

稀疏约束图正则非负矩阵分解

姜伟¹ 李宏¹ 余震国¹ 杨炳儒²

(辽宁师范大学数学学院 大连 116029)¹ (北京科技大学计算机与通信工程学院 北京 100083)²

摘要 非负矩阵分解(NMF)是在矩阵非负约束下的一种局部特征提取算法。为了提高识别率,提出了稀疏约束图正则非负矩阵分解方法。该方法不仅考虑数据的几何信息,而且对系数矩阵进行稀疏约束,并将它们整合于单个目标函数中。构造了一个有效的乘积更新算法,并且在理论上证明了该算法的收敛性。在 ORL 和 MIT-CBCL 人脸数据库上的实验表明了该算法的有效性。

关键词 非负矩阵,图正则化,稀疏编码

中图法分类号 TP181 **文献标识码** A

Graph Regularized Non-negative Matrix Factorization with Sparseness Constraints

JIANG Wei¹ LI Hong¹ YU Zhen-guo¹ YANG Bing-ru²

(School of Mathematics, Liaoning Normal University, Dalian 116029, China)¹

(School of Computer & Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China)²

Abstract Nonnegative matrix factorization(NMF) is based on part feature extraction algorithm which adds nonnegative constraint into matrix factorization. A method called graph regularized non-negative matrix factorization with sparseness constraints(GNMFSC) was proposed for enhancing the classification accuracy. It not only considers the geometric structure in the data representation, but also introduces sparseness constraint to coefficient matrix and integrates them into one single objective function. An efficient multiplicative updating procedure was produced along with its theoretic justification of the algorithmic convergence. Experiments on ORL and MIT-CBCL face recognition databases demonstrate the effectiveness of the proposed method.

Keywords Non-negative matrix, Graph Regularization, Sparse coding

1 引言

非负矩阵分解(NMF)^[1,2]是在矩阵中所有元素均为非负数约束条件之下的矩阵分解方法,目的是为非负矩阵 $V = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$ 寻找适当的非负基矩阵 $W \in \mathbb{R}^{m \times r}$ 和非负系数矩阵 $H \in \mathbb{R}^{r \times n}$, 使它们的乘积近似于原始非负矩阵 V , 可写为如下形式:

$$V \approx WH \text{ 即 } V_{ij} = (WH)_{ij} = \sum_{a=1}^r W_{ia} (H)_{aj}$$

其中, $r < \min(m, n)$, 因此, 矩阵 W 可以看作是对原始数据矩阵 V 进行逼近的一组基, 而 H 是样本集 V 在基 W 上的非负投影系数矩阵, 系数矩阵是原始数据矩阵的降维结果; 原始矩阵 V 的列向量 V_{ij} 为基矩阵 W 的列向量 W_{ij} 的线性组合, 组合系数为矩阵 H 对应的行向量。

NMF 自提出后受到了广泛的研究与应用, 为了提高 NMF 算法的有效性和识别率, 人们提出了它的许多改进算法。Hoyer^[3] 把 NMF 算法和稀疏编码结合在一起形成非负稀疏编码(NNSC)方法, 使分解后的系数具有较好的稀疏性, 这样用更少量的元素表达原有信息, 可节省存储空间提高运算效率。Li 等^[4] 在标准的 NMF 目标函数基础上对基施加空

间局部化限制提出了 LNMF 算法。Liu 等^[5] 修改了 Hoyer^[3] 中描述的方法, 用 KL 代替了欧式距离, 提出了 SNMF 算法。Hoyer^[6] 对 Liu^[5] 的工作做了进一步的发展, 提出了可精确控制稀疏性的算法(NMFSC), 它最具特点也是最具创新性的地方是以非线性投影实现对稀疏性精确控制。Wang 等^[7] 和 Zafeiriou 等^[8] 分别提出了将线性判别分析(LDA)的判别信息融合在 NMF 中来提高 NMF 算法的判别能力, 即判别非负矩阵分解(DNMF)。

许多现实世界中的数据是嵌入在高维欧氏空间中的非线性低维流形上的^[9]。然而, NMF 算法和许多改进的 NMF 算法在处理数据时, 没有考虑数据的内蕴几何结构, 它是假设数据的分布是全局线性的, 这就大大地限制了数据位于非线性流形 NMF 算法的使用。与传统的线性降维算法相比, 流形学习算法能揭示数据内在的几何结构, 寻找高维数据在低维空间中的紧致嵌入。因此, 蔡登教授等^[10,11] 针对流形数据提出了图正则非负矩阵分解算法, 该方法在矩阵分解过程中明确考虑了数据集携带的几何信息; 如果数据点在原空间是邻近点, 那么对应到新的基下也是邻近点。本文在 GNMF 的基础上加入稀疏约束, 提出了稀疏约束图正则非负矩阵分解

到稿日期:2012-03-14 返修日期:2012-06-03 本文受国家自然科学基金项目(60875029)资助。

姜伟(1969—),男,博士生,副教授,主要研究方向为数据挖掘、流形学习、模式识别, E-mail: swxxjw@yahoo.com.cn; 李宏(1987—),女,硕士生,主要研究方向为机器学习、模式识别; 杨炳儒(1943—),男,教授,博士生导师,主要研究方向为知识工程、数据挖掘和人工智能。

(Graph Regularized Non-negative Matrix Factorization with Sparseness Constraints, GNMFSFC), 该方法在考虑几何结构的同时对系数矩阵进行稀疏, 这样节省了存储空间, 提高了算法的有效性。

2 非负矩阵分解

给定非负矩阵 $V = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{m \times n}$, 寻找适当的非负基矩阵 $W \in \mathbb{R}^{m \times r}$ 和非负系数矩阵 $H \in \mathbb{R}^{r \times n}$, 其中, n 为数据样本的个数, V 的列向量 $x_i (i=1, \dots, n)$ 表示一个样本, 从数学规划的角度, NMF 要最小化如下目标函数:

$$J_{NMF} = \frac{1}{2} \|V - WH\|_F^2 \quad \text{s. t. } W \geq 0, H \geq 0 \quad (1)$$

式中, $\|\cdot\|_F$ 表示矩阵的 Frobenius 范数。

Lee 和 Seung^[12] 给出了如下的乘性迭代规则:

$$h_{ij}^{t+1} \leftarrow h_{ij}^t \frac{(W^T V)_{ij}}{(W^T W H)_{ij}} \quad (2)$$

$$w_{ij}^{t+1} \leftarrow w_{ij}^t \frac{(V H^T)_{ij}}{(W H H^T)_{ij}} \quad (3)$$

3 图正则非负矩阵分解(GNMF)

GNMF 非负分解在矩阵分解过程要求使降维后的数据保持原始数据的几何信息。在分解过程中, GNMF 做了如下的假设: 若点 x_i 和点 x_j 在原空间中是邻近点, 那么在新的基下所对应的坐标 v_i 和 v_j 也一定是接近的。

设原始数据点构成的图为 G , 其中 S_{ij} 是权矩阵, $N_p(x_i)$ 表示 x_i 的 p 个近邻, 则

$$S_{ij} = \begin{cases} 1, & x_i \in N_p(x_j) \text{ 或 } x_j \in N_p(x_i) \\ 0, & \text{其他} \end{cases}$$

定义 $L = D - S$, D 是对角矩阵, $D_{ii} = \sum_j S_{ij}$, L 是拉普拉斯矩阵。

GNMF 最终的目标函数为:

$$J_{GNMF} = \frac{1}{2} \|V - WH\|_F^2 + \frac{1}{2} \lambda \text{tr}(HLH^T) \quad (4)$$

s. t. $W \geq 0, H \geq 0$

式中, λ, β 为大于 0 的常数,

Cai 等^[11] 给出了如下的乘性迭代规则:

$$h_{ij}^{t+1} \leftarrow h_{ij}^t \frac{(W^T V)_{ij} + \lambda(HS)_{ij}}{(W^T W H)_{ij} + \lambda(HD)_{ij}} \quad (5)$$

$$w_{ij}^{t+1} \leftarrow w_{ij}^t \frac{(V H^T)_{ij}}{(W H H^T)_{ij}} \quad (6)$$

4 稀疏约束的图正则非负矩阵分解(GNMFSFC)

近年来, 信号处理领域得到广泛关注的一个问题是稀疏表示, 稀疏表示就是将一信号分解为一系列基信号的线性组合, 并希望组合系数尽可能多的为 0。Lasso (Least Absolute Shrinkage and Selection Operator) 用于描述一类有约束的优化问题, 且用自变量对因变量进行线性回归, 并限定回归系数为 ℓ_1 范数不超过某个阈值。由此可见, Lasso 算法与稀疏表示是一脉相承的。借助于机器学习技术, 稀疏表示应用于降噪及超分辨率等图像处理问题。2007 年 Cai 等通过对投影方向施加稀疏性约束提出了统一的稀疏子空间学习算法^[10]。如果对 GNMF 加上系数矩阵稀疏约束, 那么就可以得到更加有效和稀疏的系数矩阵, 从而节省了存储空间, 提高了分解质量。

综合 GNMF 和系数矩阵稀疏约束条件获得的新的最小

化目标函数为:

$$J_{GNMFSFC} = \frac{1}{2} \|V - WH\|_F^2 + \frac{1}{2} \lambda \text{tr}(HLH^T) + \beta \|H\|_1 \quad (7)$$

s. t. $W \geq 0, H \geq 0$

式中, λ, β 为大于 0 的常数。

如下, 利用最速下降法和迭代法, 可以推导出式(7)的乘性迭代规则; 首先新的目标函数可表示为:

$$\begin{aligned} J_{GNMFSFC} &= \frac{1}{2} \text{Tr}((V - WH)(V - WH)^T) + \frac{1}{2} \lambda \text{Tr} \\ &\quad (HLH^T) + \beta \|H\|_1 \\ &= \frac{1}{2} \text{Tr}(VV^T) - \text{Tr}(VH^T W^T) + \frac{1}{2} \lambda \text{Tr} \\ &\quad (WHH^T W^T) + \frac{1}{2} \text{Tr}(HLH^T) + \beta \|H\|_1 \end{aligned} \quad (8)$$

s. t. $W \geq 0, H \geq 0$

由 Lagrange 定理:

$$\begin{aligned} J_{GNMFSFC} &= \frac{1}{2} \text{Tr}(VV^T) - \text{Tr}(VH^T W^T) + \frac{1}{2} \lambda \text{Tr}(HLH^T) \\ &\quad + \frac{1}{2} \lambda \text{Tr}(WHH^T W^T) + \beta \|H\|_1 + \text{Tr}(\Psi W^T) \\ &\quad + \text{Tr}(\Phi H^T) \end{aligned}$$

则 $J_{GNMFSFC}$ 关于 W 和 H 的偏微分为:

$$\frac{\partial J_{GNMFSFC}}{\partial W} = -VH^T + WHH^T + \Psi \quad (9)$$

$$\frac{\partial J_{GNMFSFC}}{\partial H} = -W^T V + W^T W H + \lambda H L + \beta I + \Phi \quad (10)$$

式中, $I \in \mathbb{R}^{r \times n}$ 是全 1 矩阵。

由 KKT 条件 $\phi_{ij} h_{ij} = 0$ 和 $\psi_{ij} w_{ij} = 0$, 得到如下关于 h_{ij} 和 w_{ij} 的公式(即 W 和 H 的乘性迭代规则):

$$h_{ij}^{t+1} \leftarrow h_{ij}^t \frac{(W^T V)_{ij} + \lambda(HS)_{ij}}{(W^T W H)_{ij} + \lambda(HD)_{ij} + \beta} \quad (11)$$

$$w_{ij}^{t+1} \leftarrow w_{ij}^t \frac{(V H^T)_{ij}}{(W H H^T)_{ij}} \quad (12)$$

5 算法收敛性证明

这一部分研究算法的收敛性, 首先引入辅助函数的定义。

定义 1 设 $G(h, h')$ 为 $F(h)$ 的辅助函数, 满足条件: $G(h, h') \geq F(h)$, $G(h, h) = F(h)$

引理 1 如果 $G(h, h')$ 是 $F(h)$ 的辅助函数, 则 $F(h)$ 在下面迭代规则下是非增的:

$$h^{t+1} = \arg \min_h G(h, h') \quad (13)$$

证明: 由上可以得到的结论为, $h = h^{t+1}$ 时函数 $G(h, h')$ 取最小值, 根据定义 1, $G(h^{t+1}, h')$ 大于等于 $F(h^{t+1})$, 用不等式概括为:

$$F(h^{t+1}) \stackrel{\text{def}}{\leq} G(h^{t+1}, h') \stackrel{\min}{\leq} G(h^t, h') \stackrel{\text{def}}{=} F(h^t)$$

注意到 $F(h^{t+1}) = F(h^t)$ 只有当 h^t 为 $G(h, h')$ 的局部极小点时才成立。如果函数 F 的导数存在且在 h^t 的一个极小领域内连续, 则微分 $\nabla F(h^t) = 0$ 。通过式(13)可得到一系列收敛到局部极小点 $h_{\min} = \arg \min_h F(h)$ 的序列:

$$F(h_{\min}) \leq \dots \leq F(h^{t+1}) \leq F(h^t) \leq \dots \leq F(h^1) \leq F(h^0)$$

可以看到通过定义这样的辅助函数 $G(h, h')$, 对于目标函数 $J_{GNMFSFC}$, 其相应的迭代规则能满足 $h^{t+1} = \arg \min_h G(h, h')$ 。

对于 H 中的任何元素 h_{ab} , 我们用 F_{ab} 表示 $J_{GNMFSFC}$ 中仅与 h_{ab} 有关的部分, 很容易得到:

$$F'_{ab} = (\frac{\partial J_{NMFC}}{\partial H})_{ab} = (-W^T V + W^T W H + \lambda H L + \beta I)_{ab} \quad (14)$$

$$F'_{ab} = (W^T W)_{aa} + \lambda L_{ab} \quad (15)$$

引理 2 函数

$$G(h, h'_{ab}) = F_{ab}(h'_{ab}) + F'_{ab}(h'_{ab})(h - h'_{ab}) + \frac{(W^T W H)_{ab} + \lambda(HD)_{ab} + \beta}{2h'_{ab}}(h - h'_{ab})^2 \quad (16)$$

是 F_{ab} 的辅助函数。

证明:显然 $G(h, h) = F_{ab}(h)$, 只需要证明 $G(h, h') \geq F_{ab}(h)$, 为了证明这个问题, 比较 $F_{ab}(h)$ 的泰勒展开:

$$F_{ab}(h_{ab}) = F_{ab}(h'_{ab}) + F'_{ab}(h'_{ab})(h - h'_{ab}) + [(W^T W)_{aa} + \lambda L_{ab}](h - h'_{ab})^2 \quad (17)$$

与式(17)比较后得到 $G(h, h'_{ab}) \geq F_{ab}(h)$ 等价于:

$$\frac{(W^T W H)_{ab} + \lambda(HD)_{ab} + \beta}{2h'_{ab}} \geq (W^T W)_{aa} + \lambda L_{ab} \quad (18)$$

我们有:

$$(W^T W H)_{ab} + \beta \geq (W^T W H)_{ab} = \sum_{l=1}^k (W^T W)_{al} h'_{lb} \geq (W^T W)_{aa} h'_{ab} \quad (19)$$

并且

$$\begin{aligned} \lambda(HD)_{ab} &= \lambda \sum_{j=1}^m h'_{aj} D_{jb} \geq \lambda h'_{ab} D_{bb} \geq \lambda h'_{ab} (D - S)_{bb} \\ &= \lambda h'_{ab} L_{ab} \end{aligned} \quad (20)$$

引理 3 通过最小化辅助函数 $G(h, h'_{ab})$ 可以得到式(11), h'_{ab} 是第 t 次迭代的解。

证明:为了得到最优解, 辅助函数 $G(h, h'_{ab})$ 对 h 进行微分有:

$$\frac{\partial G(h, h'_{ab})}{\partial h} = F'_{ab}(h'_{ab}) + \frac{(W^T W H)_{ab} + \lambda(HD)_{ab} + \beta}{h'_{ab}}(h - h'_{ab}) = 0 \quad (21)$$

因此能够得到 h_{ij} 的更新规则为:

$$\begin{aligned} &(-W^T V + W^T W H + \lambda H L + \beta I)_{ab} + \frac{(W^T W H)_{ab} + \lambda(HD)_{ab} + \beta}{h'_{ab}}(h - h'_{ab}) = 0 \\ &-W^T V + W^T W H + \lambda H L + \beta I - W^T W H - \lambda H D - \beta I + \frac{W^T W H + \lambda H D + \beta}{h'_{ab}} h = 0 \\ &-W^T V - \lambda H S + \frac{W^T W H + \lambda H D + \beta}{h'_{ab}} h = 0 \\ &\frac{W^T W H + \lambda H D + \beta}{h'_{ab}} h = W^T V + \lambda H S \\ &h = h'_{ab} \frac{W^T V + \lambda H S}{W^T W H + \lambda H D + \beta} \end{aligned}$$

所以更新规则为:

$$h_{ij}^{t+1} \leftarrow h'_{ij} \frac{(W^T V)_{ij} + (\lambda H S)_{ij}}{(W^T W H)_{ij} + (\lambda H D)_{ij} + \beta}$$

同理可得 w_{ij} 的更新规则为:

$$w_{ij}^{t+1} \leftarrow w'_{ij} \frac{(V H^T)_{ij}}{(W H H^T)_{ij}}$$

6 实验与结果分析

实验采用了两个比较著名的人脸数据库, 分别是 ORL 人脸数据库和 MIT-CBCL 人脸数据库。在实验中, 所有人脸图像都经过了标准化处理, 即首先标出人脸图像中两眼的位置并在该位置将图像对齐, 然后剪切出图像中的面部区域并将其缩放成统一的 32×32 大小, 图像最终按列堆叠成 1024 维的向量。

6.1 数据集

ORL 人脸数据库中有 40 个人的面部图像, 每人有 10 种不同的姿态, 每幅尺寸为 92×112 的像素, 即 92×112 的矩阵; 人脸姿态有相当程度的变化, 深度旋转和平面旋转可达 20 度, 人脸的尺度也有多达 10% 的变化。本文在 ORL 人脸数据库上选取 200 图作为训练集(即每人 5 幅图), 其它的作为测试集。

MIT-CBCL 人脸数据库由 10 个人的 2000 幅人脸图像构成, 人脸姿态有很大的变化, 其平面旋转最多可达 90 度, 选取 MIT-CBCL 人脸图像数据库中的 10 个人, 随机选取每个人的 10 幅图像作为训练样本, 另外选 10 幅作为测试样本, 这样就共有 100 幅训练人脸图像样本, 100 幅测试人脸图像样本。

6.2 参数(特征维度) r 的选择

在 GNMFCSC 算法中, 基矩阵的维数 r 的取值很重要, 它影响着算法的收敛速度, 虽然从理论上讲 $r < \min(m, n)$, 但是 r 的取值过大或过小都会影响算法的收敛性和分解后的重构误差。图 1 和图 2 分别表示在数据库 ORL 和数据库 CBCL 上 r 取不同值时 GNMFCSC 算法对应的误差曲线; 由图可知, 在 ORL 数据库上取 $r=25$ 时误差最小, 而在 CBCL 数据库上 $r=49$ 时误差最小(其中参数 $\lambda=1, \beta=1$)。

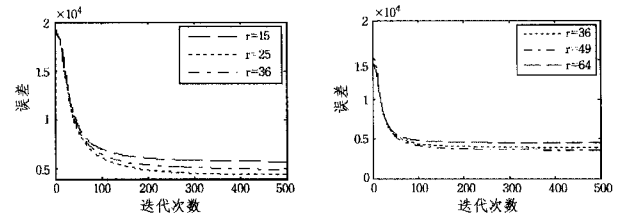


图 1 r 取不同值时的误差曲线 图 2 r 取不同值时的误差曲线

6.3 识别率分析

为了进一步验证 GNMFCSC 算法的有效性, 进行了人脸识别实验, 比较了该算法与几种经典的子空间学习算法; 这些对比的算法包括主成分分析(PCA)、标准非负矩阵分解(NMF)、稀疏非负矩阵分解(SNMF)、图正则非负矩阵分解(GNMF)。

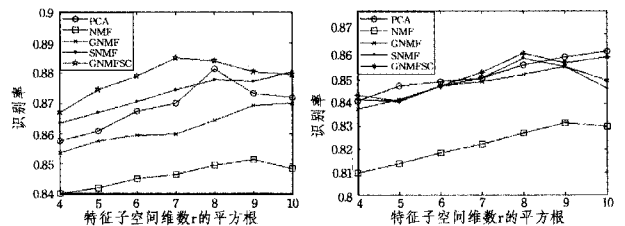


图 3 在 ORL 数据库上 r 取不同 图 4 在 CBCL 数据库上 r 取不同值时 5 种算法的识别率

图 3 和图 4 给出了在选取不同特征维数时这 5 种算法在 ORL 人脸数据库和 CBCL 数据库上的识别率; 其中将实验重复进行 10 次, 并取其识别率的平均值作为最终的实验结果。由图 3 可以看出, 在 ORL 数据库上相比其他的 4 种算法, GNMFCSC 得到了更好的识别效果, 并在特征子空间维数 $r=49$ 时, 达到了最高识别率 88.50%; 由图 4 可知, 在 CBCL 数据库上, GNMFCSC 在特征子空间维数 $r=64$ 时, 达到了最高识别率 86.15%, 同样 GNMFCSC 相比其他的 4 种算法得到了相对较高的识别率。

- [22] Li Yi, Liu Qiang. Chinese Sentence Similarity Based on Multi-feature Combination[C]//2009 WRI Global Congress on Intelligent Systems. 2009,3:751-756
- [23] Wang Rong-bo, Wang Xiao-hua, Chi zhe-ru, et al. Chinese Sentence Similarity Measure Based on Words and Structure Information[C]//2008 International Conference on Advanced Language Processing and Web Information Technology. 2008;27-31
- [24] Masrah H C F, Murad M A A, Doraisamy S C, et al. Measuring Sentence Similarity from Both the Perspectives of Commonalities and Differences[C]//22nd IEEE International Conference on Tools with Artificial Intelligence. 2010,1:318-322
- [25] Li Lin, Zhou Yi-ming, Yuan Bo-qiu, et al. Sentence Similarity Measurement Based on Shallow Parsing[C]//Sixth International Conference on Fuzzy Systems and Knowledge Discovery.

2009,7:487-491

- [26] Shan Jian-fang, Liu Zong-tian, Zhou Wen. Sentence Similarity Measure Based on Events and Content Words[M]. Sixth International Conference on Fuzzy Systems and Knowledge Discovery, 2009,7:623-627
- [27] Lee Ming-che, Zhang Jia-wei, Lee Wen-xiang, et al. Sentence Similarity Computation Based on POS and Semantic Nets[C]//2009 Fifth International Joint Conference on INC, IMS and IDC. 2009:907-912
- [28] Dolan B, Quirk C, Brockett C. Unsupervised construction of large paraphrase corpora; Exploiting massively parallel news sources[C]//Proceedings of the 20th International Conference on Computational Linguistics. 2004
- [29] <http://www.WordNet.org/>

(上接第 220 页)

6.4 基图像的稀疏度

首先定义度量稀疏度的函数^[6]为:

$$sp(x) = \frac{\sqrt{n} - (\sum |x_i|) / \sqrt{\sum x_i^2}}{\sqrt{n} - 1} \quad (22)$$

式中, n 是向量 x 的维度, $0 \leq sp(x) \leq 1$, 当且仅当 x 仅有一个非零元时, 函数值为 1; 当且仅当所有的元素相等时, 函数值为 0。

图 5 和图 6 分别表示在数据库 ORL 上特征维数 r 取值 25、在数据库 CBCL 上特征维数 r 取值 49 时, 用 NMF 算法、GNMF 算法、GNMFSC 算法对由训练图像构成的非负矩阵 V 进行矩阵分解得到的基图像, 其中用式 (22) 度量稀疏度, 把基矩阵 W 看作向量, 计算稀疏度。

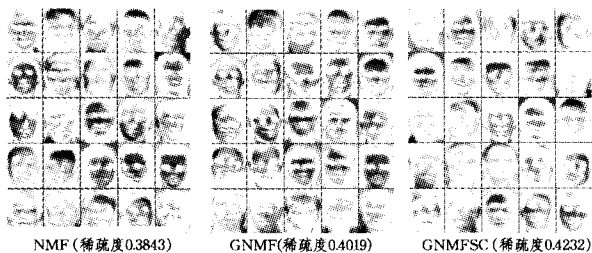


图 5 3 种算法在特征维数 r 为 25 时提取得到的基图像

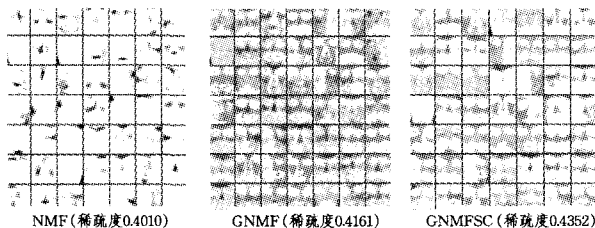


图 6 3 种算法在特征维数 r 为 49 时提取得到的基图像

由图 5 和图 6 可知, 在这两个数据库上对比这 3 种算法的基图像稀疏度, NMF 的稀疏度最差, GNMFSC 的稀疏度最高, 超过了 GNMF 基图像的稀疏度。也就是说, 与 NMF 算法、GNMF 算法相比, GNMFSC 算法得到了最佳的局部表示。

结束语 提出了稀疏约束的图正则非负矩阵分解, 并给出了迭代公式以及收敛性证明。以 ORL 人脸库和 CBCL 人

参考文献

- [1] 姜伟, 杨炳儒. 局部敏感非负矩阵分解[J]. 计算机科学, 2010, 37(12): 211-214
- [2] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature, 1999, 401(6755): 788-791
- [3] Hoyer P O. Non-negative sparse coding[C]//Processings IEEE Workshop on Neural Netw. Signal Process. 2002;557-565
- [4] Li S Z, Hou Xin-wen, Zhang Hong-jiang, et al. Learning spatially localized, parts-based representation[C]//Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2001, 1: 207-212
- [5] Liu Wei-xiang, Zheng Nan-ning, Lu Xiao-feng. Nonnegative matrix factorization for visual coding[C]//Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. 2003, 3: 293-296
- [6] Hoyer P O. Non-negative matrix factorization with sparseness constraints [J]. Journal of Machine Learning Research, 2004, 5(9): 1457-1469
- [7] Wang Yuan, Jia Yun-de, Hu Chang-bo, et al. Fisher non-negative matrix factorization for learning local features[C]//The Fifth Conference on Computer Vision. 2004
- [8] Zafeiriou S, Tefas A, Buciu I, et al. Exploiting discriminant information to frontal face verification [J]. IEEE Transactions on Neural Networks, 2006, 17(3): 683-695
- [9] Belkin M, Niyogi P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural Computations, 2003, 15(6): 1373-1396
- [10] Cai Deng, He Xiao-fei, Jia Wei-han. Spectral regression: A unified approach for sparse subspace learning[C]//IEEE International Conference on Data Mining (ICDM). 2007
- [11] Cai Deng, He Xiao-fei, Wu Xiao-yun, et al. Non-negative Matrix Factorization on Manifold[C]//The IEEE International Conference on Data Mining. 2008
- [12] Lee D D, Seung H S. Algorithms for non-negative matrix factorization[C]//Annual Conference on Neural Information Processing Systems. 2001, 13: 556-562