EXEMPLAR-EMBED COMPLEX MATRIX FACTORIZATION FOR FACIAL EXPRESSION RECOGNITION

Viet-Hang Duong1, Yuan-Shan Lee1, Jian-Jiun Ding2, Bach-Tung Pham1, Manh-Quan Bui3, Pham The Bao3, and Jia-Ching Wang1

1 Dept. of Computer Science and Information Engineering, National Central University, Jhongli, Taiwan
2 Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan
3 Fac. of Mathematics and Computer Sciences, University of Science, Ho Chi Minh City, Viet Nam

ABSTRACT

This paper presents an image representation approach which is based on matrix factorization in the complex domain and called exemplar-embed complex matrix factorization (EE-CMF). The proposed EE-CMF approach can very effectively improve the performance of facial expression recognition. Moreover, Wirtinger’s calculus was employed to determine derivatives. The gradient descent method was utilized to solve the complex optimization problem. Experiments on facial expression recognition verified the effectiveness of the proposed EE-CMF. It provides consistently better recognition results than standard NMFs.

Index Terms— Complex matrix factorization, facial expression, optimization, nonnegative matrix factorization

1. INTRODUCTION

Facial expressions contain a lot of information, such as feeling and cognitive motion. Facial expression recognition (FER) plays an important role in human communication. Age, ethnicity, gender, facial hair, the makeup style, gesture, occlusion, and environment lighting affect the performance of FER [1]. How to design an effective and robust system is a challenging topic in FER.

Feature extraction is a critical step of the FER system. Recently, many works have been done on subspace projection techniques for appearance-based feature extraction. In the subspace learning scheme, the new feature matrix is built which maps data points to a subspace. The popular subspace projection techniques, such as principal component analysis (PCA) [2], linear discriminant analysis (LDA) [3, 4, 12], and nonnegative matrix factorization (NMF) [5, 6], represent a facial image as a linear combination of low rank basis images. Lee and Seung [5, 6] found that NMF has the superior ability on parts-based representation. NMF is an unsupervised data-driven approach in which all elements of the decomposed matrix and the obtained matrix factors are forced to be nonnegative. The sparsity constraint can also be imposed on the cost function. For instance, Hoyer [7] proposed a sparse function and incorporated the sparseness into factorizing a nonnegative matrix to improve the obtained decompositions. Yuan and Oja [8] introduced a projective NMF (PNMF) which learns localized features.

Many modified NMFs have been developed to perform the FER task. Nikitidis et al. [9, 10] developed two extensions of the NMF that applied discriminant criteria as constraints, including the clustering based discriminant analysis (CDA) [11] and linear discriminant analysis (LDA) [12]. Lee and Chellappa suggested incorporating sparsity constraints to generate localized dictionaries from dense motion flow image sequences [13]. Apparently, most NMF frameworks require the addition of regularizers to improve FER performance. Moreover, nonnegative entries are usually compulsory for the data matrix in NMFs, which restrict the applications of NMF. Semi NMF and convex NMF algorithms have been proposed to deal with this limitation [14]. In particular, convex NMF algorithm (Con-NMF) further require that the basis vectors in NMF are convex or linear combinations of the mixed-sign data points. Besides, the interesting work of Liwicki et al. [15] showed the equivalence between the square Frobenius matrix norm in the complex field and the robust dissimilarity measure in the real field. These studies motivate us to propose a new model, named exemplar-embed complex matrix factorization (EE-CMF). In the proposed model, the real data is transformed to the complex domain and the complex data matrix is factorized under imitating Con-NMF framework. The objective function is minimized throughout an unconstrained complex optimization problem. In the real domain, a Con-NMF loss function is bounded optimization. Since the Cauchy-Riemann equation no longer holds, the standard complex derivative is unable to operate as usual. In complex domain, the Wirtinger calculus [16] provides an efficient tool to compute derivations and brings more advantages for complex optimization problem.

The main contributions of this work are summarized as follows.

1) An image analysis method on the complex domain, which is called EE-CMF, is proposed.
surprise sadness anger fear disgust happiness neutral

Fig. 1. Cropped face images of six facial expressions from the CK+ dataset [25].

surprise sadness anger fear disgust happiness neutral

Fig. 2. Cropped face images of six facial expressions from the JAFFE dataset [26].

where \( X' X = (X'X)^{+} - (X'X)^{-} \), \((a_{j})^{+} = \max \{0, a_{j}\}\), and \((a_{j})^{-} = \max \{0, -a_{j}\}\).

2.2. Wirtinger’s Calculus and Complex Optimization

[Definition 1] Let \( f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C} \) be a function of real variables \( x \) and \( y \) such that \( g(x, z) = f(x, y) \) where \( z = x + iy \) and \( g \) is analytic with respect to \( z \) and \( \dot{z} \). The two “partial derivative” operators \( \partial g/\partial z \) and \( \partial g/\partial \dot{z} \) are defined by:

\[
\frac{\partial g}{\partial z} = \frac{1}{2} \left( \frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right), \quad \frac{\partial g}{\partial \dot{z}} = \frac{1}{2} \left( \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right)
\]  

(5)

It is often referred to as Wirtinger’s derivative [18].

[Definition 2] If \( f \) is a real function of a complex matrix \( Z \) then the complex gradient matrix is given by [18]:

\[
\nabla f(Z) = 2 \frac{\partial f(Z)}{\partial Z} + i \frac{\partial f(Z)}{\partial \text{Re}(Z)} + i \frac{\partial f(Z)}{\partial \text{Im}(Z)}
\]  

(6)

How to find the direction of minimum rate of change will be stated in the following theorem.

[Theorem 1] Let \( g : \mathbb{C}^{n \times m} \times \mathbb{C}^{n \times m} \rightarrow \mathbb{R} \) be a real-valued function that maps complex matrices into the real domain. \( -\nabla_{\dot{Z}} g(Z, \dot{Z}) \) gives the direction where the function \( g \) has the minimum rate of change with respect to \( Z \) [19].

In fact, the first order Taylor series expansion for the real-differentiable function \( g(Z, \dot{Z}) \) has the form as:

\[
Ag(Z, \dot{Z}) = \langle \nabla_{\dot{Z}} g, \Delta Z \rangle + \langle \nabla_{\dot{Z}} g, \Delta Z \rangle = 2 \text{Re}\left( \langle \nabla_{\dot{Z}} g, \Delta Z \rangle \right)
\]  

(7)

According to the Cauchy-Bunyakovsky-Schwarz inequality [20], \( |\Delta Z' \nabla_{\dot{Z}} g| \leq \|\Delta Z\| \|\nabla_{\dot{Z}} g\| \). The equality holds when \( \Delta Z \) and \( \nabla_{\dot{Z}} g \) is collinear, i.e., the gradient \( \nabla_{\dot{Z}} g \) defines the direction of the maximum rate of change in \( g(\cdot, \cdot) \) with respect to \( Z \).

3. PROPOSED METHOD

3.1. Exemplar-Embed Complex Matrix Factorization (EE-CMF)

To get the complex data matrix \( Z \in \mathbb{C}^{n \times m} \), the original real data matrix \( X \) is first normalized and then transformed into

\[
X = \frac{X}{\|X\|_2}
\]  

(9)

The factors \( V \) and \( A \) are updated as follows [14]:

\[
V_j \leftarrow V_j \left[ ((X'X)^{+}A)^{+} + [(X'X)^{-}A]^{+} \right]
\]

\[
A_j \leftarrow A_j \left[ ((X'X)^{+}V)^{+} + [(X'X)^{-}AV]^{+} \right]
\]  

(10)

The basic matrix \( W \) contains \( K \) vectors which are linearly combined by the coefficients in \( V \) to represent the data. To solve (1), Lee and Seung [5, 6] provided the iteratively updating algorithms as follows:

\[
V_j \leftarrow V_j \frac{(W'X)^{+}}{W'WV'}; \quad W_j \leftarrow W_j \frac{(XV')^{+}}{(W'WV')} \]

(2)

To relax the constraint of nonnegative data, Ding et al. [14] proposed convex nonnegative matrix factorization (Con-NMF) where mixed-sign data matrices are applied. Con-NMF imposes a constraint that the column vectors of \( M \times N \) matrix \( X \) corresponding to an image with size \( a \) by \( b \) \((N=a \times b)\). The NMF problem is to find \( W \in \mathbb{R}^{a \times K} \) and \( V \in \mathbb{R}^{b \times K} \) that satisfies the following objective function:

\[
\min_{W \in \mathbb{R}^{a \times K}, V \in \mathbb{R}^{b \times K}} O_{\text{NMF}}(W, V) = \frac{1}{2} \| X - WV \|_F^2
\]  

(1)

where \( \| \cdot \|_F \) denotes the Frobenius norm. The equality holds when \( W \) and \( V \) are collinear, i.e., the gradient \( \nabla_{WV} \) defines the direction of the maximum rate of change in \( \nabla_{WV} \) with respect to \( W \).

The equality holds when \( W \) and \( V \) are collinear, i.e., the gradient \( \nabla_{WV} \) defines the direction of the maximum rate of change in \( \nabla_{WV} \) with respect to \( W \).
the complex field using Euler’s formula [21] by a mapping $\Omega$ from $\mathbb{R}^n$ to $\mathbb{C}^n$.

$$\Omega(x) = z = \frac{1}{\sqrt{2}} e^{i\alpha x} = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{i\alpha x_1} \\ \vdots \\ e^{i\alpha x_N} \end{bmatrix}$$

(8)

where $x_i$ denotes an $N$-dimensional vector, which corresponds to an input image and is sorted in the lexicographic order, $x_i(\cdot) \in [0,1]$ and $\alpha \in \mathbb{R}^+$. After obtaining the complex training data $Z$, we employ our proposed EE-CMF to factorize $Z$ into complex matrix factors. The basis matrix in EE-CMF is constructed by the linear combination of the complex training examples. Given the complex data matrix $Z \in \mathbb{C}^{n \times m}$, EE-CMF factorizes $Z$ into the encoding matrix $V \in \mathbb{C}^{n \times K}$ and the exemplar-embed basis matrix $W = ZW$ where $W \in \mathbb{C}^{m \times K}$. Therefore, the objective function of EE-CMF problem can be formulated as

$$\min_{W,V} O_{EE-CMF}(W,V) = \min_{W,V} \frac{1}{2} \|Z-ZWV\|_F^2$$

(9)

where

$$\|Z-ZWV\|_F^2 = \text{Tr}(Z-ZWV)^\dagger(Z-ZWV)$$

$$= \text{Tr}(Z^\dagger Z - V^\dagger W^\dagger Z^\dagger Z - V^\dagger W^\dagger Z^\dagger ZWV + V^\dagger W^\dagger Z^\dagger ZWV + V^\dagger W^\dagger Z^\dagger ZWV)$$

(10)

3.2. Optimal Solution

In order to solve the minimization problem, we use the complex gradient descent algorithm by exploiting Wirtinger’s calculus. It can be seen that (9) is a nonconvex minimization problem with respect to both variables $W$ and $V$. Therefore, it is impractical to obtain the optimal solution by the conventional method. Instead, the following scheme can be used to solve the problem in (9).

- First, fix $W$ and the objective function (9) is modified as a function of one variable $V$ as follows:

$$\min_{V} O(V) = \min_{V} \frac{1}{2} \|Z-ZWV\|_F^2$$

(11)

- Then, $W$ is updated based on the Moore–Penrose pseudoinverse [22], which is denoted by $\dagger$, and $W = (Z^\dagger Z)^\dagger$ with fixed $V$.

To solve the subproblem (11), the function $O(V)$ is treated as $O(V, V^\dagger)$ where

$$O(V, V^\dagger) = \frac{1}{2} \text{Tr}(Z^\dagger Z - (V^\dagger V)^\dagger W^\dagger Z^\dagger Z$$

$$- Z^\dagger ZWV + (V^\dagger V)^\dagger W^\dagger Z^\dagger ZWV)$$

(12)

According to Theorem 1, at a given iteration round $t$, the following update rule is employed:

$$V^{(t+1)} = V^{(t)} - 2\beta \nabla_v O(V^{(t)}, V^{(t)})$$

(13)

where $\beta$ is the learning step parameter for the $t^{th}$ iteration estimated by the Armijo rule [23]. From the Armijo rule, $\beta_t = \mu^t$, $0 < \mu < 1$, and $s$ is the first non-negative integer such that the following inequality is satisfied:

$$O(V^{(t+1)}, V^{(t+1)}) - O(V^{(t)}, V^{(t)}) \leq 2\sigma \text{Re} \left\{ \langle V_v \rangle O(V^{(t)}, V^{(t)}), V^{(t+1)} - V^{(t)} \rangle \right\}$$

(14)

The first order partial derivative with respect to $V^\dagger$ are evaluated as follows:

$$\nabla_v O(V, V^\dagger) = -W^\dagger Z^\dagger Z + W^\dagger Z^\dagger ZWV$$

(15)

The condition in (14) guarantees the decrease of the function value in each iteration. Finally, one can choose a pre-defined threshold $\varepsilon$ and set the stopping condition as follows:

$$\|V_v O(V, V^\dagger)\| \leq \varepsilon$$

(16)

4. EXPERIMENTS

In this section, we evaluated the performance of the proposed EE-CMF framework for FER. The classification capability of the derived encoding coefficient vector was compared with various NMF-based methods. We obtained the basic matrix $W_{tr}$ from $W_{tr} = Z_{tr}(V_{tr} V_{tr})^\dagger$ in the training phase. The test sample $z_{te}$ was encoded by $v_{te} = (Z_{te} W_{tr})^\dagger z_{te}$. Classification was performed by the nearest neighbor classifier after projection.

4.1. Data Description, Baselines, and Experiment Settings

The proposed model was evaluated on two publicly available databases: the Cohn-Kanade (CK) [25] and the JAFFE [26] datasets. There are seven facial expressions in these datasets, including one neutral state and six basic expressions that contains happiness, sadness, surprise, anger, disgust, and fear. Each facial image in two databases was cropped and resized to have fixed size of 32 × 32 pixels. Figures 1 and 2 show some of the images in the two datasets.

The proposed algorithm was compared to the following popular NMF algorithms: (1) basic NMF [6]; (2) semi-NMF [14]; (3) convex NMF [14] (Con-NMF); (4) weighted NMF (We-NMF) [27], which assigns binary weights to the data matrix; (5) Ne-NMF [28], which is an efficient solver that applies Nesterov’s optimal gradient method in the optimization process.

To satisfy (12), the rate of reducing the step size $\mu$ was set with the sufficient decrease condition at 0.01. The stopping criterion was as in (15) where the relative tolerance $\varepsilon$ was $10^{-4}$ or at most 10000 iterations.
In the test phase. Similarly, in the experimental results, the proposed EE framework has much better performance than other existing methods.

### 5. Conclusion

This work proposes the exemplar-embed complex matrix factorization (EE-CMF), a subspace learning framework, in the complex domain. After transforming training images into the complex domain, the basis matrix is constructed by the linear combination of the complex training examples. The gradient descent method with Wirtinger’s calculus was used to solve the complex matrix factorization problems. The proposed EE-CMF framework was tested on two facial expression datasets and very accurate recognition results are yielded. The proposed framework is much superior to the traditional and extension of NMF algorithms. Incorporating more constraints to widen applications and further improving the performance will be the future work. The complexification for tensor factorization is also a direction for future research.
6. REFERENCES
