

非负矩阵分解特征提取技术的研究进展

王科俊, 左春婷

(哈尔滨工程大学 自动化学院, 哈尔滨 150001)

摘要: 非负矩阵分解(non-negative matrix factorization, NMF)算法是在矩阵中所有元素均为非负的条件下对其实现的非负分解,基于非负矩阵分解的图像特征提取技术通过将图像表示为一系列非负基图像非减的叠加组合来提取图像的特征,这种特征提取方法不但具有良好的局部表征特性、有一定的稀疏性,而且对遮挡、光照不均及图像质量较差等情形具有卓越的效果。自正式提出以来,该方法得到了许多改进,但目前关于这些改进的综述都只是罗列了这些方法,并没有系统深入地分析,因而在大量阅读文献的基础上分析其内部联系,分类总结了非负矩阵分解的研究进展和各种改进方法的实质。首先介绍非负矩阵分解的基本思想,以手指静脉图像为例说明其应用于图像特征提取的方式,然后重点深入讨论了非负矩阵分解方法的改进算法,提出了非负矩阵分解应用中有待进一步研究的新问题。

关键词: 模式识别; 非负矩阵分解; 图像特征提取; 稀疏表示

中图分类号: TP391.4

文献标志码: A

文章编号: 1001-3695(2014)04-0970-06

doi:10.3969/j.issn.1001-3695.2014.04.003

Improvements of non-negative matrix factorization for image extraction

WANG Ke-jun, ZUO Chun-ting

(College of Automation, Harbin Engineering University, Harbin 150001, China)

Abstract: Non-negative matrix factorization(NMF) is a new factor extraction method for pattern recognition. It aims to find two non-negative matrices whose products can well approximate the original matrix, which naturally leads to parts-based representation. Image feature extraction technique based on non-negative matrix factorization expresses an image as the non-negative combination of a series of base images. This feature extraction method has good local characteristics and results in some sparsity. Furthermore, it stands out in face of random occlusion, uneven illumination and poor image quality. Since the formal inception, the method has improved much, but most reviews just simply listed these improved methods without analyzing systematically and deep. Thus, this paper summarized a large number of researches and analysed their internal relationship, classified them and decribed the substance of these improved methods. It firstly described the basic idea of non-negative matrix factorization, and then took finger vein images as an example to extracting features. In next part, it put even more weight to discuss the available improved approaches and drew some conclusion on the future improvements.

Key words: pattern recognition; non-negative matrix factorization (NMF); feature extraction; sparse representation

心理学^[1]和生理学^[2,3]研究表明,对整体的感知是由对组成整体的部分的感知组成的(纯加性),这也符合直观的理解,即整体是由部分组成的^[4]。以这些事实为依据产生了许多基于部分的图像表示方法,统称为子空间的表示方法。一些经典的子空间方法如主成分分析(principal component analysis, PCA)、独立分量分析(independent component analysis, ICA)及其改进方法在模式识别领域已经得到了广泛的应用^[5-7],这些方法得到的基图像是一种全局的描述;同时它们的像素点可以是正值也可以是负值,描述原图像的线性组合中可能会存在相减关系,缺少直观意义上的由局部合成整体的效果,而且对于图像矩阵,负值的存在使分解结果缺少可解释和明确的物理意义,失去了与实际问题的联系。针对这些问题, Lee 等人^[4]于1999年提出了一种新的特征子空间的方法——非负矩阵分解(NMF), NMF通过对基和系数的非负约束,将重构图像表示为一系列非负基图像非减的叠加组合,与传统子空间方法相比,这一重建过程更接近于由局部组合而成为整体的过程,具有明

确的物理意义。而且, NMF的分解结果还具有一定的稀疏性,能在一定程度上抑制由外界变化(如部分遮挡、光照变化和物体的旋转等)给特征提取带来的不利影响^[8],在图像提取特征任务中的效果较好。

本文首先介绍非负矩阵分解算法的基本内容,以手指静脉图像为例说明非负矩阵分解应用于图像特征提取的方法和特点,然后系统地总结了 NMF 改进算法,最后就未来有待进一步研究的方向作出了展望。

1 NMF 和基于 NMF 的特征提取技术

非负矩阵分解的本质是一种矩阵分解和投影技术,其基本原理如下:

对于非负矩阵 V , 存在 $W \geq 0, H \geq 0$, 满足:

$$V_{n \times m} \approx W_{n \times r} H_{r \times m} \quad (1)$$

其中: r 为特征维数, 通常 $nm(n+m)r$, 所以得到的 W 和 H 小

收稿日期: 2013-07-30; 修回日期: 2013-09-10

作者简介: 王科俊(1962-), 男, 黑龙江哈尔滨人, 教授, 博导, 主要研究方向为生物特征识别与智能监控技术、计算生物信息学、健身物联网技术等(wangkej@hrbeu.edu.cn); 左春婷(1990-), 女, 硕士研究生, 主要研究方向为指纹及手指静脉识别。

于原始矩阵 V , 从而实现数据的压缩, 压缩率为 $\frac{nm}{nr+mr}$ 。称 W 为基矩阵, H 为系数矩阵。

1.1 NMF 的目标函数和优化

为了描述 $V \approx WH$ 的近似效果, 需要定义目标函数, 可以通过分解前后两个非负矩阵即 V 与 WH 之间的一些距离来度量。一种方法是利用矩阵 V 与 WH 间的欧氏距离, 对应的目标函数为

$$E(V \| WH) = \|V - WH\|^2 = \sum_{ij} [V_{ij} - (WH)_{ij}]^2 \quad (2)$$

另一种方法是利用矩阵 V 与 WH 间的广义 K-L 散度, 对应的目标函数为

$$D(V \| WH) = \sum_{ij} (V_{ij} \log \frac{V_{ij}}{(WH)_{ij}} - V_{ij} + (WH)_{ij}) \quad (3)$$

式(2)(3)对应的优化问题为

a) $\min E(V \| WH)$, 并且 $W \geq 0, H \geq 0$;

b) $\min D(V \| WH)$, 并且 $W \geq 0, H \geq 0$ 。

对于 WH 上述两式是非凸的, 为此文献[5]中给出了交替迭代直至达到最优解的算法。进一步地, 为保证非负矩阵分解结果的非负性, 采用了如下的乘性迭代规则:

对目标函数 a), 相应的迭代规则为

$$W_{ik} \leftarrow W_{ik} \frac{(VH^T)_{ik}}{(WHH^T)_{ik}} \quad (4)$$

$$H_{kj} \leftarrow H_{kj} \frac{(W^T V)_{kj}}{(W^T WH)_{kj}} \quad (5)$$

对目标函数 b), 相应的迭代规则为

$$W_{ik} \leftarrow W_{ik} \frac{\sum_j H_{kj} V_{ij} / (WH)_{ij}}{\sum_j H_{kj}} \quad (6)$$

$$W_{ik} \leftarrow \frac{W_{ik}}{\sum_j W_{jk}} \quad (7)$$

$$H_{kj} \leftarrow H_{kj} \frac{\sum_i W_{ik} V_{ij} / (WH)_{ij}}{\sum_i W_{ik}} \quad (8)$$

按照上述迭代规则, 当 W 和 H 趋于固定不变时, 目标函数收敛, Seung 等人[5]给出了两种迭代算法收敛的完整证明。图1是目标函数与迭代次数的关系曲线。

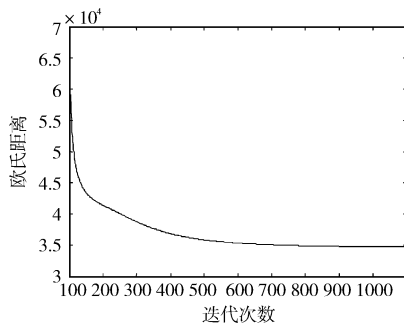


图1 非负矩阵分解收敛曲线

从图1中可以看出, 迭代一定次数后, 目标函数趋于稳定, 还可以从图1中读出收敛需要的最少迭代次数及目标函数稳定后的阈值, 因此这两个参数可以作为设定 NMF 算法结束的依据。

1.2 基于 NMF 的特征提取

使用 NMF 特征提取方法, 得到的基矩阵代表的是图像中起关键作用的特征, 系数矩阵存放的是将这些局部特征累加为整体的权系数。下面以 NMF 提取手指静脉图像特征为例说明 NMF 用于特征提取的方式方法, 采用的图片为本实验室手指静脉图像库的感兴趣区域样本库。

a) 将手指静脉图像向量化。使手指静脉图片的每列像素按扫描顺序依次排列, 从而将其由 $p \times q$ 的二维矩阵 $A_i (i = 1, 2, \dots, m)$ 转换成 $n \times 1$ 一维列向量 v_i , 其中 $(n = p \times q)$, 那么 m 个样本就组成了 $n \times m$ 矩阵 $V = [v_1 \ v_2 \ \dots \ v_m]$, 矩阵的每一列代表一个样本图片。

b) 将式(1)写成列的形式 $v_i = Wh_i$, 即 V 的一列 v_i 和 H 相对应的一列 h_i 对应, 换言之, v_i 为基矩阵 W 的所有行向量的线性组合, 权系数为 h_i 。所以, 在 W 子空间下, 对应较小维数的 h_i 就可以代替原始样本 v_i , 求取待识别手指静脉图像 NMF 的系数矩阵, 就得到了其图像特征。投影系数 h_i 的好坏由基矩阵 W 确定, W 中相对较少的基向量应满足对原始数据的描述与逼近。

从图2(b)(c)两种方法提取手指静脉特征的结果对比可以看出, NMF 提取的特征与 PCA 提取的特征有较大区别, NMF 提取的局部特征较为稀疏, 并且重构图像中静脉纹路比较深, 而 PCA 的基矩阵表现比较稠密, 接近静脉的整体特性。

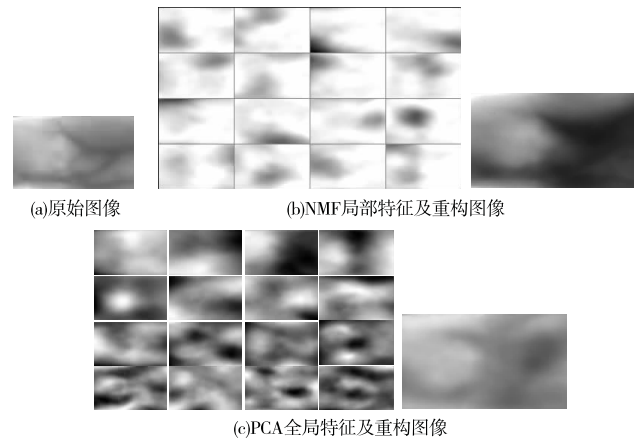


图2 手指静脉特征基和重构图像

2 NMF 的改进

NMF 算法自提出以来, 经过众多学者的不懈努力取得了长足发展, 各种改进方法相继提出。研究发现, 大多数 NMF 算法都需要将图像矩阵向量化, 但向量化操作不仅会产生大量的高维数据, 导致算法耗时较长, 而且有可能丢失一些结构信息, 影响最终的识别率。针对这个问题, 又产生了一些直接基于图像矩阵的分解算法。本章将 NMF 的改进算法分为基于向量的算法和基于二维图像的算法两大类分别展开讨论。

2.1 基于向量的改进算法

传统 NMF 算法以欧氏距离或 K-L 散度作为目标函数, 采用乘性迭代算法对 W 、 H 进行交替更新, 因此改进操作也分别从这两方面进行。

2.1.1 迭代算法的改进

传统的 NMF 采用乘性迭代交替更新, 其实质是梯度法, 该算法的收敛速度慢, 用于特征提取时需要较长的训练时间, 很难适应大规模数据处理, 因此找到快速的迭代更新方法是 NMF 推广应用需解决的一大难题。

对 NMF 而言, 无论是基于欧氏距离还是 K-L 散度的目标函数, 对 WH 都是非凸的, 但对于 W 或 H 单独而言却是凸的。所以基于这两类目标函数的改进 NMF 算法一般也沿用交替迭代更新的方式, 每一步都求出当前的最优解, 从而得到问题的稳定点或最优解。

NMF 的目标函数和约束条件如式(2)(3)所示,要得到这样的约束优化问题的最优解,梯度法可能是最简单的实现方式,但收敛速度很慢,而且梯度法对步长的选取十分敏感^[9];共轭梯度法收敛要比梯度法快,但是其实现方式要复杂得多^[10]。

对基于欧氏距离的 NMF 算法, Lin^[11] 针对梯度法的缺陷提出了投影梯度法,其实质是在每步梯度法更新后进行简单的非负投影,即将其中非负元素修正为 0,改进后的算法比乘性迭代能更快收敛;文献[12]也表明基于此迭代方法的稀疏非负矩阵分解在人脸识别中的分解时间及识别率均好于传统 NMF 和稀疏非负矩阵分解;Kim 等人^[13]将 NMF 转换为序列的非负最小二乘问题,其思想是:如果固定 W ,则对于单变量的目标函数可以转换为求解一序列的非负最小二乘问题,反之亦然。基于此提出了基于最小非负二乘问题的积极集法(active method)迭代收敛在实验中表现出较为理想的效果;Li 等人^[14]则将 NMF 转换为秩 1 序列部分和的形式,提出固定 H 和 $\{W_{\cdot 1}, W_{\cdot 2}, \dots, W_{\cdot r}\} / W_{\cdot i}$ 依次更新 $W_{\cdot i} (i=1, 2, \dots, r)$ 再固定 W 和 $\{H_1, H_2, \dots, H_r\} / H_j$ 依次更新 H_j 的快速收敛算法,文献[15]实验证明该算法有很好的收敛速度。

对基于 K-L 散度的 NMF 算法,目标函数如式(3)所示,如果对模型固定变量 W ,则目标函数可以表示为

$$D(V \| WH) = \sum_j D(V_{\cdot j} \| WH_{\cdot j}) \quad (9)$$

因此求解式(5)就转换为求解式(9)所示的多个子问题。为表示简单,此处用 $D(v \| Wh)$ 表示子问题的目标函数,子问题即为

$$\begin{aligned} \min D(v \| Wh) \\ \text{s. t. } h \geq 0 \end{aligned} \quad (10)$$

由乘性迭代算法知其迭代式为

$$h_a^{(k+1)} = h_a^{(k)} \frac{\sum_i W_{ia} v_i / (Wh^{(k)})_i}{\sum_i W_{ia}} \quad (11)$$

这一求解方法实际上就是 1982 年 Shepp 等人^[16]在电子扫描极大似然图像重建中所提出的最大期望算法(expectation-maximization algorithm, EM),即基于 K-L 散度的约束优化问题等价于 EM 算法,那么对于 EM 的各种改进算法^[17-21],可以考虑适当修正后运用到 K-L 散度的 NMF 算法改进中。例如文献[22]就基于 EM 算法的块迭代加速技术提出了 NMF 的块迭代算法,不但保留了乘性迭代的性质,单调性也得到了保障,实验中表明大多数时候这种算法都优于传统的 NMF 算法。

2.1.2 目标函数的改进

综合诸多文献发现,传统 NMF 中存在的一些问题如非正交的基矩阵导致数据冗余性较大、无监督训练过程不利于后续识别率提高等,这些缺陷可通过添加一些约束得到改善,约束的改变往往表现为目标函数的形式变化。另外,除了 Lee 等人^[22,23]提出的基于欧氏距离和 K-L 散度的目标函数外,基于 dual-KL 散度、 α 散度和 β 散度等的目标函数也得到了一些研究,由于这些目标函数的算法有时是可以通用的^[23],因篇幅有限此处仅讨论由约束引起的目标函数的改进。

1) 稀疏方面的约束改进

子空间方法进行特征提取时,往往希望基矩阵 W 尽量稀疏,得到足够细节的局部特征,系数矩阵尽量稀疏,待测图片与尽可能少的基图像相关,从而更有利于识别。虽然文献[4]中表明 NMF 的分解结果具有一定的稀疏性,但其稀疏程度远远

不能达到要求,有什么办法可以进一步稀疏化 NMF 的分解结果呢?能否显性直观地控制基矩阵或系数矩阵的稀疏程度?一些文献中对这些问题作出了探索,重要成果可以总结如下。

(1) 加局部约束的 NMF

为了得到更为局部的特征, Li 等人^[24]提出了局部非负矩阵分解算法(local nonnegative matrix factorization, LNMF),通过对 K-L 散度目标函数施加基的列正交性约束以最小化基向量的个数,并减少基向量之间的冗余。LNMF 的损失函数为

$$D(X \| WH)_{\text{LNMF}} = D(X \| WH)_{\text{NMF}} + \alpha \sum_{ik} u_{ik} - \beta \sum_{kk} v_{kk} \quad (12)$$

其中: $[u_{jk}] = U = W^T W$, $[v_{jk}] = V = HH^T$ 和 $\alpha > 0, \beta > 0$ 为常量。LNMF 的任务就是在下面三个约束条件下,求取上式损失函数的最小值:a) $\min \sum_j u_{jj}$ (系数矩阵最大的稀疏性,以生成更局部的特征); b) $\min \sum_{j \neq k} u_{jk}$ (基矩阵的正交性,以最小化基图像的冗余); c) $\max \sum_j v_{jj}$ (基图像有最大的代表性),得到基矩阵和系数矩阵的迭代式:

$$h_{jk} \leftarrow \sqrt{\frac{h_{kj} \sum_i w_{ki} x_{ij}}{\sum_k w_{ik} h_{kj}}} \quad (13)$$

$$w_{ik} \leftarrow \frac{w_{ik} \sum_j x_{ij} h_{jk}}{\sum_j h_{kj}} \quad (14)$$

$$w_{ik} \leftarrow \frac{w_{ik}}{\sum_i w_{ik}}, \forall k \quad (15)$$

基于人脸重构和识别的数值实验^[25]表明, LNMF 分解得到的特征图像比 NMF 更稀疏,在识别率和训练速度方面也更优越。

(2) 稀疏增强的 NMF

尽管 NMF 在某种程度上产生了稀疏的非负矩阵^[4], LNMF 通过添加局部约束增加了基矩阵的稀疏程度,但是利用对矩阵分解直接引入稀疏约束,可以得到更稀疏的表示。Wright 等人^[26]在人脸识别中提出对表达系数进行稀疏约束,有很好的识别效果和一定的鲁棒性。使基矩阵更稀疏,可以得到更为细节的局部基图像;使系数矩阵更稀疏,则更少的基图像就可以线性表示原图像。

根据文献[27]中的稀疏编码原则, Hoyer^[28]结合 NMF 和稀疏编码提出了非负稀疏编码(non-negative sparse coding, NNSC),并给出了简单有效的乘性迭代规则。

稀疏非负矩阵分解(sparse non-negative matrix factorization, SNMF)^[29]就是这样一种非负稀疏矩阵分解算法,其思想是以传统 K-L 距离测量近似误差,对系数矩阵 H 进行稀疏化,其中基矩阵 W 的每列的一范数为 1,目标函数为

$$\min D_{\text{SNMF}}(V \| WH) = \sum_{ij} (V_{ij} \log \frac{V_{ij}}{(WH)_{ij}} - V_{ij} + (WH)_{ij}) + \alpha \sum_{kj} H_{kj} \quad (16)$$

$$\text{s. t. } W \geq 0, H \geq 0, \sum_i W_{ik} = 1, \forall k$$

相比于 NMF, SNMF 通过最小化所有编码系数的和对系数矩阵进行稀疏约束。式(16)中,正常数 α 为稀疏因子,当 $\alpha = 0$ 时, SNMF 等同于 NMF。上式对应的乘性迭代规则如下:

$$W_{ik} \leftarrow W_{ik} \sum_j H_{kj} V_{ij} / (WH)_{ij} \quad (17)$$

$$W_{ik} \leftarrow \frac{W_{ik}}{\sum_l W_{lk}} \quad (18)$$

$$H_{kj} \leftarrow \frac{H_{kj} \sum_i W_{ik} V_{ij} / (WH)_{ij}}{1 + \alpha} \quad (19)$$

SNMF 在保证分解结果非负的基础上,还考虑了稀疏化系数矩阵 H ,使得原始矩阵与尽可能少的基向量相关。该算法全部采用乘性迭代规则,能很好地保持数据的非负特性。

以 NMF 欧氏距离目标函数为基础,引入稀疏度约束,通过显式控制因子矩阵的稀疏度获取更局部化的分解,得到对应的目标函数为

$$\min E(W, H) = \|V - WH\|^2 \quad (20)$$

$$\text{spareness}(w_i) = S_w, \forall i; \text{spareness}(h_i) = S_h, \forall i \quad (21)$$

其中: S_w 表示基矩阵 W 的第 i 列 w_i 的稀疏度; S_h 表示系数矩阵 H 的第 i 行 h_i 的稀疏度,同样变化范围为 $[0, 1]$, 稀疏因子的公式为

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

这就是文献[30]中将稀疏因子引入到 NMF 框架中,提出的非负矩阵分解的稀疏约束(non-negative matrix factorization with sparseness constraints, NMFSC)算法。与 SNMF 性比,该算法可以对系数矩阵、基矩阵稀疏度同时直观地控制,其迭代过程如下:

a) 设定 W 和 H 的初始值,由程序自动生成随机的正矩阵。

b) 如果对基矩阵 W 进行稀疏约束,那么将 W 的每一列元素 w_i 变为非负的,并保证它在 L_1 范数上满足稀疏要求, L_2 范数不变。

c) 如果对系数矩阵 H 进行稀疏约束,那么将 H 的每一行元素 h_i 变为非负的,并保证它在 L_1 范数上满足稀疏要求, L_2 范数不变。

d) 根据以下迭代规则进行迭代,达到稀疏要求目标函数稳定时迭代停止。

(a) 如果对 W 的列有稀疏约束,则迭代规则为

$$W_{\text{new}} = W - \mu_w (WH - V) H^T \quad (23)$$

如果对 W 的列没有稀疏约束,则按照标准的乘法规则进行:

$$W_{\text{new}} = W * (V H^T) ./ (W H H^T) \quad (24)$$

(b) 如果对 H 的行有稀疏约束,则迭代规则为

$$H_{\text{new}} = H - \mu_h W^T (WH - V) \quad (25)$$

如果对 H 的行没有稀疏约束,则按照标准的乘法规则进行:

$$H_{\text{new}} = H * (W^T V) ./ (W^T W H) \quad (26)$$

注:这里的 $*$ 和 $./$ 都是基于矩阵元素的乘除法。式(23) (25)中, μ_w 和 μ_h 是迭代步长,通常取为非常小的正数值,保证算法收敛。

2) 类别方面的约束改进

传统的 NMF 算法是未考虑类别信息的非监督过程,若加入分类约束,使分解的基矩阵包含类别信息,系数矩阵也会更稀疏,不但重构图像的线性组合有明确的物理意义,而且能大大地降低训练时间。这里所说的分类约束,可以是类内距最小、类间距最大等条件,也可以是 PCA、ICA、LDA 等子空间方法的内容。

在 LNMF 的基础上继续修改目标函数,加上最大类间和最小类内约束,就得到了判别非负矩阵分解(discriminant non-negative matrix factorization, DNMF)算法,其目标函数表示为

$$D_{\text{DNMF}}(V \| WH) = D_{\text{LNMF}}(V \| WH) + \gamma \sum_{c=1}^C \sum_{l=1}^{n_c} (h_{cl} - \mu_c) (h_{cl} - \mu_c)^T - \delta \sum_{c=1}^C (\mu_c - \mu) (\mu_c - \mu)^T \quad (27)$$

其中: μ_c 表示第 c 类样本对应分解系数的均值; μ 表示所有样本对应分解系数的均值; n_c 为第 c 类训练样本数量; C 为总类

别数。

DNMF 的分解结果的稀疏程度介于 NMF 和 LNMF 之间,就面部表情识别而言, DNMF 优于 LNMF,这说明在 LNMF 致力于获得局部化表示的过程中可能存在一些重要特征的丢失^[31]。

Wang 等人^[32]提出为 NMF 加入 Fisher 判别约束用于人脸图像的 FNMF 特征提取方法,并证明其识别效果比 LNMF 更优越。而 Zafeiriou 等人^[33]提出了 NMF plus LDA 的有监督的两阶段的特征提取方法——NMFfaces 方法,在人脸识别实验中的效果好于著名的 Fisherfaces 和 Eigenfaces 方法。

2.1.3 突破 NMF 框架的算法

另外,一些学者尝试将 NMF 的思想作为约束条件运用到其他特征提取技术中并取得了一定的成效,如文献[34]提出非负稀疏主成分分析(non-negative sparse PCA, NSPCA)算法,该算法将非负稀疏分解与 PCA 算法结合,从根本上打破了非负矩阵分解框架,而且此算法训练时仅需要计算基矩阵,避免了 NMF 迭代更新需要同时计算基矩阵和系数矩阵,使得算法十分耗时的问题(系数矩阵维数一般很大,不利于使用迭代法求解)。

2.2 基于二维图像的改进算法

上述基于一维向量的 NMF 算法均存在如下两方面的问题:a) 图像规模较大时,图像向量化常常导致高维的图像矢量空间,使得 NMF 需要耗费大的计算代价,而且尽可能接近地重构原始图像很困难;b) 这种将 2D 矩阵生硬写成 1D 向量的变换,很可能会丢失原始 2D 图像中隐藏着的结构信息。针对这些问题,可以直接使用图像矩阵进行非负矩阵分解,这就是二维非负矩阵分解(two-dimensional non-negative matrix factorization, 2DNMF)。

文献[35]先在列向量然后在行向量上应用 NMF,以得到相应的一维的列向量基和行向量基,然后计算这些列向量、行向量的向量积作为 2DNMF 的 2D 基,求取待测图像在 2D 基矩阵下的投影系数作为图像特征。

将 m 个 $p \times q$ 维样本图片 A_k 排列成 $p \times qm$ 矩阵 $X = [A_1 A_2 \dots A_m]$, 其中 A_k 代表一个原始样本;然后采用如下的步骤训练得到 2DNMF 的 2D 基矩阵,从而求得待测图片的系数矩阵作为提取到的特征。

1) 列 NMF 采用 NMF 算法找到 $p \times q$ 维非负矩阵 L 和 $d \times qm$ 维非负矩阵 H 使得 $X \approx LH$, 其中 $H = [H_1 H_2 \dots H_m]$, 则有 $A_k = LH_k$, H 矩阵的一列对应原始某图像的一列像素,即第 k 个图片是基矩阵 L 各列的加权和,故称 L 为列基矩阵,此次非负矩阵分解称为列 NMF。

2) 行 NMF 将 H 改写成 $q \times dm$ 维矩阵 $H' = [H_1^T H_2^T \dots H_m^T]$, 采用 NMF 算法找到 $q \times g$ 维矩阵 R 和 $g \times dm$ 维矩阵 C 使得 $H' \approx RC$, 其中 $C = [C_1 C_2 \dots C_m]$, 则有 $H_k^T = RC_k$ 。因为 H' 矩阵的一列对应原始某图像的一行,所以 C 的一列也对应此原始图像的一行,故称 R 为行基矩阵,此次非负矩阵分解称为行 NMF。

故 $X = LC^T R^T$, 记 $L = [l_1 l_2 \dots l_d]$, $R = [r_1 r_2 \dots r_g]$, $C^T = D = [D_{ij}^k]$, $E_{ij} = l_i \cdot r_j^T$, 可以得到:

$$A_k = \sum_{i=1}^d \sum_{j=1}^g d_{ij}^k l_i r_j^T = \sum_{i=1}^d \sum_{j=1}^g D_{ij}^k E_{ij} \quad (28)$$

这样就得到了 2D 基矩阵 $(E_{ij})_{p \times q}$, 而 $D^k = \frac{L^T A_k R}{L^T L R^T R}$ 为系数

矩阵,为了简便可以直接用 $F = L^T A_k R$ 作为待测图片的特征。

由于 2DNMF 可以视为基于行分解的 NMF,仅保留了图像行向量之间的相关信息,而不能覆盖所有的结构信息,高宏娟等人^[36]将原始的人脸图像进行对角化,提出基于图像对角变换的 diagonal-2DNMF 算法解决了这一问题。因为经过转换的对角图像中的行(列)同时整合了原始图像中行和列的信息,通过行和列信息的关联,diagonal-2DNMF 可以使得对角线信息得以充分的利用,实验证明对角化后其识别率得到了提高。但此方法与基矩阵正交化^[37]相比,还是后者方法能得到更高的识别率。所以文献[38]中又继续改进提出了 $(2D)^2$ NMF 方法:a)两次非负矩阵分解并行进行,大大缩短了计算时间;b)将基矩阵进行正交化,消除特征间的相关性,以更适合欧氏距离的相似性度量,从而大大地提高了识别精度;c)图像对角化、镜像变换后再进行分解、正交化,将得到的基矩阵加权,相当于增加了训练样本的数量,进一步增加了图像的结构信息。关于人脸识别的实验证明, $(2D)^2$ NMF 方法在识别率方面大大优于传统 NMF 方法,训练速度大大优于 2DNMF 方法,可以完全取代 2DNMF。

在基于向量的 NMF 算法的基础上,还可以进行适当的二维改进,得到新的二维改进 NMF 算法。文献[38]中提到对 2DNMF 分解后的基矩阵进行正交化就是文献[37]中基于向量的基矩阵正交化 NMF 的二维改进;文献[39]对文献[34]的 NSPCA 方法进行二维改进,提出了二维投影非负矩阵分解(two-dimensional projective non-negative matrix factorization, 2DPNMF)的人脸识别算法,进一步降低了计算复杂度。该算法的思想是寻找一个 $p \times r$ 维的非负矩阵 W ,使得图像 A_k 在 W 上的投影最大,同时要求 W 不同列之间尽量单位正交,即

$$\max f(W) = \sum_{i=1}^m |W^T A_i| - \alpha |W^T W - I|, \text{ s. t. } W \geq 0 \quad (29)$$

基于二维图像的算法需要的训练时间较短,尤其本方法不需要计算投影系数,所以训练时间很短。但需要指出的是,基于二维图像算法的投影系数较多,此算法的测试时间要长于基于向量的 NSPCA 算法。

3 有待进一步研究的问题

非负矩阵分解得到了许多改进,但至今为止还没有一个很好的评价体系来统一衡量各种算法,此外,如何最小化目标函数?如何进行非负矩阵分解的初始化?这些问题都还没有一个十分卓越的方法,在其推广应用的道路上还存在着一些开放性难题,值得进一步探索。

1) 子空间维数的选取

如何确定低维特征数 r 的大小以获得最佳的分解效果?目前这个问题还没有很好的解决方法,这是因为 r 的最佳取值与图像数据有关。但是在分解时,只有 $r \ll \min\{m, n\}$ 降维效果才能显著,但 r 太小又会导致特征提取的信息不足,难以满足后续分类的要求。现在通用的一般方法是先选取一系列的 r 值分别进行分解实验,取得最佳识别率的子空间维数确定为最终结果。

2) 基矩阵和系数矩阵的初始化

非负矩阵分解得到的是局部最优解,其结果往往与初始化有关。关于迭代规则的初始值,一些学者尝试了各种方法,如 Wild 等人^[40]球形聚类的初始化方法、Boutsidis 等人基于奇异

值分解(singular value decomposition, SVD)的初始化方法^[41]、Buciu 等人^[42]则在原始数据库中选取构造基向量。

3) 最优化问题

非负矩阵分解作为一个约束优化问题,需要寻找合理的目标函数和相关约束,然后应用最合适的最优化手段进行优化。目前应用最广的是文献[5]中的两种方法,其效率均不高,用于中等规模问题时,一般至少需要几百乃至上千步的迭代才可接近收敛。目前有一些改进算法对此有一定的提高,但探索更为卓越的优化手段、甚至找到全局最优化解决办法将能更好地推广 NMF 的使用,也将是今后非负矩阵分解一个重要的发展方向。

4) 非线性特征提取的 NMF 方法

标准的 NMF 是对图像数据的线性分解,近来有一些 NMF 改进方法提取非线性特征的研究,如文献[43]提出核 NMF 方法提取图像的非线性特征。另外,寻求将非负性约束和其他方法相结合的新的分解方法,如非负独立成分分析^[44]和非负 Boltzmann 机^[45],也是值得探索的。

4 结束语

非负矩阵分解是一种专门针对非负数据进行线性非负分解的矩阵分解方法。与传统子空间方法提取图像特征的过程一致,利用非负矩阵分解的方法进行特征提取,需要先训练求解出图像的基矩阵即子空间向量集,然后求取待测图片在子空间上的投影系数作为提取的图像特征。非负矩阵分解算法,在功能上,它实现了对大脑的基于部分感知功能的模拟;在算法上,采用简单有效的乘性迭代规则,很好地保持了非负性;在应用中,非负性的数据大量存在,且非负分解的结果具有明确的物理含义。作为一种低秩逼近算法,它能有效节约存储和计算资源。

除了图像特征提取,非负矩阵分解算法在网络安全入侵检测、灰度图像的数字水印技术、基因及细胞分析和脑电信号处理方面也有广泛的应用,由此可见非负矩阵分解在智能信息处理和模式识别的大规模多维数据处理与分析中有着十分重要的应用意义。从已有的对非负矩阵分解的改进算法以及上面的讨论来看,需要在理论上对非负矩阵分解作更深入的研究,寻求新的分解算法,并进一步探讨该方法在智能信息处理和模式识别领域中的应用。

参考文献:

- [1] PALMER S E. Hierarchical structure in perceptual representation [J]. *Cognitive Psychol*, 1977, 9(4): 441-474.
- [2] WACHSMUTH E, ORAM M W, PERRETT D I. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque [J]. *Cerebral Cortex*, 1994, 4(5): 509-522.
- [3] LOGOTHETIS N K, SHEINBERG D L. Visual object recognition [J]. *Annual Review of Neuroscience*, 1996, 19(1): 577-621.
- [4] LEE D, SENG H. Learning the parts of objects by non-negative matrix factorization [J]. *Nature*, 1999, 401(6755): 788-791.
- [5] LEE D D, SEUNG H S. Algorithms for non-negative matrix factorization [C]// *Advances in Neural Information Processing Systems*. Berlin: Springer, 2001: 556-562.
- [6] TURK M A, PENTLAND A P. Face recognition using eigenfaces [C]// *Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington DC: IEEE Computer Socie-

- ty, 1991: 586-591.
- [7] BARLETT M S. Independent component representations for face recognition[M]//Face Image Analysis by Unsupervised Learning. New York: Springer, 2001: 39-67.
- [8] SPRATLING M W. Learning image components for object recognition[J]. *Journal of Machine Learning Research*, 2006, 7(2): 793-815.
- [9] GUAN Nai-yang, TAO Da-cheng, LUO Zhi-gang, *et al.* NeNMF: an optimal gradient method for nonnegative matrix factorization[J]. *IEEE Trans on Signal Processing*, 2012, 60(6): 2882-2898.
- [10] WANG Wen-wu, ZOU Xuan. Non-negative matrix factorization based on projected nonlinear conjugate gradient algorithm[C]//Proc of ICA Research Network International Workshop. 2008: 5-8.
- [11] LIN C J. Projected gradient methods for non-negative matrix factorization[J]. *Neural Computation*, 2007, 19(10): 2756-2779.
- [12] 平沙沙, 褚蕾蕾. 基于投影梯度法的非负矩阵分解稀疏算法[J]. *计算机与数字工程*, 2012, 40(12): 20-22.
- [13] KIM H, PARK H. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method[J]. *SIAM Journal on Matrix Analysis and Applications*, 2008, 30(2): 713-730.
- [14] LI Le, ZHANG Yu-jin. FastNMF: highly efficient monotonic fixed-point nonnegative matrix factorization algorithm with good applicability[J]. *Journal of Electronic Imaging*, 2009, 18(3): 4-33.
- [15] 刘勾连. 实用快速非负矩阵分解算法的研究及应用[D]. 大连: 大连理工大学, 2010.
- [16] SHEPP L A, VARDI Y. Maximum likelihood reconstruction for emission tomography[J]. *IEEE Trans on Medical Imaging*, 1982, 1(2): 113-122.
- [17] HUDSON H M, LARKIN R S. Accelerated image reconstruction using ordered subsets of projection data[J]. *IEEE Trans on Medical Imaging*, 1994, 13(4): 601-609.
- [18] LANGE K, HUNTER D R, YANG I. Optimization transfer using surrogate objective function[J]. *Journal of Computational and Graphical Statistics*, 2000, 9(1): 1-20.
- [19] NEAL R M, HINTON G E. A view of the EM algorithm that justifies incremental, sparse, and other variants[M]//Learning in Graphical Models. Cambridge: MIT Press, 1998: 355-368.
- [20] FESSLER J A, ERDOGAN H. A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction[C]//Proc of IEEE Conference Record of Nuclear Science Symposium. 1998: 1132-1135.
- [21] BROWNE J, De PIERRO A B. A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography[J]. *IEEE Trans on Medical Imaging*, 1996, 15(5): 687-699.
- [22] SRA S. Block-iterative algorithms for non-negative matrix approximation[C]//Proc of the 8th IEEE International Conference on Data Mining. 2008: 1037-1042.
- [23] 蒋娇娇. 非负矩阵分解算法的改进及应用[D]. 北京: 北京工业大学, 2011.
- [24] LI S Z, HOU Xin-wen, ZHANG Hong-jiang, *et al.* Learning spatially localized, parts-based representation[C]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2001: 207-212.
- [25] 宋星光, 夏利民, 赵桂敏. 基于LNMF分解的人脸[J]. *计算机工程与应用*, 2005, 41(5): 42-43.
- [26] WRIGHT J, YANG A Y, GANESH A, *et al.* Robust face recognition via sparse representation[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2009, 31(2): 210-227.
- [27] OLSHAUSEN B A, FIELD D J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images[J]. *Nature*, 1996, 381(6583): 607-609.
- [28] HOYER P O. Non-negative sparse coding[C]//Proc of the 12th IEEE Workshop on Neural Networks for Signal Processing. 2002: 557-565.
- [29] LIU Wei-xiang, ZHENG Nan-ning, LU Xiao-feng. Non-negative matrix factorization for visual coding[C]//Proc of IEEE International Conference on Acoustics, Speech, and Signal Processing. 2003: 293-296.
- [30] HOYER P O. Non-negative matrix factorization with sparseness constraints[J]. *Journal of Machine Learning Research*, 2004, 5(12): 1457-1469.
- [31] BOCIU I, PITAS I. A new sparse image representation algorithm applied to facial expression recognition[C]//Proc of the 14th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing. 2004: 539-548.
- [32] WANG Yuan, JIA Yun-de, HU Chang-bo, *et al.* Fisher non-negative matrix factorization for learning local features[C]//Proc of Asian Conference on Computer Vision. 2004: 27-30.
- [33] ZAFEIROUS S, TEFAS A, PITAS I. Discriminant NMF-faces for frontal face verification[C]//Proc of IEEE International Workshop on Machine Learning for Signal Processing. 2005: 355-359.
- [34] ZASS R, SHASHUA A. Nonnegative sparse PCA[M]//Neural Information Processing Systems. Cambridge: MIT Press, 2007: 1561-1568.
- [35] ZHANG Dao-qiang, CHEN Song-can, ZHOU Zhi-hua. Two-dimensional non-negative matrix factorization for face representation and recognition[C]//Proc of the 2nd International Workshop on Analysis and Modeling of Faces and Gestures. Berlin: Springer, 2005: 350-363.
- [36] 高宏娟, 潘晨. 基于非负矩阵分解的人脸识别算法的改进[J]. *计算机技术与发展*, 2007, 17(11): 63-66.
- [37] 刘维湘. 非负矩阵分解及其应用[D]. 西安: 西安交通大学, 2005.
- [38] 高宏娟, 潘晨. 基于 $(2D)^2$ NMF及其改进算法的人脸识别[J]. *计算机应用*, 2007, 27(7): 63-66.
- [39] 方蔚涛, 马鹏, 成正斌, 等. 二维投影非负矩阵分解算法及其在人脸识别中的应用[J]. *自动化学报*, 2012, 38(9): 1503-1512.
- [40] WILD S M, CURRY J, DOUGHERTY A. Improving non-negative matrix factorizations through structured initialization[J]. *Pattern Recognition*, 2004, 37(11): 2217-2232.
- [41] BOUTSIDIS C, GALLOPOULOS E. SVD-based initialization: a head start for nonnegative matrix factorization[J]. *Pattern Recognition*, 2008, 41(4): 1350-1362.
- [42] BUCIU I, NIKOLAIDIS N, PITAS I. On the initialization of the DNMF algorithm[C]//Proc of IEEE International Symposium on Circuits and Systems. [S.l.]: IEEE Press, 2006: 4.
- [43] BUCIU I, NIKOLAIDIS N, PITAS I. Nonnegative matrix factorization in polynomial feature space[J]. *IEEE Trans on Neural Networks*, 2008, 19(6): 1090-1100.
- [44] PLUMBLEY M D. Conditions for non-negative independent component analysis[J]. *IEEE Signal Processing Letters*, 2002, 9(6): 177-180.
- [45] DOWNS O B, MACKAY D J C, LEE D D. The nonnegative Boltzmann machine[C]//Advances in Neural Information Processing Systems. 2000: 428-434.
- [46] HEILER M, SCHORR C. Learning sparse representations by non-negative matrix factorization and sequential cone programming[J]. *Journal of Machine Learning Research*, 2006, 7(12): 1385-1407.