022)
- Hamilton, W.R.: On quaternions; or on a new system of imaginaries in algebra. *Philos. Mag.* **25**(3), 489–495 (1844)
- Tian, Y.: Matrix theory over the complex quaternion algebra. *arXiv Prepr. math/0004005* (2000).
- Zhang, F.: Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251**, 21–57 (1997)
- Kösal, H.H.: Least-squares solutions of the reduced biquaternion matrix equation  $AX=B$  and their applications in colour image restoration. *J. Mod. Opt.* **66**(18), 1802–1810 (2019)
- Schutte, H., Wenzel, J.: Hypercomplex numbers in digital signal processing. In: *1990 IEEE International Symposium on Circuits and Systems (ISCAS)*, 1990, pp. 1557–1560 vol.2.
- Ueda, K., Takahashi, S.: Digital filters with hypercomplex coefficients. *Electron. Commun. Japan (Part III Fundam. Electron. Sci., vol. 76, no. 9, pp. 85–98* (1993).
- Pei, S.C.; Chang, J.H.; Ding, J.J.: Commutative reduced biquaternions and their Fourier transform for signal and image processing applications. *IEEE Trans. Signal Process.* **52**(7), 2012–2030 (2004)
- Xu, D.; Mandic, D.P.: The theory of quaternion matrix derivatives. *IEEE Trans. Signal Process.* **63**(6), 1543–1556 (2015)
- Nefian, A. V.: Georgia Tech face database. *Georg. Inst. Technol.* (1999).
- Phillips, P.J.; Wechsler, H.; Huang, J.; Rauss, P.J.: FERET database and evaluation procedure for face-recognition algorithms. *Image Vis. Comput.* **16**(5), 295–306 (1998)
- Sangwine, S. J., Le Bihan, N.: *Quaternion and octonion toolbox for Matlab* (2013).
- Zhao, M.; Jia, Z.; Gong, D.: Improved two-dimensional quaternion principal component analysis. *IEEE Access* **7**, 79409–79417 (2019)