2018 17th International Symposium on Distributed Computing and Applications for Business Engineering and Science

Constrained Multi-view NMF with Graph Embedding for Face Clustering

Bin Qian
Traffic Management Research Institute of the Ministry of Public Security, Wuxi, China
311062198@njust.edu.cn

Xiguang Gu
Traffic Management Research Institute of the Ministry of Public Security, Wuxi, China
guxiguang@nuaa.edu.cn

Zhenqiu Shu
School of Computer Engineering, Jiangsu University of Technology, Changzhou, China
shuzhenqiu@163.com

Xiaobo Shen
School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
njust.shenxiaobo@gmail.com

***Abstract*—Face clustering that aims to group faces from the same people is a key component in face tagging and attribute analysis. Nonnegative matrix factorization (NMF) has shown competitiveness for clustering, but lacks of discrimination in practical tasks. In this paper, we propose a constrained multi-view NMF method with graph embedding (GCMNMF) for face clustering. GCMNMF incorporates the graph constraint and label constraint into an unified framework. GCMNMF aims to seek latent discriminative representations for multiple views, and maintain the within-view geometric structure simultaneously. In addition, an iterative optimization algorithm based on multiplicative rules is developed to efficiently solve GCMNMF. Experimental results on two real-world datasets demonstrate the effectiveness of the proposed method on face clustering tasks.**

***Keywords*-nonnegative matrix factorization; multi-view learning; face clustering; graph; constrained**

I. INTRODUCTION

Multi-view face clustering has recently attracted a lot of attention in face tagging and video retrieval tasks[1][2]. Multiple views, e.g., color, shape and texture features can provide complementary information for the same face. Integrating these multiple views appropriately has been shown to outperform using only single view[3][4]. Until now many multi-view clustering methods have been proposed. Among them, there are two main clustering categories: spectral based method and subspace based method.

Spectral based methods are mainly extended from single-view clustering methods[5]. Co-training[6][7] trains the classifiers by maximizing the mutual agreement on two distinct views of the unlabeled data. Co-regularized spectral clustering[8] implicitly combines graphs from multiple views to achieve better clustering. Subspace based multi-view clustering methods first project multi-view data into a common low-dimensional subspace and then apply any clustering method such as K-means[9]. A typical subspace method is Canonical Correlation Analysis (CCA)[10], which analyzes linear correlations among multiple views. Nonnegative matrix factorization (NMF)[11] that obtains interpretable representations with the nonnegative constraint is a popular subspace method for clustering[12][13]. However, most existing NMFs are designed only for single view, which limits its use in many real-world applications. A naive way to exploit multiple views is to concatenate all the views into a single one, and then apply NMF to the single feature. However, this simple approach ignores the differences of statistical properties among different views and usually cannot obtain satisfactory results.

Recently, several NMF variants have been applied to multi-view clustering, and have already achieved promising results. In [14], collective NMF (ColNMF) is proposed for relational learning. ColNMF treats multi-view clustering as a latent space searching problem, and decomposes each view into two matrices, i.e., projection matrix and shared coefficient matrix. Liu *et al*. [15] propose multi-view NMF (MNMF) by posing a novel normalization strategy. For the semi-supervised scenario, Wang *et al*. [16] propose semi-supervised multi-view NMF (SMNMF), which takes the label information as additional hard constraint. NMF assumes that data points are sampled from euclidean space, thus it fails to exploit the geometric structure. Zhang *et al*. [17] propose a graph regularized multi-view NMF (GMNMF) to consider the manifold information. In GMNMF, the graph regularization is imposed on the latent representation and the parameters are set automatically according to the data.

Motivated by recent progress in NMF, we propose an improved multi-view NMF for face clustering by exploiting two constraints, including graph-based and label information from few labeled faces. The basic idea is to seek latent discriminative representations for multiple views, and to maintain the within-view geometric structure simultaneously. GCMNMF can substantially improve clustering by considering both with-view and cross-view correlations. An efficient optimization scheme is developed. Experiments on two real-world world datasets validate the effectiveness of the proposed method.

In the following, we first briefly review some related works. Then, we present the proposed method and report the experimental results. Finally, we conclude the paper.

978-1-5386-7445-1/18/$31.00 ©2018 IEEE
DOI 10.1109/DCABES.2018.00036

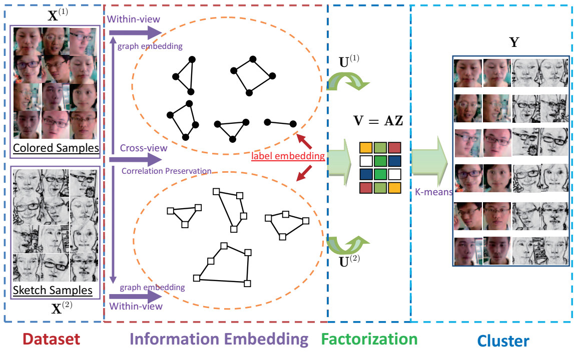


Figure 1. The flowchart of the proposed GCMNMF. First, multiple features are extracted from original images. Then the features are further factorized by the proposed multi-view learning model, and graph structure and partial label constraints are incorporated. Finally, the shared coefficient matrice is used for clustering.

II. Related works

This section briefly introduces a basic multi-view NMF method, which is also named collective matrix factorization (ColNMF) [14]. Some definitions are first provided.

Given a data set $\mathbf{X} = \{(\mathbf{x}_i^1, \mathbf{x}_i^2, \cdots, \mathbf{x}_i^m, y_i), i = 1, \cdots, n\}$, where $\mathbf{x}_i = (\mathbf{x}_i^1, \mathbf{x}_i^2, \cdots, \mathbf{x}_i^m)$ is the i^{th} example, $\mathbf{x}_i^q \in \mathrm{R}^{p_q}$ is the instance of the i^{th} example in the q^{th} view, m is the number of views, and y_i is its cluster label. ColNMF aims at clustering $\mathbf{x}_i$ into it's corresponding cluster label y_i through each single view's factorization with the shared coefficient matrix. The objective can be formulated as

$$\min_{\mathbf{U}^q, \mathbf{V}} F(\mathbf{U}^q, \mathbf{V}) = \sum_{q=1}^{m} \|\mathbf{X}^q - \mathbf{U}^q \mathbf{V}^T\|^2, s.t.\ \mathbf{U}^q \geq 0, \mathbf{V} \geq 0, \tag{1}$$

where $\mathbf{X}^q$ represents all the samples of the q^{th} view, $\mathbf{U}^q$ represents the projection matrix of the q^{th} view, and $\mathbf{V}$ is the shared coefficient matrix.

III. Approach

ColNMF learns a shared view representation. However, it fails to discover the geometrical structure of inner-view space. On the other hand, partial label priori information is not considered in ColNMF, which limits its use in real-world semi-supervised applications. This section introduces the proposed constrained multi-view NMF with graph embedding model (GCMNMF) to avoid these limitations.

A. Formulation

1) Multi-view NMF Clustering Framework with Graph Embedding: In multi-view learning, it is critical to analyze the relationships within and without views. It is commonly known that if data described in different views are related to similar person, they are expected to share a certain common structure[18]. On the other hand, the instances in the same view have certain latent geometric structure[19]. In order to make the fusion of different views meaningful, we optimize the following problem:

$$\begin{aligned}\min_{\mathbf{U}^q,\mathbf{V}^q,\mathbf{V}^*} &= \sum_{q=1}^{m} \|\mathbf{X}^q - \mathbf{U}^q(\mathbf{V}^q)^T\|_F^2 + \lambda^q\|\mathbf{V}^q - \mathbf{V}^*\|_F^2 \\ &+ \mu\sum_{q=1}^{m} \lambda^q \mathrm{Tr}((\mathbf{V}^q)^T\mathbf{L}^q\mathbf{V}^q), \\ s.t.\ &\mathbf{U}^q \geq 0, \mathbf{V}^q \geq 0, \mathbf{V}^* \geq 0, \|\mathbf{U}^q_{*,j}\| = 1,\end{aligned} \tag{2}$$

where $\mathbf{U}^q_{*,j}$ denotes the j^{th} column of $\mathbf{U}$. $\mathbf{L}^q = \mathbf{D}^q - \mathbf{W}^q$ denotes the Laplacian matrix of the q^{th}, where $\mathbf{D}^q$ is a diagonal matrix and $\mathbf{D}^q_{j,j} = \sum_l \mathbf{W}^q_{jl}$.

2) Discriminative Label Embedding: Assuming the first l data points are labeled with c classes, then an indicator matrix $\mathbf{C}$ can be constructed, where $c_{ij} = 1$ if v_i is labeled with j^{th} class; or $c_{ij} = 0$ otherwise. Then, the label constraint matrix $\mathbf{A}$ can be defined as follows:

$$\mathbf{A} = \begin{pmatrix} \mathbf{B}_{l\times c} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{n-l} \end{pmatrix}, \tag{3}$$

To incorporate label information, we introduce an auxiliary matrix $\mathbf{Z}$ with $\mathbf{V} = \mathbf{AZ}$, which can guarantee that data from the same person have the same representation.

3) Overall Objective Function: Note that the constraint $\|\mathbf{U}^q_{*,j}\| = 1$ in Eq. (2) will make the optimization computation difficult. In order to simplify the optimization process, we introduce auxiliary variables $\mathbf{Q}^q$ for $\mathbf{U}^q$, where $\mathbf{Q}^q$ is defined as:

$$\mathbf{Q}^q = Diag(\sum_j \mathbf{U}^q_{j,1}, \sum_j \mathbf{U}^q_{j,2}, \cdots, \sum_j \mathbf{U}^q_{j,r}), \tag{4}$$

where $Diag(\cdot)$ denotes a diagonal matrix.

Considering Eq. (3) and Eq. (4), the problem in Eq. (2) is equivalent to the following optimization problem:

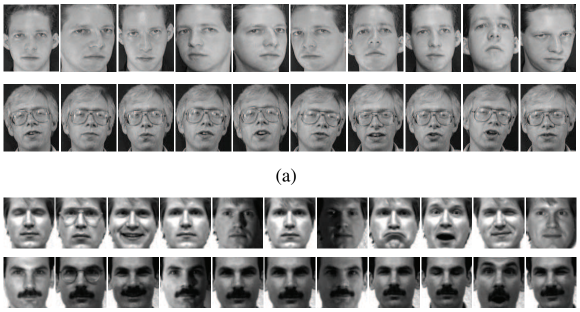
$$\begin{aligned}\min_{\mathbf{U}^q,\mathbf{Z}^q,\mathbf{Z}^*} &= \sum_{q=1}^{m} \|\mathbf{X}^q - \mathbf{U}^q(\mathbf{Z}^q)^T\mathbf{A}^T\|_F^2 + \lambda^q\|\mathbf{Z}^q\mathbf{Q}^q - \mathbf{Z}^*\|_F^2 \\ &+ \mu\sum_{q=1}^{m} \lambda^q \mathrm{Tr}((\mathbf{Z}^q)^T\mathbf{A}^T\mathbf{L}^q\mathbf{A}\mathbf{Z}^q), \\ s.t.\ &\mathbf{U}^q \geq 0, \mathbf{Z}^q \geq 0, \mathbf{Z}^* \geq 0.\end{aligned} \tag{5}$$

B. Optimization

To minimize the objective function in Eq. (5), we adopt iterative updating procedure.

1) Fixing $\mathbf{Z}^*$, update $\mathbf{U}^q$ and $\mathbf{Z}^q$: When $\mathbf{Z}^*$ is given, each view is independent. For simplicity, $\mathbf{U}$, $\mathbf{Z}$ and $\mathbf{Q}$ represent $\mathbf{U}^q$, $\mathbf{Z}^q$ and $\mathbf{Q}^q$. The Lagrange function for each view is as follow:

$$\begin{aligned}J &= \|\mathbf{X} - \mathbf{U}(\mathbf{Z})^T\mathbf{A}^T\|_F^2 + \lambda^q\|\mathbf{ZQ} - \mathbf{Z}^*\|_F^2 \\ &+ \mu\lambda^q\mathrm{Tr}((\mathbf{Z})^T\mathbf{A}^T\mathbf{LAZ}) + \mathrm{Tr}(\Phi\mathbf{U}^T) + \mathrm{Tr}(\Psi\mathbf{Z}^T).\end{aligned} \tag{6}$$



(a)

(b)

Figure 2. Some sample images from two face databases. (a) ORL database. (b) Yale database.

Taking partial derivative of J with respect to $\mathbf{U}$ and $\mathbf{Z}$, we can derive the following updating rules:

$$\mathbf{U}_{ik} = \mathbf{U}_{ik} \frac{(\mathbf{XAZ})_{ik} + \lambda^q \sum_{j=1}^{n-l+c} \mathbf{Z}_{jk}\mathbf{Z}^*_{jk}}{(\mathbf{UZ}^T\mathbf{A}^T\mathbf{AZ})_{ik} + \lambda^q \sum_{f=1}^{p_q} \mathbf{U}_{fk} \sum_{j=1}^{n-l+c} \mathbf{Z}^2_{jk}},$$
$$\mathbf{Z}_{jk} = \mathbf{Z}_{jk} \frac{(\mathbf{A}^T\mathbf{X}^T\mathbf{U})_{jk} + \lambda^q \mathbf{Z}^*_{jk} + \lambda^q \mu (\mathbf{A}^T\mathbf{WAZ})_{jk}}{(\mathbf{A}^T\mathbf{AZU}^T\mathbf{U})_{jk} + \lambda^q \mathbf{Z}_{jk} + \lambda^q \mu (\mathbf{A}^T\mathbf{DAZ})_{jk}}. \tag{7}$$

When computing $\mathbf{U}^q$ and $\mathbf{Z}^q$, we first compute $\mathbf{U}^q$ and then normalize column vectors of $\mathbf{U}^q$ and $\mathbf{Z}^q$ using $\mathbf{Q}^q$ as:

$$\mathbf{U}^q \leftarrow \mathbf{U}^q(\mathbf{Q}^q)^{-1}, \quad \mathbf{Z}^q \leftarrow \mathbf{Z}^q\mathbf{Q}^q. \tag{8}$$

2) Fixing $\mathbf{U}^q$ and $\mathbf{Z}^q$, update $\mathbf{Z}^*$: When $\mathbf{U}$ and $\mathbf{Z}$ are computed over each view, we take the derivative of loss function (6) over $\mathbf{Z}^*$ and get close-form solution to $\mathbf{Z}^*$:

$$\mathbf{Z}^* = \frac{\sum_{q=1}^{m} \lambda^q \mathbf{Z}^q \mathbf{Q}^q}{\sum_{q=1}^{m} \lambda^q}. \tag{9}$$

After several iterations, the loss value can converge. Finally, we can obtain the representative features $\mathbf{V}^* = \mathbf{AZ}^*$. After we obtain the latent matrix $\mathbf{V}^*$, the cluster label $\mathbf{Y}$ is computed by K-means algorithm in this paper.

IV. Experiments

This section evaluates the proposed methods by performing face clustering on two public image databases: ORL database and Yale database [20]. Some sample images are shown in Fig.2. For simplicity, raw pixel values $\mathbf{x}^1 \in R^{1024}$ and the local binary pattern feature $\mathbf{x}^2 \in R^{59}$ are extracted for further multi-view data fusion. To evaluate the effectiveness and efficiency, we present quantitative evaluations of our proposed GCMNMF, and compare it with some related methods: ConcatNMF[21], ColNMF[14], MNMF[15], SMNMF[16], and GMNMF[17]. Additionally, NMFs with single view (SV1, SV2) are also adopted as two baselines for comparison. We utilize K-means to cluster the low-dimensional shared representation and set the number of clusters as the classes of faces. The clustering performance is evaluated by clustering Accuracy (AC), which has been widely used for clustering. For the semi-supervised label embedding scene, we randomly choose 20% images from each person as the available label information, and use them to construct the label constraint matrix. Finally, we mix the labeled images and unlabeled images as a whole for face clustering tasks[22].

Table 1 and 2 report the clustering results on ORL and Yale databases, respectively. From the experimental results, we have the following conclusions:

(1) GCMNMF always outperforms two single-view NMFs, which validates that GCMNMF can explicitly integrate the visual information among different views and improve the face clustering performance.

(2) We observe that GCMNMF outperforms SMNMF on both databases. This is mainly attributed to the fact that the graph embedding technique is applied. It shows that graph Laplacian regularizer can effectively reveal the intrinsic geometrical structure within each view.

(3) GCMNMF is superior to GMNMF. GCMNMF considers the semi-supervised label as additional information, while GMNMF is unsupervised. It implies that available semi-supervised priori information plays an important role for face clustering.

V. Conclusion

This paper proposed an improved NMF algorithm by considering local geometrical structure and partial label information for multi-view clustering. Our model considers both the inner-view and inter-view relatedness among multi-view data. For the formulated non-convex objective function, we propose an alternative update scheme. The experimental results on face clustering validate its superiority over single-view and multi-view NMF methods.

Acknowledgment

This paper is supported by national key research and development program of China (Grant No. 2017YFC0804806), national natural science foundation of China (Grant No. 61603159) and natural science foundation of Jiangsu province (Grant No. BK20160293).

References

[1] P. Huang, Y. Wang, M. Shao, "A new method for multi-view face clustering in video sequence," IEEE International Conference on Data Mining Workshops, 2008, pp. 869-873.

[2] X. Cao, C. Zhang, H. Fu, S. Liu, H. Zhang, "Diversity-induced multi-view subspace clustering," International Conference on Computer Vision and Pattern Recognition, 2015, pp. 586-594.

[3] X. B. Shen, F. M. Shen, Q. S. Sun, Y. Yang, Y. H. Yuan, H. T. Shen, "Semi-paired discrete hashing: learning latent hash codes for semi-paired cross-view retrieval," IEEE Transactions on Cybernetics, 2017, 47: 4275-4288.

[4] X. B. Shen, Q. S. Sun, Y. H. Yuan, "A unified multiset canonical correlation analysis framework based on graph embedding for multiple feature extraction," Neurocomputing, 2015, 148: 397-408.

Table I
CLUSTERING PERFORMANCE ON ORL DATABASE (AC%)

c	SV1	SV2	ConcatNMF	ColNMF	MNMF	SMNMF	GMNMF	GCMNMF
3	61.76	79.33	75.41	77.00	79.88	84.98	81.02	**86.21**
4	60.46	79.13	76.82	78.02	80.97	85.79	82.33	**86.99**
5	63.58	79.60	76.89	78.86	82.04	85.90	82.09	**88.01**
6	56.71	75.67	73.90	75.23	78.75	83.51	80.39	**85.33**
7	52.89	74.57	68.09	70.93	74.21	79.14	76.90	**82.12**
8	51.07	74.06	67.12	68.07	74.57	78.91	77.31	**80.71**
9	50.88	73.17	60.12	64.21	72.99	77.00	74.84	**79.84**
10	51.89	73.75	61.93	64.69	71.08	75.28	74.45	**76.96**
Avg	56.16	76.16	70.03	72.13	76.81	81.31	78.67	**83.27**

Table II
CLUSTERING PERFORMANCE ON YALE DATABASE (AC%)

c	SV1	SV2	ConcatNMF	ColNMF	MNMF	SMNMF	GMNMF	GCMNMF
3	55.90	66.36	67.91	67.12	70.32	76.17	71.88	**79.90**
4	52.53	60.80	62.38	61.86	65.97	71.04	68.92	**76.24**
5	45.32	55.09	58.79	56.91	60.65	67.61	62.54	**72.09**
6	40.86	52.50	55.12	56.77	59.33	65.96	62.18	**63.42**
7	34.77	48.90	50.31	52.30	58.93	63.01	59.82	**60.11**
8	35.80	47.95	50.22	48.06	53.79	57.88	55.13	**54.31**
9	37.91	48.84	49.81	47.83	50.11	54.28	53.16	**58.77**
10	35.12	47.05	48.90	47.45	50.04	54.10	51.93	**51.60**
Avg	42.28	53.44	56.59	54.79	58.64	63.76	60.70	**64.56**

[5] V. R. D. Sa, "Spectral clustering with two views," ICML Workshop on Learning with Multiple Views, 2005, pp. 20-27.

[6] K. Nigam, R. Ghani. "Analyzing the effectiveness and applicability of co-training," International Conference on Information and Knowledge Management, 2000, pp. 86-93.

[7] A. Kumar, H. D. Iii, "A co-training approach for multi-view spectral clustering," International Conference on Machine Learning, 2011, pp. 393-400.

[8] A. Kumar, P. Rai, "Co-regularized multi-view spectral clustering," International Conference on Neural Information Processing Systems, 2011, pp. 1413-1421.

[9] X. Wang, X. Guo, Z. Lei, C. Zhang, S. Z. Li, "Exclusivity-consistency regularized multi-view subspace clustering," IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 923-931.

[10] X. B. Shen, Q. S. Sun, "A novel semi-supervised canonical correlation analysis and extensions for multi-view dimensionality reduction," Journal of Visual Communication and Image Representation, 2014, 25, 1894-1904.

[11] D. D. Lee, H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," Nature, 1999, 401, 788-791.

[12] G. Cui, X. Li, Y. Dong, "Subspace clustering guided convex nonnegative matrix factorization," Neurocomputing, 2018, 292, 38-48.

[13] N. Guan, T. Liu, Y. Zhang, et al, "Truncated Cauchy non-negative matrix factorization for robust subspace learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 1-1.

[14] A. P. Singh, G. J. Gordon, "Relational learning via collective matrix factorization," In Proceedings of the 14th ACM SIGKDD international conference on Knowledge Discovery and Data Mining, 2008, pp. 650-658.

[15] J. Liu, C. Wang, J. Gao, J. Han, "Multi-view clustering via joint nonnegative matrix factorization," In Proceedings of the 2013 SIAM International Conference on Data Mining, 2013, pp. 252–260.

[16] J. Wang, X. Wang, F. Tian, C. H. Liu, H. Yu, Y. Liu, "Adaptive multi-view semi-supervised nonnegative matrix factorization," International Conference on Neural Information Processing, 2016, pp. 435–444.

[17] X. Zhang, H. Gao, G. Li, J. Zhao, J. Huo, J. Yin, "Multi-view clustering based on graph-regularized nonnegative matrix factorization for object recognition," Information Sciences, 2018, 432, 463-478.

[18] B. Qian, X. B. Shen, Y. Y. Gu, Y. H. Ding, "Double constrained NMF for partial multi-view clustering," International Conference on Digital Image Computing: Techniques and Applications, 2016, pp. 1-7.

[19] P. Luo, J. Peng, Z. Guan, J. Fan, "Dual regularized multi-view non-negative matrix factorization for clustering," Neurocomputing, 2018, 294: 1-11.

[20] X. Pei, Y. Wu, "Neighbors isometric embedding nonnegative matrix factorization for image representation," Multidimensional Systems and Signal Processing, 2017, 28(2): 529-547.

[21] W. Xu, X. Liu, Y. Gong, "Document clustering based on non-negative matrix factorization," In Proc ACM Conference Research and Development in Information Retrieval, 2003, pp. 267-273.

[22] X. Shen, W. Liu, I. W. Tsang, Q. S. Sun and Y. S. Ong, "Multilabel Prediction via Cross-View Search," IEEE Transactions on Neural Networks and Learning Systems, 2018, 1-15.