國立中央大學
National Central University

IMAGE REPRESENTATION USING SUPERVISED AND UNSUPERVISED LEARNING METHODS ON COMPLEX DOMAIN

Manh-Quan Bui

Viet-Hang Duong

Yung-Hui Li

Tzu-Chiang Tai

Jia-Ching Wang

Outline

1. • Introduction
2. • Euler formula for space transformation
3. • The proposed methods
4. • Experiments
5. • Conclusions

1. Introduction (1)

- Image representation (IR) system is to transform the input signal into a new representation which reduces its dimensionality and explicates its latent structures
- Matrix factorization (MF) has been gained considerable popularity for data exploration, analysis and interpretation
  - Principal component analysis (PCA)
  - Linear discriminant analysis (LDA)
  - Nonnegative matrix factorization (NMF)
  - Complex matrix factorization (CMF)
- NMF and CMF are a relative novel paradigm for dimensionality reduction. NMF factorizes a nonnegative data matrix into nonnegative matrix factors; meanwhile CMF is not only limited to sign of the data but also works on complex domain.

1. Introduction (2)

Matrix factorization on complex domain for feature learning

- Unsupervised PCMF
- Supervised DPCMF

Facial expression recognition

- The goal of this paper is to develop two complex matrix factorization models projective complex matrix factorization (PCMF) and discriminant projective complex matrix factorization (DPCMF)

2. Euler's Formula for Space Transformation (1)

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*

$$\mathbf{x}_i \ (i=1,2);\ \mathbf{x}_i \in \mathbb{R}^N;\ \mathbf{x}_i(c) \in [0,1]$$

The map $Z: \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$

$$Z(\mathbf{x}_i) = \frac{1}{\sqrt{N}}\left[\cos(\mathbf{x}_i)^T \ \sin(\mathbf{x}_i)^T\right]^T$$

$$\cos(\mathbf{x}_i) = \left[\cos(\mathbf{x}_i(1)), \cos(\mathbf{x}_i(2)), \ldots, \cos(\mathbf{x}_i(N))\right]^T$$

$$\sin(\mathbf{x}_i) = \left[\sin(\mathbf{x}_i(1)), \sin(\mathbf{x}_i(2)), \ldots, \sin(\mathbf{x}_i(N))\right]^T$$

$$\left\|Z(\mathbf{x}_i)\right\| = 1$$

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

The mapping function from $\mathbb{R}^N$ to $\mathbb{R}^{2N}$ is equivalent to a mapping function

$$f: \mathbb{R}^N \rightarrow \mathbb{C}^N$$

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

where the Euler's formula is

$$e^{i\alpha\pi\mathbf{x}_i} = \cos(\alpha\pi\mathbf{x}_i) + i\sin(\alpha\pi\mathbf{x}_i).$$

V. H. Duong, Y. S. Lee, J. J. Ding, B. T. Pham, M. Q Bui, P. T. Bao, J. C. Wang, "Exemplar-Embed Complex Matrix Factorization for Facial Expression Recognition", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1837-1841, U.S.A 2017

2. Euler's Formula for Space Transformation (2)

The cosine dissimilarity distance of a pair of data in the input **real space** equal to **the Euclidean distance** of the corresponding data in **complex field**

If $d(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) = \frac{1}{2}\|Z(\mathbf{x}_1) - Z(\mathbf{x}_2)\|_F^2$ → $\mathbb{C}^N$

Then $d(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) = 1 - \frac{1}{N}\sum_{c=1}^{N}\cos(\mathbf{x}_1(c) - \mathbf{x}_2(c))$ → $\mathbb{R}^{2N}$

If $I_1 \approx I_2$ e.g. $\forall c, \mathbf{x}_1(c) - \mathbf{x}_2(c) \approx 0$

Then $d(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) \to 0$

if the two images are unrelated, then their local elements be unmatched.

$$Z: \mathbb{R}^N \to \mathbb{R}^{2N} \Leftrightarrow f: \mathbb{R}^N \to \mathbb{C}^N$$

Cosine Dissimilarity

Frobenious norm

3. The Proposed Method (1)

We denote the given training data set as a complex matrix $\mathbf{D} = [\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_M]$, where $\mathbf{D}_i \in \mathbb{C}^N$, and M is the total number of training samples; the number of vectors in the i^{th} class as n_i, and the number of classes as C.

Projective Complex Matrix Factorization (PCMF)

Model of PCMF:

$$\mathbf{D} \approx \mathbf{B}\mathbf{B}^{H}\mathbf{D}$$

Objective function of PCMF

$$\min_{\mathbf{B}} f_{\text{PCMF}}(\mathbf{B}) = \min \frac{1}{2}\left\|\mathbf{D} - \mathbf{B}\mathbf{B}^{H}\mathbf{D}\right\|_F^2$$

$$\left\|\mathbf{D} - \mathbf{B}\mathbf{B}^{H}\mathbf{D}\right\|_F^2 = Trace[(\mathbf{D}^{H}\mathbf{D} - 2\mathbf{D}^{H}\mathbf{B}\mathbf{B}^{H}\mathbf{D} + \mathbf{D}^{H}\mathbf{B}\mathbf{B}^{H}\mathbf{B}\mathbf{B}^{H}\mathbf{D})]$$

3. The Proposed Method (2)

Discriminant Projective Complex Matrix Factorization (DPCMF)

Integrating the Fisher's criterion [6] into PCMF to utilize the label information

→ minimizing the distance between any two samples of the same class

→ maximizing the distance of the samples in different classes

Objective function of DPCMF

$$\min_{\mathbf{B}} f_{\text{DPCMF}}(\mathbf{B}) = \min \frac{1}{2}\left\|\mathbf{D} - \mathbf{B}\mathbf{B}^H\mathbf{D}\right\|_F^2 + \frac{1}{2}\alpha Trace(\mathbf{B}^H(\lambda\mathbf{S}_w - \mathbf{S}_b)\mathbf{B})$$

$$\mathbf{S}_{\mathrm{w}} = \frac{1}{C}\sum_{i=1}^{C}\frac{1}{n_i}\sum_{j=1}^{n_i}(\mathbf{d}_j - \mu_i)(\mathbf{d}_j - \mu_i)^T$$

$$\mathbf{S}_b = \frac{1}{C(C-1)}\sum_{i=1}^{C}\sum_{j=1}^{C}(\mu_i - \mu_j)(\mu_i - \mu_j)^T$$

[6] S. Zafeiriou, A. Tefas, I. Buciu, I. Pitas, "Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification," *IEEE Transactions on Neural Networks*, vol. 17, pp. 683-695, 2006

3. The Proposed Method (3)

Gradient descent method for optimal solutions

The update rules

$$\mathbf{B}^{(t+1)} = \mathbf{B}^{(t)} - 2\beta_t \nabla_{\mathbf{B}^*} f(\mathbf{B}^{(t)})$$

The Wirtinger's calculus

$$\nabla_{\mathbf{B}^*} f(\mathbf{B}) = \frac{\partial f(\mathbf{B})}{\partial(\mathrm{Re}\,\mathbf{B})} + i\frac{\partial f(\mathbf{B})}{\partial(\mathrm{Im}\,\mathbf{B})}$$

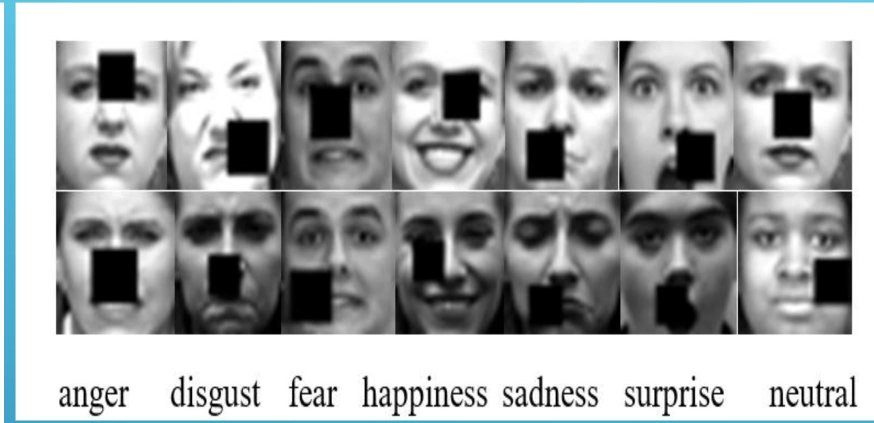
$$\nabla_{\mathbf{B}^*} f_{PCMF}(\mathbf{B}) = -2\mathbf{D}^H\mathbf{D}\mathbf{B} + \mathbf{B}\mathbf{B}^H\mathbf{D}\mathbf{D}^H\mathbf{B} + \mathbf{D}\mathbf{D}^H\mathbf{B}\mathbf{B}^H\mathbf{B}$$

$$\nabla_{\mathbf{B}^*} f_{DPCMF}(\mathbf{B}) = -2\mathbf{D}^H\mathbf{D}\mathbf{B} + \mathbf{B}\mathbf{B}^H\mathbf{D}\mathbf{D}^H\mathbf{B} + \mathbf{D}\mathbf{D}^H\mathbf{B}\mathbf{B}^H\mathbf{B} + \alpha(\lambda S_w - S_b)\mathbf{B}$$

4. Experiments (1)

Facial Expression Recognition

Sample images from the CK+ and JAFFE dataset



Occluded images from the CK+ dataset

Baseline

(1) EE-CMF: Exemplar-embed complex matrix factorization

(2) PNMF: Projective nonnegative matrix factorization

(3) DPNMF: Discriminant projective nonnegative matrix factorization

(4) NMF: Nonnegative matrix factorization.

(5) DNMF: Discriminant nonnegative matrix factorization.

4. Experiments (2)

PCMF and DPCMF for Facial Expression Recognition

Feature learning

- Learn on training dataset $\mathbf{D}_{train}$

→ dictionary $\mathbf{B}_{train}$ ($\mathbf{D}_{train}=\mathbf{B}_{train}(\mathbf{B}_{train})^H\ \mathbf{D}_{train}$)

- Get the learned feature $\mathbf{C}_{train} = (\mathbf{B}_{train})^H\ \mathbf{D}_{train}$
- Project the tested samples $\mathbf{D}_{test}$ onto the feature space and obtaining the encode $\mathbf{C}_{test} = (\mathbf{B}_{train})^H\ \mathbf{D}_{test}$.

Recognizing

- Feed $\mathbf{C}_{train}$ and $\mathbf{C}_{test}$ as input data to a NN classifier.

4. Experiments (3)

TABLE I
FACIAL EXPRESSION RECOGNITION RATE (%) USING THE CK DATASET WITH DIFFERENT SUBSPACE DIMENSIONALITIES

No. Base	DPCMF	PCMF	EE-CMF	DPNMF	PNMF	DNMF	NMF
20	**96.78**	95.95	95.43	95.89	77.25	24.24	85.41
30	**97.31**	97.00	92.25	96.32	80.83	25.19	90.99
40	**97.15**	96.96	91.24	96.69	81.30	28.26	93.88
50	97.11	**97.15**	95.06	96.65	85.32	28.68	94.5
60	**97.31**	97.17	96.14	97.02	84.54	38.74	95.06
70	**97.44**	97.05	96.59	96.80	86.39	38.49	95.18
80	**97.27**	97.21	96.74	97.07	87.47	38.93	95.93
90	**97.25**	97.19	96.63	96.96	86.89	40.48	95.95
100	**97.20**	97.11	96.78	97.17	87.99	45.76	96.03
Ave.	**97.20**	96.98	95.21	96.73	84.22	34.31	93.66

4. Experiments (5)

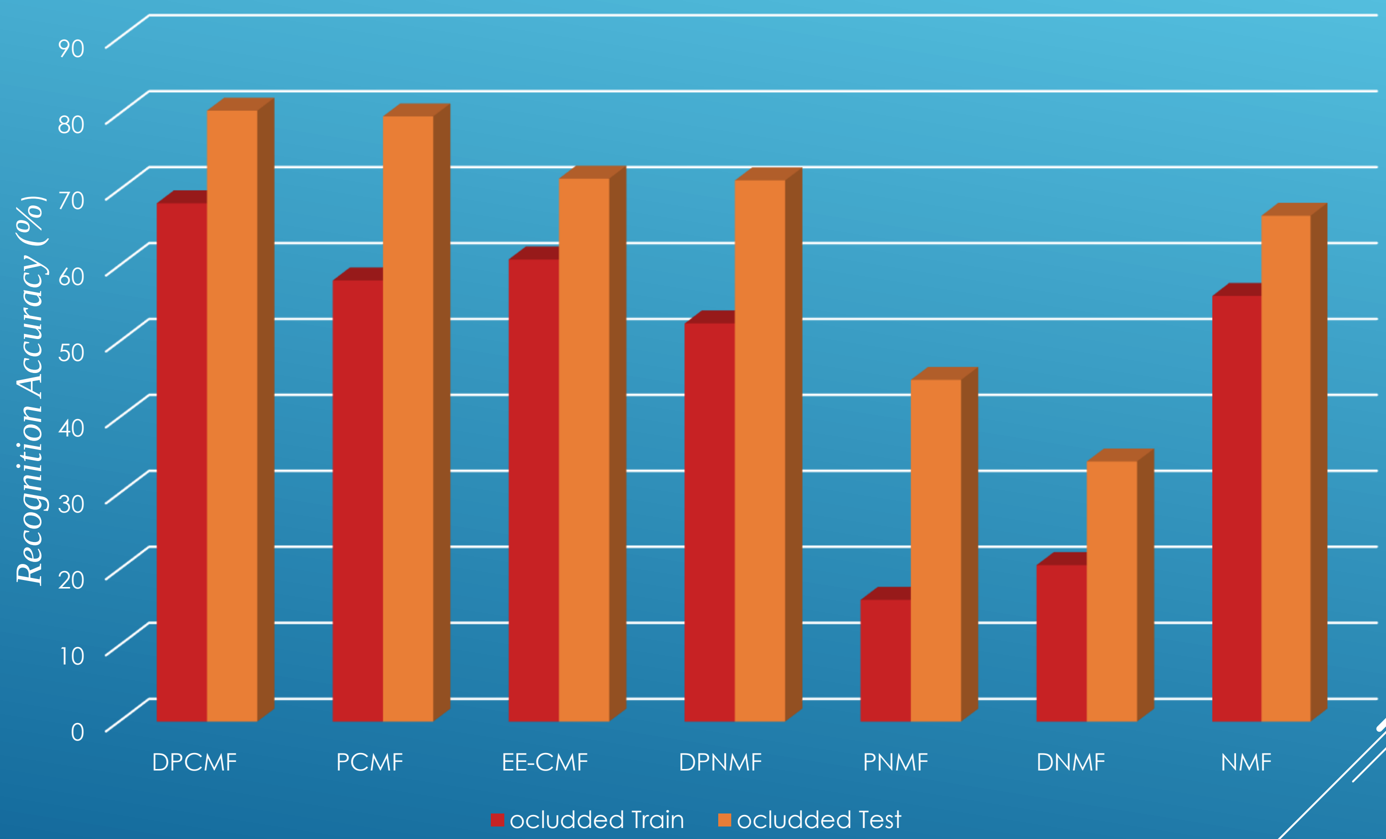
TABLE III

FACIAL EXPRESSION RECOGNITION RATE (%) USING THE JAFFE DATASET WITH DIFFERENT SUBSPACE DIMENSIONALITIES

No. Base	DPCMF	PCMF	EE-CMF	DPNMF	PNMF	DNMF	NMF
20	**70.42**	69.58	66.99	63.01	50.00	15.31	65.24
30	**70.98**	69.02	66.36	66.78	56.01	14.90	68.11
40	**72.31**	71.26	**72.31**	69.58	57.83	15.10	70.84
50	**72.66**	71.40	72.03	69.58	60.21	15.59	71.68
60	72.24	**72.45**	**72.45**	70.49	57.34	15.66	71.12
70	**72.45**	72.38	72.31	70.07	60.56	15.24	69.79
80	**73.01**	71.75	72.59	71.75	62.03	15.38	26.15
90	72.17	**72.80**	71.68	71.54	63.64	15.45	16.01
100	72.73	72.24	**73.57**	72.31	61.54	17.48	18.60
Ave.	**72.11**	71.43	71.14	69.46	58.80	15.57	53.06

4. Experiments (8)

Experiments on Occlusion CK+ Images



5. Conclusions

- This paper has presented two efficient algorithms for robust image representation system. The novel approaches take advantages on complex matrix factorization to learn subspace. The combination with Fisher's criteria provides a supervised model which is reliable and stable to extract the meaningful features and make the classification task much easier.
- Without limiting the sign of data, the developed methods are able to be applied on real-world applications, particularly the field of complex-valued data processing, such as communication and acoustic, etc..
- Future works: extending the proposed approaches to the nonlinear representation and also testing their performance on various type of dataset.

THANK YOU
FOR YOUR ATTENTION

Q & A