

RESEARCH

Open Access



Projective complex matrix factorization for facial expression recognition

Viet-Hang Duong¹, Yuan-Shan Lee¹, Jian-Jiun Ding², Bach-Tung Pham¹, Manh-Quan Bui¹, Pham The Bao² and Jia-Ching Wang^{1,3*}

Abstract

In this paper, a dimensionality reduction method applied on facial expression recognition is investigated. An unsupervised learning framework, projective complex matrix factorization (proCMF), is introduced to project high-dimensional input facial images into a lower dimension subspace. The proCMF model is related to both the conventional projective nonnegative matrix factorization (proNMF) and the cosine dissimilarity metric in the simple manner by transforming real data into the complex domain. A projective matrix is then found through solving an unconstrained complex optimization problem. The gradient descent method was utilized to optimize a complex cost function. Extensive experiments carried on the extended Cohn-Kanade and the JAFFE databases show that the proposed proCMF model provides even better performance than state-of-the-art methods for facial expression recognition.

Keywords: Complex matrix factorization, Facial expression recognition, Nonnegative matrix factorization, Projected gradient descent

1 Introduction

Facial expression recognition (FER) plays an important role in many applications related to computer vision and pattern recognition, such as human computer interface, surveillance, and multimedia [1–5]. Intensive techniques have been proposed for FER problem [6–14]. Much attention is put on the facial action coding system (FACS) approach which attempts to decompose facial expressions into varied action units and facial expression could be recognized based on the mixture of action units [6, 7]. A new perspective of using deep neural networks [8–11] to extract powerful temporal features hidden in facial images is also an interesting approach for FER, in particular for dynamic facial expression recognition. The dimensionality reduction technique that reduces the size of the feature space has been widely utilized because of its effectiveness on feature representation [12–14]. It is known that the changes of local appearance (e.g., eyes, nose, mouth) are usually related to facial expression variations; meanwhile, the global

features that are related to the whole facial image may fail on expression analysis. Local facial components contain more discriminative information and outperform global features for face recognition [15]. A nonholistic representation and a low rank approximation of the data make better performance for an FER system [16].

Nonnegative matrix factorization (NMF) [12, 13] is among the most popular dimensionality reduction methods by its natural part-based representation ability. NMF decomposes the high-dimensional nonnegative data matrix into two low-rank nonnegative matrices with additional non-negativity constraints. Most NMF techniques estimate the linear subspace of the given data by the square of the Euclidean distance or the generalized Kullback-Leibler divergence. They work well when image noise is independent and identically distributed. However, for the data corrupted by outliers, the estimated subspace can be arbitrarily biased [17]. To overcome this drawback, researchers improved NMF in different distance metrics. Specifically, D. Kong et al. proposed a robust NMF by using L_2 and L_1 norms, where the noise was assumed to follow the Laplacian distribution [18]. Similarly, Earth Mover's distance (EMD) and the Manhattan distance were also suggested in the work of Sandler et al. [19] and N. Guan et al. [20], respectively.

* Correspondence: jcw@csie.ncu.edu.tw

¹Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan

³Faculty of Mathematics and Computer Science, University of Science, Ho Chi Minh, Vietnam

Full list of author information is available at the end of the article

Besides, integrating constraints or using different metrics to modify the structure is also another attractive strategy to extend NMF. Starting from the ideas of SVD and NMF, for instance, Yuan et al. proposed a novel method called projective nonnegative matrix factorization (proNMF) [21]. Instead of operating two parameter matrices as NMF, there is only one computed matrix in proNMF. The coefficient matrix is replaced by the inner product of the base and the input matrix. As a result, there exists much fewer free-learned parameters. Furthermore, in spite of having no any extra regularization term, proNMF is still able to learn more spatially localized and part-based representations of visual patterns.

Recently, many FER algorithms have adopted NMF. Buciu and Pitas employed NMF for FER by enforcing local constraints to create a local NMF (LMF) model [22]. Nikitidis et al. incorporated discriminant constraints to build the supervised NMF learning method [23, 24]. Based on clustering-based discriminant analysis (CDA), the algorithm in [25] efficiently decomposes the provided data to discriminant parts and successfully extended the well-known NMF to subclass discriminant NMF (SDNMF). R. Zhi et al. [26] provided another extended NMF which imposed both sparse constraints on the basis matrix and graph-preserving criterions to improve the classification performance. Interpreting expressive images into two subspaces including the identity subspace and the expression subspace was a new model investigated in dual subspace nonnegative matrix factorization (DSNMF) [27] and dual subspace nonnegative graph embedding [28]. It can be seen that, in order to extract representative expression facial features, most existing algorithms tried to integrate label information via different constraints to derive supervised NMF. Therefore, the overall performance highly depends on the expression label which is insufficient to characterize data by the variations of pose, illumination, etc.

How to develop an unsupervised learning algorithm that performs efficiently FER was aimed in this paper. Encouraged by the localized and part-based representations of proNMF as well as the challenges on choosing a suitable metric that quantifies the approximation error in the NMF model, we proposed a new dimensionality reduction approach, named projective complex matrix factorization (proCMF). The proCMF model attempts to learn the localized basis that highlights the salient features of facial expressions; meanwhile, the reconstruction error affected by outliers was also minimized as much as possible. Moreover, the underlying intrinsic cosine dissimilarity referenced from the equivalence between two different metrics is also exploited. We also adopt a cosine-based distance measure to define an explicit mapping function from the space of pixel intensity into a high-dimensional sphere where complex projection is performed. The Euler

formula was used for constructing isomorphism mapping from the real space to the complex space. In this context, the cosine dissimilarity for approximating real data was replaced by the Frobenius norm to evaluate the reconstruction error of complex approximation. After transforming the real data into the complex field, the obtained complex matrix was approximately factorized into a product of two low-rank matrices through an unconstrained complex optimization problem. The proposed method addresses the general problem of finding projections that enhance class separability without attached label information in the reduced dimensional space. There is no any extra regularization term in the objective functions that reduce the complication on solving the optimization problem. Furthermore, the representation of newly coming samples in the FER framework by using only the trained projection matrix is one of the important factors on decreasing the computational complexity.

In summary, the contributions of this paper are shown as follows.

First, we construct a new dimensionality reduction model for learning low rank projection in the complex domain. The adoption of cosine dissimilarity and the Frobenius norm significantly enhance the performance of the FER system.

Second, proCMF can be performed without limiting the sign of data. Our proposed method can be applied to both negative and positive data, which yields extension on real-world applications. Therefore, several operations that can extract complex features, such as the short-time Fourier transform, are going to be utilized directly instead of their absolute values (magnitude/power spectrogram) on sound processing.

To satisfy nonnegative requirement, NMF usually uses various strategies on minimizing a function, which leads to computational complexity. On the contrary, the significant superiority compared to NMF approaches of the proposed proCMF is to construct an unconstrained optimization problem that simplified the framework of extracting the basis and intrinsic features.

The rest of this paper is organized as follows. In Section 2, we recall some basic facts about proNMF. The motivation, formulation, and algorithms of the proposed model as well as its optimal solution are presented in Section 3. Some optimization techniques and computation methods are also described in this section. The convincing experimental results of the proposed approach on FER are depicted in Section 4. Finally, conclusion and future work are drawn in Section 5.

2 Projective nonnegative matrix factorization

The problem on ensuring the learned basis vectors to be part-based standard NMF was treated by projective nonnegative matrix factorization (proNMF) [29, 30]. ProNMF approximately factorizes a projection matrix \mathbf{P} into a positive low-rank matrix \mathbf{W} and its transpose \mathbf{W}^T such

that $\mathbf{X} = \mathbf{P}\mathbf{X}$. The proNMF optimization problem has the following form:

$$\min_{\mathbf{W} \geq 0} O_{proNMF}(\mathbf{W}) = \frac{1}{2} \|\mathbf{X} - \mathbf{W}\mathbf{W}^T\mathbf{X}\|_F^2. \quad (1)$$

It can be seen that the projection matrix is constructed only from the basic matrix, so the optimization problem in (1) mainly takes account of one variable \mathbf{W} . Using the popular multiplicative update algorithm [13, 29] to solve (1), one can update \mathbf{W} iteratively from

$$\mathbf{W} \leftarrow \mathbf{W} \frac{\mathbf{X}\mathbf{X}^T\mathbf{W}}{\mathbf{W}\mathbf{W}^T\mathbf{X}\mathbf{X}^T\mathbf{W} + \mathbf{X}\mathbf{X}^T\mathbf{W}\mathbf{W}^T\mathbf{W}} \quad (2)$$

3 Projective complex matrix factorization

3.1 Observations on cosine dissimilarity

Assume that we are given the representations of two images \mathbf{I}_1 and \mathbf{I}_2 that are written by the N -dimensional vector \mathbf{x}_i ($i = 1, 2$) in the lexicographic order. First, $\mathbf{x}_i \in \mathbb{R}^N$ is normalized to get $\mathbf{x}_i(c) \in [0, 1]$ where c is the element vector index or the vector spatial location. Then, \mathbf{x}_i is mapped into the \mathbb{R}^{2N} sphere by

$$\mathbf{Z}(\mathbf{x}_i) = \frac{1}{\sqrt{N}} \begin{bmatrix} \cos(\mathbf{x}_i)^T & \sin(\mathbf{x}_i)^T \end{bmatrix}^T \quad (3)$$

where

$$\begin{aligned} \cos(\mathbf{x}_i) &= [\cos(\mathbf{x}_i(1)), \cos(\mathbf{x}_i(2)), \dots, \cos(\mathbf{x}_i(N))]^T, \\ \sin(\mathbf{x}_i) &= [\sin(\mathbf{x}_i(1)), \sin(\mathbf{x}_i(2)), \dots, \sin(\mathbf{x}_i(N))]^T, \end{aligned}$$

and $\|\mathbf{Z}(\mathbf{x}_i)\| = 1$.

We have

$$\mathbf{Z}(\mathbf{x}_1)^T \mathbf{Z}(\mathbf{x}_2) = \frac{1}{N} \sum_{c=1}^N \cos(\mathbf{x}_1(c) - \mathbf{x}_2(c)). \quad (4)$$

Recall that the cosine distance measure between two vectors \mathbf{x} and \mathbf{y} is given by

$$d(\mathbf{x}, \mathbf{y}) = 1 - \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (5)$$

If the distance between $\mathbf{Z}(\mathbf{x}_1)$ and $\mathbf{Z}(\mathbf{x}_2)$ have the form

$$d(\mathbf{Z}(\mathbf{x}_1), \mathbf{Z}(\mathbf{x}_2)) = \frac{1}{2} \|\mathbf{Z}(\mathbf{x}_1) - \mathbf{Z}(\mathbf{x}_2)\|_F^2, \quad (6)$$

then it is equal to a cosine-based distance measure and

$$d(\mathbf{Z}(\mathbf{x}_1), \mathbf{Z}(\mathbf{x}_2)) = 1 - \frac{1}{N} \sum_{c=1}^N \cos(\mathbf{x}_1(c) - \mathbf{x}_2(c)). \quad (7)$$

It can be seen that, if $\mathbf{I}_1 \approx \mathbf{I}_2$, e.g., $\forall c, \mathbf{x}_1(c) - \mathbf{x}_2(c) \approx 0$, then $d(\mathbf{Z}(\mathbf{x}_1), \mathbf{Z}(\mathbf{x}_2)) \rightarrow 0$. This implies that if the two

images are unrelated, then their local elements are unmatched.

Moreover, the mapping function (3) from \mathbb{R}^N to \mathbb{R}^{2N} is equivalent to a mapping function $f: \mathbb{R}^N \rightarrow \mathbb{C}^N$ defined by:

$$f(\mathbf{x}_t) = \mathbf{z}_t = \frac{1}{\sqrt{2}} e^{i\alpha\pi\mathbf{x}_t} = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{i\alpha\pi\mathbf{x}_t(1)} \\ \vdots \\ e^{i\alpha\pi\mathbf{x}_t(N)} \end{bmatrix} \quad (8)$$

where the Euler's formula [31] is

$$e^{i\alpha\pi\mathbf{x}_t} = \cos(\alpha\pi\mathbf{x}_t) + i\sin(\alpha\pi\mathbf{x}_t) \quad (9)$$

Therefore, the cosine dissimilarity of a data pair in the input real space equals to the Frobenius distance of the corresponding data pair in the complex domain. It is known that the robustness of cosine dissimilarity in the real domain has been found to suppress outliers [32]. With the idea of utilizing different robust similarity metrics to extend NMF, we introduce a new dimensionality reduction method (proCMF) which relates to conventional proNMF and uses the cosine dissimilarity metric as a measurement of the reconstruction error. However, the complexity of optimizing the real function with the cosine distance is wisely addressed by converting to a complex optimization problem with the Frobenius norm which is described in the next sections.

3.2 Problem formulation

In this section, we formulate the problem of multi-variants data factorization within the framework of complex data decomposition. Given the sample dataset $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]$, $\mathbf{x}_i \in \mathbb{R}^N$, we convert the real data matrix $\mathbf{X} \in \mathbb{R}^{N \times M}$ to a complex matrix $\mathbf{Z} \in \mathbb{C}^{N \times M}$ by the mapping (8) and perform matrix factorization in this complex feature space.

The basic idea of proCMF is the coefficient of each data point \mathbf{z}_i ($i = 1, 2, \dots, M$) that lies within the subspace spanned by the column vectors of one projection matrix. The coefficient matrix $\mathbf{H} \in \mathbb{C}^{K \times M}$ is obtained by linear transformation from samples. More specifically, given a matrix $\mathbf{Z} \in \mathbb{C}^{N \times M}$, we need to find out two matrices $\mathbf{W} \in \mathbb{C}^{N \times K}$ and $\mathbf{H} \in \mathbb{C}^{K \times M}$ to minimize $\|\mathbf{Z} - \mathbf{W}\mathbf{H}\|_F^2$ s.t $\mathbf{H} = \mathbf{V}\mathbf{Z}$ where $\mathbf{V} \in \mathbb{C}^{K \times N}$ is the projection matrix. The proCMF objective function is as follows:

$$\min_{\mathbf{W}, \mathbf{V}} O_{proCMF}(\mathbf{W}, \mathbf{V}) = \min_{\mathbf{W}, \mathbf{V}} \frac{1}{2} \|\mathbf{Z} - \mathbf{W}\mathbf{V}\mathbf{Z}\|_F^2 \quad (10)$$

3.3 Complex-valued gradient decent method

It can be seen that (10) is a nonconvex minimization problem with respect to both variables \mathbf{W} and \mathbf{V} . Therefore, they are impractical to obtain the optimal solution. This NP-hard problem can be tackled by applying the

block coordinate descent (BCD) with two matrix blocks [33] to obtain a local solution by the following scheme:

Given an initial $\mathbf{W}^{(0)}$, we find the optimal solution $\mathbf{V}^{(t+1)}$ such that:

$$\begin{aligned} \mathbf{V}^{(t+1)} &= \arg \min_{\mathbf{V}} \text{O}_{\text{proCMF}}(\mathbf{W}^{(t)}, \mathbf{V}) \\ &= \frac{1}{2} \|\mathbf{Z} - \mathbf{W}^{(t)} \mathbf{V} \mathbf{Z}\|_F^2. \end{aligned} \quad (11)$$

Because of no nonnegative constraint, the basis can be updated simply by the Moore–Penrose pseudoinverse [34]

$$\mathbf{W}^{(t+1)} = \mathbf{Z}(\mathbf{V}^{(t+1)} \mathbf{Z})^\dagger. \quad (12)$$

To find optimal solutions of (11), we use the complex-valued gradient descent algorithm (CGD). Here, (11) is considered as a real-valued function that is needed to be minimized subject to the complex variable V . Generally, the problem of one scalar function of a complex variable (11) is defined as to solve the following unconstrained optimization problem:

$$\min_{\mathbf{V}} f(\mathbf{V}) \quad (13)$$

where

$$\begin{aligned} f(\mathbf{V}) &= \frac{1}{2} \|\mathbf{Z} - \mathbf{W} \mathbf{V} \mathbf{Z}\|_F^2 = \frac{1}{2} \text{Trace}(\mathbf{Z} - \mathbf{W} \mathbf{V} \mathbf{Z})^H (\mathbf{Z} - \mathbf{W} \mathbf{V} \mathbf{Z}) \\ &= \frac{1}{2} \text{Trace}(\mathbf{Z}^H \mathbf{Z} - \mathbf{Z}^H \mathbf{V}^H \mathbf{W}^H \mathbf{Z} - \mathbf{Z}^H \mathbf{W} \mathbf{V} \mathbf{Z} \\ &\quad + \mathbf{Z}^H \mathbf{V}^H \mathbf{W}^H \mathbf{W} \mathbf{V} \mathbf{Z}), \end{aligned} \quad (14)$$

and $(\cdot)^H$ is the matrix Hermitian operation.

Let $V = \text{Re}(V) + i\text{Im}(V)$ where i is the imaginary unit and $i^2 = -1$. Then, $f(\mathbf{V})$ can be viewed as a real bivariate function of its real and imaginary components.

In most of complex-variable optimization problems, the objective functions are the real-valued functions of complex arguments. They are not analytic on the complex plane and do not satisfy Cauchy-Riemann conditions. Brandwood's analytic theory [35] can be applied to overcome this common problem. Recall that if $f(\mathbf{V})$ satisfies Brandwood's analytic condition, i.e., $f(\mathbf{V})$ is analytic with respect to the complex-valued variable V and its complex conjugate \bar{V} where \mathbf{V} and $\bar{\mathbf{V}}$ are treated as independent variables, then the first-order Taylor expansion of $f(\mathbf{V}, \bar{\mathbf{V}})$ is as follows:

$$\Delta f = \langle \nabla_{\bar{\mathbf{Z}}} f, \Delta \mathbf{Z} \rangle + \langle \nabla_{\mathbf{Z}} f, \Delta \bar{\mathbf{Z}} \rangle = 2 \text{Re} \{ \langle \nabla_{\bar{\mathbf{Z}}} f, \Delta \mathbf{Z} \rangle \} \quad (15)$$

and the complex gradient of $f(\mathbf{V})$ is defined as

$$\nabla_{\bar{\mathbf{V}}} f(\mathbf{V}, \bar{\mathbf{V}}) = \frac{\partial f(\mathbf{V})}{\text{Re}(\mathbf{V})} + i \frac{\partial f(\mathbf{V})}{\text{Im}(\mathbf{V})}. \quad (16)$$

Algorithm 1: Complex-Valued Gradient Decent Method (CGD)

Input: \mathbf{Z} , $0 < \mu < 1$, $0 < \sigma < 1$

Output: \mathbf{V}

1. Initialize any feasible $\mathbf{V}^{(0)}$

Set $\beta^{(0)} = 1$

2. **repeat**

(a) Compute the gradient $\nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)})$

(b) **for** $t = 1, 2, \dots$

(b1) Assign $\beta^{(t)} \leftarrow \beta^{(t-1)}$

(b2) **if** $\beta^{(t)}$ satisfies (20)

repeat

$\beta^{(t)} = \beta^{(t)} / \mu$

until Either β_t does not satisfy (20)

or $\mathbf{V}(\beta^{(t)} / \mu) = \mathbf{V}(\beta^{(t)})$

else

repeat

$\beta^{(t)} = \beta^{(t)} \cdot \mu$

until $\beta^{(t)}$ satisfies (20)

(b3) Set $\mathbf{V}^{(t+1)} = \mathbf{V}^{(t)} - \beta^{(t)} \nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)})$

until the stopping criterion is satisfied

Therefore, the function $f(\mathbf{V})$ is treated as $f(\mathbf{V}, \bar{\mathbf{V}})$, where

$$f(\mathbf{V}, \bar{\mathbf{V}}) = \frac{1}{2} \text{Trace}[\mathbf{Z}^H \mathbf{Z} - \mathbf{Z}^H (\bar{\mathbf{V}})^T \mathbf{W}^H \mathbf{Z} - \mathbf{Z}^H \mathbf{W} \mathbf{V} \mathbf{Z} + \mathbf{Z}^H (\bar{\mathbf{V}})^T \mathbf{W}^H \mathbf{W} \mathbf{V} \mathbf{Z}] \quad (17)$$

The gradient of $f(\mathbf{V}, \bar{\mathbf{V}})$ with respect to \mathbf{V} is given by:

$$\nabla_{\bar{\mathbf{V}}} f(\mathbf{V}, \bar{\mathbf{V}}) = -\mathbf{W}^H \mathbf{Z} \mathbf{Z}^H + \mathbf{W}^H \mathbf{W} \mathbf{V} \mathbf{Z} \mathbf{Z}^H. \quad (18)$$

The gradient decent method for the unconstrained optimization problem in (13) builds a sequence $\{\mathbf{V}^{(t)}\}_{t \in \mathbb{N}}$ according to the following iterative form:

$$\mathbf{V}^{(t+1)} = \mathbf{V}^{(t)} - \beta^{(t)} \nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)}) \quad (19)$$

where $\beta^{(t)}$ is the step size variable, which is a small positive constant minimizing $f(\mathbf{V}^{(t)} - \beta^{(t)} \nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)}))$ over \mathbb{R} . In this paper, backtracking line search, which is also known as the Armijo rule [36], is used to estimate the step size. In this rule, $\beta^{(t)} = \mu^{k_t}$, where $0 < \mu < 1$ and k_t is the first non-negative integer k such that:

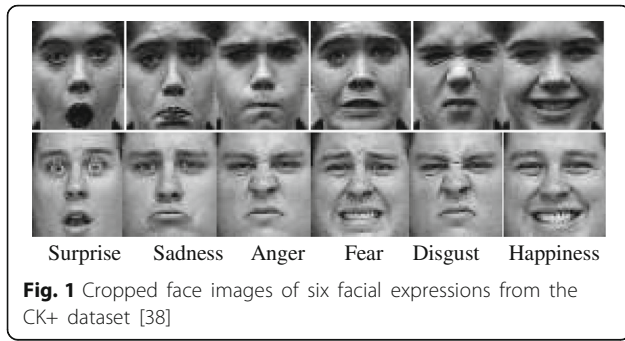


Fig. 1 Cropped face images of six facial expressions from the CK+ dataset [38]

$$f(\mathbf{V}^{(t+1)}, \bar{\mathbf{V}}^{(t+1)}) - f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)}) \leq 2\sigma \operatorname{Re} \left\{ \left\langle \nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)}), \mathbf{V}^{(t+1)} - \mathbf{V}^{(t)} \right\rangle \right\}. \quad (20)$$

There always exists a step length $\beta^{(t)}$, which is among $1, \mu^1, \mu^2, \dots$. A stationary point of (13) also exists among the limit points of $\{\mathbf{V}^{(t)}\}_{t \in \mathbb{N}}$ [37]. The iteration will be stopped when the solution is close to a stationary point. Practically, a common condition to check if a point $\mathbf{V}^{(t)}$ is close to a stationary point is:

$$\left\| \nabla_{\bar{\mathbf{V}}} f(\mathbf{V}^{(t)}, \bar{\mathbf{V}}^{(t)}) \right\|_F \leq \varepsilon \quad (21)$$

where ε is a pre-defined threshold.

The following Algorithm 2 summarizes the optimization process of the proposed proCMF model.

Algorithm 2: ProCMF

Input: \mathbf{Z}

Output: \mathbf{V}, \mathbf{W}

1. Initialize any feasible $\mathbf{W}^{(0)}, t=0$
2. repeat
 - (a) Use CGD (Algorithm 1) to update $\mathbf{V}^{(t+1)}$
 - (b) Update $\mathbf{W}^{(t+1)}$ by (12)
 - (c) $t \leftarrow t+1$

until the stopping criterion is satisfied

3. $\mathbf{V} = \mathbf{V}^{(t)}, \mathbf{W} = \mathbf{W}^{(t)}$

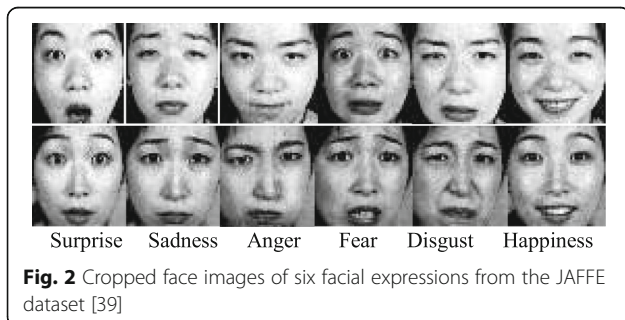


Fig. 2 Cropped face images of six facial expressions from the JAFFE dataset [39]

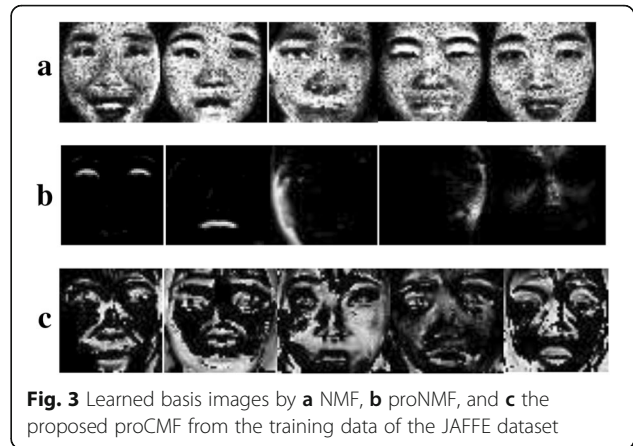


Fig. 3 Learned basis images by **a** NMF, **b** proNMF, and **c** the proposed proCMF from the training data of the JAFFE dataset

4 Experiments

We evaluated the proposed proCMF on FER. The classification capability of the derived encoding coefficient vectors was compared with various NMF-based methods.

Here, the minimization of the objective function $\frac{1}{2} \|\mathbf{Z}_{tr} - \mathbf{W}_{tr} \mathbf{V}_{tr} \mathbf{Z}_{tr}\|^2$ of proCMF was operated on the training phase that extracts new feature $\mathbf{H}_{tr} = \mathbf{V}_{tr} \mathbf{Z}_{tr}$ from the image training set \mathbf{Z}_{tr} . After creating a projection matrix \mathbf{V}_{tr} , for each new testing sample \mathbf{z}_{te} in a query face image \mathbf{Z}_{te} , we can easily obtain its corresponding projection \mathbf{h}_{te} such that $\mathbf{h}_{te} = \mathbf{V}_{tr} \mathbf{z}_{te}$. To make a decision about the facial expression class, the projected expression representation \mathbf{h}_{te} will be fed to the nearest neighbor classifier. Then, the Euclidean distance between \mathbf{h}_{te} and each training datum of \mathbf{H}_{tr} is computed. The tested image is assigned to the class with the closest training datum.

4.1 Data description

The proposed algorithm is tested on two well-known datasets, the extended Cohn-Kanade (CK+) [38] and the Japanese Female Facial Expression (JAFFE) [39] datasets

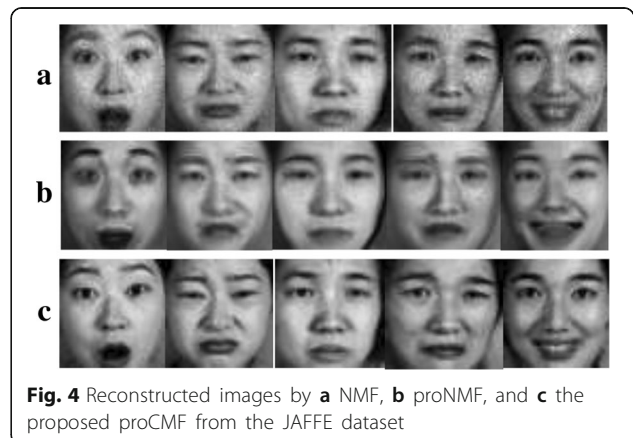


Fig. 4 Reconstructed images by **a** NMF, **b** proNMF, and **c** the proposed proCMF from the JAFFE dataset

Table 1 Facial expression recognition rate (%) using the CK+ dataset with different subspace dimensionalities (case 1: no. training = 1)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	94.86	85.41	73.78	41.28	85.06	94.30	66.94	<i>95.60</i>	30.87	32.35	48.82	87.00
30	96.09	90.99	78.74	58.99	91.07	89.30	78.93	<i>96.69</i>	40.93	71.97	56.84	92.98
40	<i>96.55</i>	93.88	82.21	73.55	94.17	90.27	80.99	96.01	45.37	92.65	62.54	94.34
50	<i>97.02</i>	94.50	84.84	80.27	94.75	91.80	86.78	96.71	51.94	93.18	66.94	95.08
60	<i>97.02</i>	95.06	83.66	87.40	94.92	92.34	88.02	96.57	55.45	94.88	72.44	95.31
70	<i>97.21</i>	95.18	85.31	90.42	95.62	93.31	89.05	96.71	55.19	95.25	73.45	94.77
80	<i>97.21</i>	95.93	85.89	91.57	95.58	93.27	91.12	96.84	60.29	94.70	75.12	94.65
90	<i>97.29</i>	95.95	86.53	92.42	95.87	93.26	91.53	96.82	62.46	94.74	77.87	94.30
100	<i>97.15</i>	96.03	88.55	92.03	95.62	94.17	91.53	96.86	78.60	94.65	78.99	93.68
Ave.	<i>96.71</i>	93.66	83.28	78.66	93.63	92.45	84.99	96.54	53.46	84.93	68.11	93.57

The best results are indicated in italics

of six “basic” facial expressions (happiness, sadness, surprise, anger, disgust, and fear).

The CK+ dataset consists of 593 video sequences from 123 subjects. Each video sequence shows distinct facial expressions. For each expression of a subject, the last five frames in the videos are selected and these frames were treated as static different facial expressions. Some samples in the CK+ dataset are shown in Fig. 1.

Altogether, the JAFFE dataset has 213 grayscale images of ten subjects posed 2–4 examples for each expressions. Figure 2 shows some figures in the JAFFE dataset.

Each cropped facial image in the datasets was isotropically scaled to the fixed size of 32×32 pixels.

4.2 Baselines and experiment settings

The proposed algorithm is compared to the following popular PCA and NMF algorithms: (1) basic NMF [4]; (2) projective NMF (proNMF) [9]; (3) convex NMF (conNMF) [40]; (4) weighted NMF (weiNMF), which assigns binary weights to the data matrix [41]; (5) NeNMF, which applies efficient Nesterov’s optimal gradient method in the optimization process [42]; (6) principal

component analysis (PCA) [43]; (7) graph preserving sparse nonnegative matrix factorization (GSNMF) [26]; (8)(9) unsupervised and supervised robust nonnegative graph embedding (uRNGE) and (sRNGE), which are robust nonnegative graph embedding methods that replace the L_2 -norm with the L_{21} -norm [44, 45]; and (10)(11) unsupervised and supervised robust semi-nonnegative graph embedding (usRNGE) and (ssRNGE), which are robust semi-nonnegative graph embedding methods that impose no constraint on the base matrix [44].

In the implementation of the complex valued conjugate gradient decent method with the Amijo rule (CGD), we set the decreasing rate of the step size μ to satisfy (19) with the sufficient decreasing condition at 0.01 and the stopping criteria of 10,000 times of iterations or that the error ε is smaller than 10^{-4} .

4.3 Visualization of learned basis components and reconstructed images

In order to visualize the basis and reconstructed images, we computed the complex matrix \mathbf{z} back to the real domain by applying the angle operator \angle , which returns

Table 2 Facial expression recognition rate (%) using the CK+ dataset with different subspace dimensionalities (case 2: no. training = 2)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	96.78	84.44	77.25	29.37	84.49	94.75	72.73	<i>98.04</i>	28.95	78.18	41.74	91.76
30	97.99	91.38	80.83	41.85	91.32	94.08	78.24	<i>98.26</i>	35.65	87.19	53.28	96.69
40	98.40	94.55	81.30	53.83	94.60	94.41	85.67	<i>98.71</i>	42.56	95.43	61.18	97.02
50	<i>98.57</i>	96.03	85.32	65.37	95.15	94.52	90.36	98.54	47.77	97.33	66.39	97.66
60	<i>98.68</i>	97.02	84.54	76.78	96.47	94.38	92.56	96.36	51.93	98.07	70.39	97.93
70	<i>98.76</i>	96.70	86.39	82.26	97.11	93.72	92.84	95.65	54.66	98.29	70.74	97.91
80	<i>98.59</i>	97.91	87.47	86.28	97.66	94.21	93.94	97.74	59.26	98.21	74.74	98.18
90	<i>98.76</i>	97.88	86.89	89.84	97.49	94.16	95.32	98.21	66.25	98.29	74.68	98.15
100	<i>98.79</i>	97.85	87.99	92.78	97.47	93.28	95.87	98.21	70.74	98.29	74.35	98.24
Ave.	<i>98.37</i>	94.86	84.22	68.71	94.64	94.17	88.61	97.75	50.86	94.36	65.28	97.06

The best results are indicated in italics

Table 3 Facial expression recognition rate (%) using the JAFFE dataset with different subspace dimensionalities (case 1: no. training = 1)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	71.33	68.11	50.00	45.53	63.36	66.85	51.05	68.11	22.38	31.68	27.62	60.77
30	71.19	70.84	56.01	49.09	68.32	61.05	57.34	70.07	22.10	60.77	33.78	62.31
40	71.47	71.68	57.83	48.60	68.95	63.28	58.74	69.23	28.18	64.27	39.79	60.28
50	70.14	71.12	60.21	52.17	69.02	61.82	57.34	69.09	32.66	63.50	43.15	62.73
60	69.65	69.79	57.34	46.36	72.38	63.71	57.34	70.28	36.78	64.41	47.41	61.82
70	70.07	26.15	60.56	27.34	69.16	62.52	57.34	70.98	42.24	71.40	51.19	70.77
80	69.09	16.01	62.03	27.90	23.50	62.87	NA	69.93	46.64	13.43	59.86	12.31
90	70.28	18.60	63.64	26.01	26.71	61.40	NA	70.91	49.37	13.15	64.55	15.38
100	70.56	52.94	61.54	35.59	15.25	63.36	NA	70.56	53.64	14.06	66.36	15.17
Ave.	70.42	51.69	58.80	39.84	52.96	62.98	57.62	69.91	37.11	44.07	48.19	46.84

the angle of a complex number. The inverse mapping f^{-1} of (8) gives the pre-image \mathbf{x} in the following form:

$$\mathbf{x} = \frac{\mathcal{L}\mathbf{z}}{\alpha\pi}. \quad (22)$$

Some basis images learned by NMF, proNMF, and proCMF from the JAFFE dataset are shown as in Fig. 3. One can see that the NMF bases are generally nearly holistic and less sparse than others bases, while proNMF extracted more localized and non-overlapped bases which correspond to several facial parts. Unlike small single regions as proNMF bases, the basic components of the proposed proCMF are combined from several localized regions to highlight specific local facial features which are salient areas on the face, such as the mouth, the nose, the eyes, and the eyebrows. This can be attributed to the fact that the proposed proCMF is more robust to extract discriminant facial features and has potential superior for FER. Figure 4 gives the reconstructed images of the proposed proCMF, NMF, and proNMF. The faces reconstructed by proCMF are much clearer than those reconstructed by NMF and proNMF. Due to outliers and noise, NMF and proNMF are hard to identify well test samples.

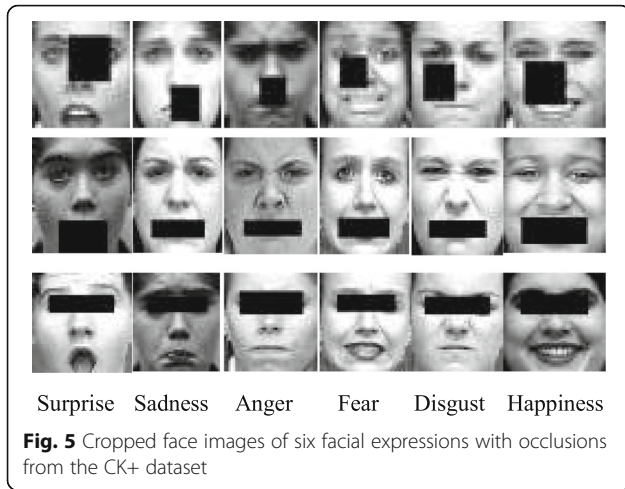
4.4 Facial expression recognition on CK+ dataset

Generally, the training set and the testing set were created by dividing the dataset into two subsets. On the CK+ dataset, we designed two experiments using different numbers of selected images for training and testing. In case 1, one image among five frames of each expression per subject is used for training and the remaining four are used for testing. In case 2, for each expression per subject, two collected images formed the training set and the remaining three formed the testing set. The average recognition rates of two cases versus different subspace dimensionalities are described in Tables 1 and 2, respectively. The best results are indicated in italics.

Most of the algorithms tend to get higher recognition rates when the subspace dimension increases. NMF performs better than proNMF and conNMF but worse than weiNMF and NeNMF. As mentioned in Section 4.3, the facial bases extracted by proNMF are spatially localized and lack expression-related information. These small basis regions of proNMF are clearly not appropriate for facial expression discrimination and result in unsatisfactory performance. The eigenface approach PCA, a well-known framework for dimensionality reduction, just rated at 86.80% in average achievement. The

Table 4 Facial expression recognition rate (%) using the JAFFE dataset with different subspace dimensionalities (case 2: no. training = 2)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	77.53	70.55	52.60	44.66	67.39	74.79	60.27	75.34	20.27	39.32	29.59	70.68
30	80.96	73.70	52.87	44.93	75.48	74.93	64.38	78.77	27.53	73.01	33.15	75.07
40	82.33	75.61	55.07	47.67	77.26	73.70	68.49	79.59	30.82	74.66	42.47	75.75
50	82.05	78.49	56.57	50.68	80.00	75.34	69.86	81.51	34.52	76.44	43.29	75.34
60	82.46	79.31	58.35	55.20	80.55	70.68	71.23	79.86	38.9	75.75	47.67	75.34
70	81.51	80.68	59.86	58.49	82.19	72.05	68.49	81.23	42.6	76.3	54.66	77.26
80	83.84	81.51	56.57	56.71	82.05	70.82	69.86	81.51	52.33	78.08	60.68	74.38
90	83.84	82.46	60.96	55.61	82.46	68.35	68.49	82.74	67.4	75.89	64.93	74.11
100	84.38	82.46	59.45	56.30	82.88	67.39	71.23	82.33	73.97	75.07	71.37	74.79
Ave.	82.10	78.31	56.92	52.25	78.92	72.01	68.03	80.32	43.15	71.61	49.76	74.75



supervised scheme of the NGE model herein is better than unsupervised ones in all cases, and no nonnegative constraint on the base matrix of semi-NMF has significant effect on FER. GSNMF and the proposed proCMF are superior to baseline methods. Their average performances are over 97% with the highest rate of 97.51% is achieved by the proposed proCMF algorithm.

4.5 Facial expression recognition on JAFFE dataset

Regarding to the JAFFE dataset, one image for each expression per person is taken randomly to construct the training data and the rest of the images are used to the test phase (case 1). Similarly, the case 2 was conducted by collecting two images for the training set. However, the JAFFE dataset is more challenging and the recognition rate is much lower than that achieved for the CK+ dataset. The recognition results are detailed in Tables 3 and 4. The proposed proCMF attained the best performance and reaches the average recognition rates of 82.10 and 70.42%, respectively, which are 1.15 ~ 36.13% higher than those of other methods. In case 1 of experiments,

when the number of learned basis K exceeds the size of the training data, the overfitting problem occurs on PCA and seminonnegative models.

4.6 Facial expression recognition on occluded CK+ images

As mentioned in previous sections, the proposed proCMF exhibits the best performance on facial expression recognition in both CK+ and JAFFE datasets. In this section, we aim to firm the robustness of the proposed method when the dataset is corrupted by partial occlusion. The occlusion experiments were conducted on the facial images in the CK+ dataset. We constructed three experimental assessments through varying the occluded position on the tested face images. The mouth and eyes occlusions are simulated by placing a mask at mouth or eyes position, respectively. The sheltering patches of size 70×70 were put randomly on the original image of size 640×490 to create the random occlusion case. Some samples from occluded CK+ images are shown in Fig. 5. In a similar way to the unocclusion scene, the five last frames of each video sequence were treated as static images. Two of them were processed to become occlusion images and collected to form the testing set, the remaining three images were used for training set. The detailed results of random, mouth, and eyes occlusions are reported in Tables 5, 6, and 7, respectively. It can be seen that the highest recognition rate is achieved repeatedly by our proposed method, proCMF. Moreover, proCMF performs much more robust and stable than other methods. In this occlusion situation, the proposed proCMF maintains its discriminating ability, while the performances of other methods decrease significantly. In order to witness this statement, we calculated the gap between the non-occlusion and the occlusion cases on the CK+ dataset. The smallest decline on the average recognition rate is achieved by proCMF and uRNGE (15.98%), followed by ssRNGE with 17.86%. Inversely, proNMF, NeNMF, and conNMF dropped by the biggest proportions of 34.43, 34.29, and 40.40%,

Table 5 Facial expression recognition rate (%) using the occluded CK+ with different subspace dimensionalities (case of occluded randomly images)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	66.82	50.62	34.38	20.04	48.60	68.97	53.31	50.50	21.78	45.54	26.9	59.13
30	73.43	58.39	35.87	24.59	58.97	63.88	60.33	75.21	22.93	52.02	32.69	70.7
40	79.42	62.27	43.93	26.16	63.43	61.07	62.81	74.09	28.55	58.93	35.33	72.85
50	80.95	65.29	44.05	27.52	65.00	59.71	66.12	68.22	28.06	62.85	39.21	73.26
60	82.23	70.37	45.91	27.19	67.85	55.54	68.18	59.01	34.17	68.06	40.95	73.47
70	83.51	70.33	46.61	26.82	72.36	52.98	69.01	67.85	34.92	73.18	45.08	76.07
80	83.39	73.31	49.88	27.81	72.40	47.23	73.55	65.62	38.60	75.79	46.61	78.55
90	84.96	73.39	52.64	27.73	73.31	39.09	77.69	72.98	41.78	78.10	48.97	79.71
100	86.03	75.25	51.36	29.34	74.38	37.60	78.51	70.08	40.54	75.58	54.79	80.08
Ave.	80.08	66.58	44.96	26.36	66.26	54.01	67.72	67.06	32.37	65.56	41.17	73.76

Table 6 Facial expression recognition rate (%) using the occluded CK+ with different subspace dimensionalities (case of occluded mouth)

No. Base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	76.28	55.21	40.83	21.74	53.36	62.75	61.57	61.57	24.30	46.28	31.98	60.80
30	81.92	63.39	45.29	25.87	56.26	59.68	70.66	74.73	30.83	55.19	37.69	67.48
40	77.85	68.18	47.85	31.69	65.35	63.52	69.83	77.80	29.67	63.34	43.06	72.03
50	85.21	67.58	50.66	33.64	69.78	65.53	70.66	75.74	32.48	74.26	45.87	77.20
60	85.87	71.61	48.26	37.48	69.19	69.72	76.45	62.51	37.85	79.81	47.52	80.63
70	84.85	73.42	47.60	43.22	76.27	71.25	77.69	75.09	41.65	84.77	48.93	82.10
80	85.21	74.91	49.92	45.79	76.62	63.05	79.34	74.20	45.95	87.66	51.82	84.41
90	86.03	78.29	50.00	49.30	78.93	58.74	79.34	77.86	47.36	89.73	53.80	86.36
100	87.60	81.90	49.83	50.91	78.87	43.80	81.82	75.03	46.42	89.02	58.93	88.12
Ave.	83.42	70.50	47.81	37.74	69.40	62.00	74.15	72.73	37.39	74.45	46.62	77.68

respectively. The recognition rates of the rest of the methods range from 20.84 to 27.54%. In overall, it can be stated that the proposed method works well not only in recognizing unoccluded faces but also in the occlusion case.

The confusion matrix is figured out in terms of unocclusion, eyes and mouth occlusion for the proposed proCMF. The merged results are given in Table 8. It is observed that, in case of no occlusion, happiness and neutral can be classified well with highest accuracy (100%), while the other five expressions are recognized with lower accuracy (96–99%). For occluded mouth images, sadness is recognized with the lowest accuracy (71.86%) since sadness is highly confused to disgust and surprise. In contrast, it seems that eyes occlusion most affect to recognizing disgust expression. There is a significant decreasing to 63.16% on discriminating disgust by the confusing with sadness and anger.

Regarding the evaluation of the influence of different parts on facial image to the expression recognition rates, we considered on both occlusion and unocclusion local regions. From the aligned images, we cropped fixed

regions on faces to create the testing set. Figure 6 shows the 12 local facial regions with the corresponding facial expression recognition accuracies. It is shown that the eye regions have more important information than the nose and mouth regions. In fact, in case of local region experiments, the eyes-nose regions obtain 60.77% in accuracy rate; meanwhile, the mouth-nose regions just reach 33.35%. The eye regions perform 30.71% accuracy which is better than the mouth regions 9.10%. In case of occlusion experiment settings, 62.82% is the average accuracy of all algorithms on recognizing the occluded mouth images and this rate is degraded to 59.9% for the occluded eyes images.

5 Conclusions

We explored a new dimensionality reduction model which employs an implicitly cosine dissimilarity metric by transforming the real data to the complex domain and sets up a complex projection matrix in the setting of an unbounded optimization problem. Under the simple framework without any added label information, the proposed method, proCMF, successfully extracted local

Table 7 Facial expression recognition rate (%) using the occluded CK+ with different subspace dimensionalities (case of occluded eyes)

No. base	proCMF	NMF	proNMF	conNMF	weiNMF	NeNMF	PCA	GSNMF	uRNGE	usRNGE	sRNGE	ssRNGE
20	75.50	58.93	38.00	21.43	50.41	70.43	38.43	76.39	28.93	25.62	25.80	65.58
30	78.02	69.70	47.41	24.09	62.57	65.29	49.17	71.72	24.88	38.10	37.01	71.19
40	77.52	73.88	51.95	28.93	67.00	69.01	54.55	70.48	26.94	42.89	41.15	75.97
50	86.86	72.26	55.32	31.94	67.83	64.76	57.02	73.38	35.45	45.21	42.80	81.88
60	83.26	74.38	51.90	39.43	71.66	65.11	56.20	62.46	39.09	48.51	48.64	85.06
70	84.46	78.29	62.00	39.32	75.62	66.35	56.61	69.54	44.38	51.98	50.41	83.53
80	80.50	75.93	63.67	45.51	78.39	59.15	60.74	73.49	46.61	50.74	53.31	87.60
90	80.41	82.64	63.62	45.69	79.75	44.86	64.88	79.99	49.75	53.97	57.62	90.20
100	84.05	81.17	63.53	45.45	79.75	44.27	66.53	82.05	52.94	59.67	60.33	87.43
Ave.	81.18	74.13	55.27	35.75	70.33	61.03	56.01	73.28	38.77	46.30	46.34	80.94

Table 8 Confusion matrix of 7-class facial expression recognition using proCMF on CK+ dataset with unocclusion and eyes and mouth occlusions (%)

	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Neutral	
Anger	96.08	1.96	0	1.96	0	0	0	Unocclusion
	94.12	2.94	0	2.94	0	0	0	Mouth
	100	0	0	0	0	0	0	Eyes
Disgust	0	98.25	0	0	1.75	0	0	Unocclusion
	2.63	92.11	0	0	0	5.26	0	Mouth
	13.16	63.16	0	0	23.68	0	0	Eyes
Fear	2.22	0	97.78	0	0	0	0	Unocclusion
	0	6.67	86.67	3.33	0	3.33	0	Mouth
	0	0	100	0	0	0	0	Eyes
Happiness	0	0	0	100	0	0	0	Unocclusion
	4.00	6.00	0	86.00	0	4.00	0	Mouth
	4.00	2.00	0	90.00	4.00	0	0	Eyes
Sadness	2.08	0	0	0	97.92	0	0	Unocclusion
	3.13	12.50	0	0	71.86	9.38	3.13	Mouth
	9.37	0	0	0	90.63	0	0	Eyes
Surprise	0	0	1.28	0	0	98.72	0	Unocclusion
	0	0	0	0	1.92	98.08	0	Mouth
	1.92	0	5.77	0	9.62	82.69	0	Eyes
Neutral	0	0	0	0	0	0	100	Unocclusion
	0	0	0	0	0	0	100	Mouth
	0	0	0	0	16.67	0	83.33	Eyes

facial features which are salient areas in facial expression formation. The experiments validate convincingly that the proposed proCMF algorithm can much outperform the popular NMF algorithms for facial expression recognition. The proposed model shows its potential on dealing with diverse scenarios that involve a whole face and an occluded face. In the future, we will develop the current model by incorporating more constraints and improving optimization methods to enhance the FER performance. In order to firm the effectiveness of the proposed method, we will extend evaluation on spontaneous facial expressions and apply directly complex spectrogram features on sound processing, especially the problems about sound source separation.

Acknowledgements

This research was financially supported in part by the National Central University, Taiwan, through the NCU International Student Scholarship. The authors would like to thank the anonymous reviewers for their valuable comments that significantly improved the quality of this paper.

Authors' contributions

VHD made the main contributions to the conception and tracking algorithms' design, as well as drafting the article. BTP contributed to the design of experiments. YSL, JJD, and PTB provided the technical advices. MQB checked the manuscript and contributed to the rearrangement of the materials. JJW offered critical suggestions on the algorithms' design, provided significant revising for important intellectual content, and gave final approval of the current version to be submitted. All authors read and approved the final manuscript.

Consent for publication

Informed consent was obtained from all authors included in the study.

Competing interests

The authors declare that they have no competing interests.

Ethical approval

All data and procedures performed in paper were in accordance with the ethical standards of research community. This paper does not contain any studies with human participants or animals performed by any of the authors.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

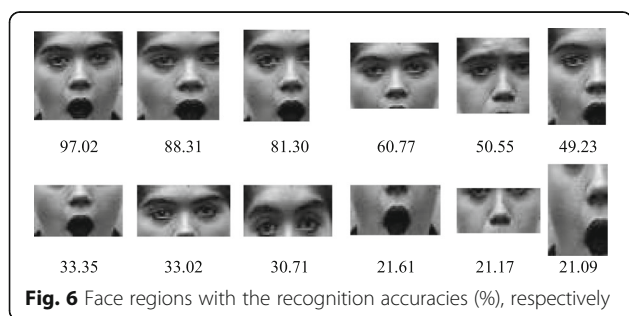


Fig. 6 Face regions with the recognition accuracies (%), respectively

Author details

¹Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan. ²Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan. ³Faculty of Mathematics and Computer Science, University of Science, Ho Chi Minh, Vietnam.

Received: 28 June 2017 Accepted: 11 December 2017

Published online: 02 February 2018

References

- Zeng, M, Pantic, GI, Roisman, TS, Huang, A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(1), 39–58 (2009)
- FS Abousaleh, TK Lim, WH Cheng, NH Yu, MA Hossain, MF Alhamid, A novel comparative deep learning framework for facial age estimation. *EURASIP J Image Video Process.* 2016;2016(1):47.
- TH Tsai, WH Cheng, CW You, MC Hu, AW Tsui, HY Chi, Learning and recognition of on-premise signs (OPSS) from weakly labeled street view images. *IEEE Trans. Image Process.* **23**(3), 1047–1059 (2014)
- SC Hidayati, CW You, WH Cheng, KL Hua, Learning and recognition of clothing genres from full-body images. *IEEE Trans. Cybern.* (2017)
- WH Cheng, CW Wang, JL Wu, Video adaptation for small display based on content recomposition. *IEEE Trans. Circuits Syst. Video Technol.* **17**(1), 43–58 (2007)
- P. Ekman, W. V. Friesen, and J. C. Hager, *Facial Action Coding System: The Manual on CD ROM*. A Human Face, 2002.
- S Kaltwang, S Todorovic, M Pantic, in *IEEE CVPR*. Latent trees for estimating intensity of facial action units (2015), pp. 296–304
- H Jung, S Lee, J Yim, S Park, J Kim, in *IEEE ICCV*. Joint fine-tuning in deep neural networks for facial expression recognition (2015), pp. 2983–2991
- JS Riera, K Srinivasan, KL Hua, WH Cheng, MA Hossain, MF Alhamid, Robust RGB-D hand tracking using deep learning priors. *IEEE Trans. Circuits Syst. Video Technol.* (2017)
- K He, X Zhang, S Ren, J Sun, in *IEEE CVPR*. Deep residual learning for image recognition (2016), pp. 770–778
- K Zhang, Y Huang, Y Du, L Wang, Facial expression recognition based on deep evolutionary spatial-temporal networks. *IEEE Trans. Image Process.* **26**(9), 4193–4203 (2017)
- D Lee, H Seung, Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 755–791 (1999)
- D Lee, H Seung, in *NIPS*. Algorithms for non-negative matrix factorization (2000), pp. 556–562
- B Wu, T Mei, WH Cheng, YD Zhang, *Unfolding Temporal Dynamics: Predicting Social Media Popularity Using Multi-Scale Temporal Decomposition*. The 30th AAAI Conference on Artificial Intelligence (AAAI - Association for the Advancement of Artificial Intelligence, Phoenix, 2016), pp. 12–17
- B Heisele, P Ho, J Wu, T Poggio, Face recognition: component-based versus global approaches. *Comput. Vis. Image Underst.* **91**, 6–12 (2003)
- J Ellison, D Massaro, Featural evaluation, integration, and judgment of facial affect. *J. Exp. Psychol. Hum. Percept. Perform.* **23**(1), 213–226 (1997)
- A Cichocki, R Zdunek, S Amari, in *Int. Conf. Independent Component Analysis and Signal Separation*. Csiszar's divergences for non-negative matrix factorization: Family of new algorithms (2006), pp. 32–39
- D Kong, C Ding, H Huang, in *ACM Int. Conf. Information and Knowledge Management*. Robust nonnegative matrix factorization using $L_{2,1}$ norm (2011), pp. 673–682
- R Sandler, M Lindenbaum, Nonnegative matrix factorization with earth mover's distance metric for image analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(8), 1590–1602 (2011)
- N Guan, D Tao, Z Luo, J Shawe-Taylor. (2012), MahNMF: Manhattan Non-negative Matrix Factorization. [Online]. Available: <https://arxiv.org/abs/1207.3438v1>. Accessed 14 July 2012
- Z Yuan, E Oja, in *Scandinavian Conf. Image Analysis*. Projective nonnegative matrix factorization for image compression and feature extraction (2005), pp. 333–342
- I Buciu, I Pitas, in *ICPR*. Application of non-negative and local non negative matrix factorization to facial expression recognition (2004), pp. 288–291
- S Nikitidis, A Tefas, N Nikolaidis, I Pitas, in *ICIP*. Facial expression recognition using clustering discriminant non-negative matrix factorization (2011), pp. 3001–3004
- S Nikitidis, A Tefas, N Nikolaidis, I Pitas, Subclass discriminant nonnegative matrix factorization for facial image analysis. *Pattern Recogn.* **45**(12), 4080–4091 (2012)
- X Chen, T Huang, Facial expression recognition: a clustering-based approach. *Pattern Recogn. Lett.* **24**(9), 1295–1302 (2003)
- R Zhi, M Flierl, Q Ruan, W Kleijin, Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition. *IEEE Trans. Syst. Man Cybern.* **41**(1), 38–52 (2011)
- Y Tu, C Hsu, in *IEEE ICPR*. Dual subspace nonnegative matrix factorization for person-invariant facial expression recognition (2012), pp. 2391–2394
- H Kung, Y Tu, C Hsu, Dual subspace nonnegative graph embedding for identity-independent expression recognition. *IEEE Trans. Inf. Forensics Secur.* **10**(3), 626–638 (2015)
- Z Yang, Z Yuan, J Laaksonen, Projective non-negative matrix factorization with applications to facial image processing. *Int. J. Pattern Recognit Artif Intell.* **21**(8), 1353–1362 (2007)
- Z Yang, E Oja, Linear and nonlinear projective non-negative matrix factorization. *IEEE Trans. Neural Netw.* **21**, 734–749 (2010)
- M Moskowitz, *A Course in Complex Analysis in One Variable* (World Scientific Publishing Co., Singapore, 2002), p. 7
- S Liwicki, G Tzimiropoulos, S Zafeiriou, M Pantic, Euler principal component analysis. *Int. J. Comput. Vis.* **101**(3), 498–518 (2013)
- J Kim, Y He, H Park, Algorithms for nonnegative matrix and tensor factorizations: a unified view based on block coordinate descent framework. *Glob. Optim.* **58**, 285–319 (2013)
- J Barata, M Hussein, The Moore–Penrose pseudoinverse: a tutorial review of the theory. *Physics* **42**, 146–165 (2012)
- DH Brandwood, A complex gradient operator and its application in adaptive array theory. *IEEE Proc. F* **130**(1), 11–16 (1983)
- C Lin, Projected gradient methods for non-negative matrix factorization. *Neural Comput.* **19**, 2756–2779 (2007)
- D Bertsekas, On the Goldstein-Levitin-Polyak gradient projection method. *IEEE Trans. Automat. Contr.* **21**(2), 174–184 (1976)
- P Lucey, JF Cohn, T Kanade, J Saragih, Z Ambadar, I Matthews, in *Proceedings of IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*. The extended Cohn–Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression (2010), pp. 94–101
- M Lyons, S Akamatsu, M Kamachi, J Gyoba, in *Proceedings of IEEE Conf. Autom. Face Gesture Recog.* Coding facial expressions with Gabor wavelets (1998), pp. 200–205
- C Ding, T Li, M Jordan, Convex and semi-nonnegative matrix factorizations. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(1), 45–55 (2010)
- D Wang, T Li, C Ding, in *Proceedings of IEEE ICDM*. Weighted feature subset nonnegative matrix factorization and its applications to document understanding (2010), pp. 541–550
- N Guan, D Tao, Z Luo, B Yuan, NeNMF: an optimal gradient method for non-negative matrix factorization. *IEEE Trans. Signal Process.* **60**(6), 2882–2898 (2012)
- K Fukunaga, *Statistical Pattern Recognition* (Academic Press Professional, Inc., San Diego, 1990)
- J Yang, S Yang, Y Fu, X Li, T Huang, in *Proceedings of IEEE CVPR*. Nonnegative graph embedding (2008), pp. 1–8
- H Zhang, ZJ Zha, Y Yang, S Yan, TS Chua, Robust (semi) nonnegative graph embedding. *IEEE Trans. Image Process.* **23**(1), 2996–3012 (2014)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com