

可行方向算法与模拟退火结合的 NMF 特征提取方法

陈卫刚, 戚飞虎

(上海交通大学计算机科学与工程系, 上海 200030)

摘 要: NMF 子空间特征提取被表示成一个大规模线性约束非线性优化问题. 为了获得更优性能的基图像, 设计了一个可行方向算法结合模拟退火算法的混合算法来求解这个优化问题. 以基于梯度的可行方向算法作为局部寻优的手段, 加快收敛速度; 以模拟退火算法作为全局寻优的手段, 避免优化过程陷入局部极小点. 同时, 在模拟退火操作中, 采用对比度增强算法, 使获得的基图像更加地空间局部化. 实验表明, 本文的可行方向算法比采用归一化实现等式约束的原算法在学习的最后阶段有更好的收敛速度, 所获得的基图像更加地空间局部化, 而且在人脸重建的应用中有更好的性能.

关键词: 子空间特征提取; NMF; 可行下降方向算法; 模拟退火; 人脸重建

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2003) 12A2190-04

Learning NMF Representation Using a Hybrid Method Combining Feasible Direction Algorithm and Simulated Annealing

CHEN Weigang, QI Feihu

(Department of Computer Science & Engineering, Shanghai Jiao Tong University, Shanghai 200030, China)

Abstract: Nonnegative matrix factorization (NMF) is formulated as a large-scale optimization problem with linear equality constraints. To get better performance, a hybrid combining simulated annealing (SA) and gradient-based algorithm is designed. The simulated annealing gradually produces better solutions with the gradient-based algorithm serving as an accelerator. Experimental results are presented to compare our method and the original method for learning NMF representation, which demonstrate the proposed method can learn basis images more spatially localized and perform better than the original one at later training stages. The comparison study on face reconstruction also shows that the proposed method leads to better results than PCA and the original NMF.

Key words: subspace representation; nonnegative matrix factorization; feasible direction method; simulated annealing; face reconstruction

1 引言

子空间分析技术在基于外观的人脸建模和识别等领域一直占有非常重要的地位. 通过对选定的人脸图像进行学习使人脸模型得到定义, 同时, 实现人脸图像高维空间的低维特征提取. 主成分分析 (PCA)^[1,2] 和独立成分分析 (ICA)^[3] 是二类最常用的方法. 前者的基互相两两正交, 而后的基互相统计独立. 由于 PCA 和 ICA 的基图像的像素点可以是正值也可以是负值, 所以这二种表示方法都缺少直观意义上的由部分合成整体的效果^[5]. 近来, Lee 等人提出了一种新的特征提取方法: Nonnegative matrix factorization (NMF)^[6]. NMF 对基图像的像素点和用于重建的系数都施加了非负性约束, 使得重建图像是由基图像非减的叠加组合而成, 使重建过程更接近于由部分组合而成为整体的过程. 与 PCA 和 ICA 的基图像相比, NMF 的基图像能更好地反映人脸的局部信息, 因此 NMF 也被称作

基于部分 (parts-based) 的特征提取方法.

文献[5~8]的学习算法本质上是基于梯度的算法, 只能求得局部最优解. 而且, 由于在每一个迭代步中采用归一化的方法来保证线性等式约束, 当接近某个局部最优解时将会发生明显的锯齿 (zigzagging) 现象. 对于计算 NMF 基图像这样的大规模约束优化问题, 算法的效率和收敛速度是必须考虑的因素. 本文提出了一种新的 NMF 学习算法, 设计了一个混合算法来求解对应于 NMF 特征提取的大规模非线性优化问题. 以模拟退火算法作为全局寻优的手段, 结合基于梯度的可行方向算法 (Feasible direction method) 来增强局部搜索能力. 由于没有采用归一化或罚函数的方法来保证线性等式约束, 本文的可行下降方向算法能有效地减少优化最后阶段的锯齿现象. 同时, 本文在模拟退火过程中对基图像进行对比度增强操作, 使基图像逐步地实现空间局部化.

2 NMF

设 m 幅训练人脸图像构成一个 $n \times m$ 维矩阵 $V = [v_1, v_2, \dots, v_m]$, 其中 $v_j \in \mathbb{R}^n$, 是对第 j 幅训练图像按列堆叠而成, 且满足下式:

$$v_{ij} \geq 0 \\ \sum_{i=1}^n v_{ij} = 1, \quad j = 1, 2, \dots, m$$

NMF 在下式所示的矩阵分解的框架下, 将人脸图像表示成基图像的线性叠加^[6]:

$$V = WH$$

其中 $W = [w_1, w_2, \dots, w_r]$, 第 i 个基图像 w_i 是一个 n 维向量. H 是一个 $r \times m$ 维矩阵, 它的每一列称作一个编码 (encoding), 且与训练集中的图像一一对应, 其中的每一个元素代表了合成人脸图像时每个基图像的权重. 文献[7]给出了二类不同的目标函数. 相应地, NMF 被定义为形式不同, 但目标相同的两个优化问题:

问题 1

$$\min_{W, H} D(V, WH) = \sum_{i=1}^n \sum_{l=1}^m \left[v_{il} \log \left(\frac{v_{il}}{(WH)_{il}} \right) - v_{il} + (WH)_{il} \right] \quad (1)$$

问题 2

$$\min_{W, H} \|V - WH\|^2 = \sum_{i=1}^n \sum_{l=1}^m [v_{il} - (WH)_{il}]^2 \quad (2)$$

其中 W 和 H 满足以下关系式:

$$\sum_{i=1}^n w_{ij} = 1, \quad j = 1, 2, \dots, r \\ W, H \geq 0 \quad (3)$$

通过乘法形式的修正公式^[7], 可以得到上述优化问题的局部最优解, 从而获得 NMF 基图像. 以问题 1 为例, 其迭代公式如下^[6]:

$$w_{kl} = w_{kl} \frac{v_{kl}}{(WH)_{kl}} h_{li} \quad (4)$$

$$w_{kl} = \frac{w_{kl}}{\sum_j w_{jl}}$$

$$h_{li} = h_{li} \frac{v_{il}}{(WH)_{il}} w_{il} \quad (5)$$

注意上述迭代公式本质上是一种基于梯度的算法 (见附录 A). 通过归一化运算来保证式(3)的线性等式约束, 从而将原优化问题转化为仅包含非负性约束的优化问题. 这样虽然使问题得到了简化, 在优化过程的初始阶段也有很好的收敛性能, 但归一化操作会在优化的末尾阶段引起明显的锯齿现象, 使收敛很慢.

3 本文的 NMF 学习算法

本文用训练图像和对应的重建图像的像素差值的平方和作为目标函数, NMF 被定义成如问题 2 所示的大规模线性约束非线性优化问题. 显然, 当重建图像与训练图像一致时式(2)取最小值. 为了避免在优化过程的最后阶段由于归一化操

作而引起锯齿现象, 本文提出了一种基于可行下降方向的优化算法.

3.1 可行方向算法

可行方向算法 (如投影梯度法和简约梯度法) 被证明是求解如式(2)所示的线性约束优化问题的有效方法^[9]. 本文采用迭代的方法, 在当前可行点处计算可行下降方向 (Improving feasible direction), 结合线性搜索算法, 确定下一个可行解.

注意到式(3)中第 j 个等式约束仅依赖于第 j 个基向量, 将目标函数和约束条件表示成仅考虑一个基图像的形式, 单独地计算对应基向量 w_j 的搜索方向:

$$\min f(w_j) = \|V - WH\|^2 \\ \text{s.t. } g(w_j) = \sum_{i=1}^n w_{ij} = 1 \\ w_{ij} \geq 0 \quad (6)$$

不难看出, 当以下条件满足时, d_j 是对应于 w_j 的可行下降方向: $1^\circ \nabla f(w_j)^T d_j < 0$; $2^\circ \sum_i d_{ij} = 0$; $3^\circ w_j + \lambda d_j \geq 0$. 其中 $\nabla f(w_j) = [5f/5w_{1j}, \dots, 5f/5w_{nj}]^T$, λ 是搜索步长. 条件 1° 保证 d_j 与当前的负梯度方向一致; 条件 2° 和 3° 使 d_j 满足约束 $g(\#)$ 和非负性.

设 I_{pw} 和 I_{nw} 分别为向量 $\nabla f(w_j)$ 中非负元素和负元素的下标集合. 为了满足上述条件, d_j 按以下方法计算:

$$d_{ij} = -\frac{A}{5f/5w_{ij}}, \quad \text{if } i \in I_{pw} \quad (7)$$

$$d_{ij} = -\frac{B}{5f/5w_{ij}}, \quad \text{if } i \in I_{nw} \quad (8)$$

$$\text{其中 } A = \min \left[\frac{w_{ij}}{5f/5w_{ij}}, 1 \right], B = \frac{\sum_{i \in I_{pw}} d_{ij}}{\sum_{i \in I_{nw}} |5f/5w_{ij}|}$$

式(7)和式(8)的基本思想为: 对于 $\nabla f(w_j)$ 中的正元素, 需要与一个合适的系数相乘后, 作为 d_j 的元素, 以避免修正后的基向量违反非负性约束; d_j 中的正元素之和与负元素之和的绝对值应相等, 以满足约束 $g(\#)$.

按相似的方法, 可以确定对应于 H 的可行下降方向.

3.2 模拟退火算法

基于梯度的优化算法虽然有很快的收敛速度, 但需要好的初始值, 且容易陷入局部极小. SA 是一种有效的随机全局优化算法, 它通过模拟物理退火过程以避免陷入局部极值. 由于没有利用梯度信息, SA 搜索通常是个冗长的过程. 本文的基本思想是将模拟退火算法与基于梯度的算法相结合, 取二者之长, 补二者之短. 另外, 由于希望基图像反映人脸的局部特征, 所以本文以一定的概率进行如下操作:

$\hat{0}$ 以二维形式将第 i 个基图像表示成 B_i , 且进行平滑操作;

$\hat{1}$ 按以下形式的点运算^[10]对基图像进行对比度增强操作:

$$b_i(x, y) = \begin{cases} b_i(x, y), & b_i(x, y) > \text{Thr} \\ Sb_i(x, y), & \text{otherwise} \end{cases}$$

其中 Thr 是一个大于灰度均值的阈值, SI (0, 0.6) 是一个随

机数。
0 归一化各个基图像,使它们满足 $g(\#)$ 的等式约束,且按列堆叠,构成矩阵 W 。

需要指出的是,虽然上述过程有针对性地提高了较亮像素与较暗像素之间的对比度,但并没有改变 SA 随机搜索的本质。

根据以上讨论,我们将算法概括总结如下:
(1)置 E_1 和 E_2 为接近 0 的正实数,置 $k=1$,置较大的初始温度 T (本文取 25)。初始化 W 和 H ,使它们满足式(3)中的约束条件;

(2) $j=1$ Br,循环执行 1 ~ »
1 按式(7)和式(8)计算 d_j ;
° 通过线性搜索解如下最小化问题来确定步长 K_j

$$\min f(W\tilde{W}, H)$$
其中 $\tilde{W} = [w_1, \dots, w_j, \dots, w_r]$,而 $\tilde{w}_j = w_j + K_j d_j$, $K_j \in [0, 1]$
» $W = \tilde{W}$ 。

(3)按相似的方法计算 $d^{(H)}$,确定步长 $K^{(H)}$,计算 H ;
(4)若 $\|W^{(k+1)} - W^{(k)}\| < E_1$ 且 $\|H^{(k+1)} - H^{(k)}\| < E_2$,输出 W 和 H ,结束;否则转(5);

(5)若 $k / (\text{iters}) \geq \text{iters} = k$ (其中 $\text{iters} \in \mathbb{Z}^+$,用于控制二次 SA 操作间的迭代次数,本文取 15),按概率 $p = e^{-1/T}$ 执行上文所述的平滑和对比度增强算法,置 $T = GT$ 。本文取 $G=0.85$;

(6) $k = k + 1$,从步骤(2)重复循环。

4 实验结果

训练样本来自两个图像库:120 幅来自 Feret 人脸库^[11],40 幅来自 AR 人脸库^[12]。包括不同的姿态和表情。为了避免背景中较亮的像素对基图像产生影响,本文以手工的方式将背景区域设置成黑色。图 1 是其中的 8 个样本。



图 1 训练样本实例,图像大小:48@72

4.1 基图像

利用上述训练样本集,本文分别计算了 $r=56, 64, 81$ 的基图像。

图 2 是 $r=56$ 的实例(图中较亮的像素对应基图像中较小的灰度值)。由于采用了如 3.1.2 节所述的对比度增强算法,本文算法得到的基图像比 Lee 等人的方法得到的基图像在空间局部化方面有所增强。

图 3 为本文的可行下降方向算法和 Lee 等人的算法(即式(4)和式(5))的迭代过程的比较,其中 F 为目标函数值。可以看出, Lee 等人的算法在训练的初始阶段具有良好的性能,而本文算法在训练的最后阶段收敛得更好。

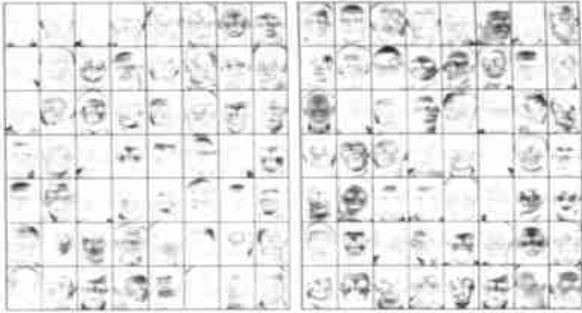


图 2 $r=56$ 的 NMF 基图像实例
左图:本文结果,右图:原方法结果

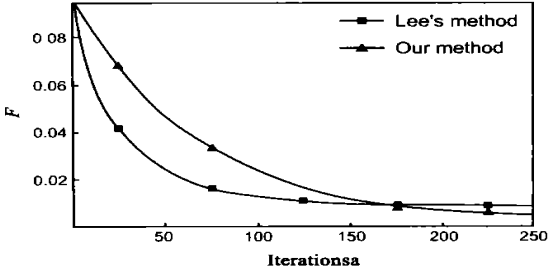


图 3 本文可行下降方向算法与 Lee 等人算法迭代过程的比较
4.1.2 人脸重建

为了评价基图像是否有效地提取了人脸的低维特征,本文对 NMF 基图像和 PCA 基图像在人脸重建方面作了实验比较。测试图像同样来自两个图像库,且不包含于训练集中。其中 180 幅来自 Feret 人脸数据库,60 幅来自 AR 人脸数据库。以手工方式作粗略的分割,背景设置成黑色。

按如下公式来衡量重建图像与原图像的吻合程度:

$$p_i = \left(1 - \frac{\|x_i - \hat{x}_i\|^2}{\|x_i\|^2 + \|\hat{x}_i\|^2} \right) @100\%$$

其中, x_i 是测试图像, \hat{x}_i 为对应的重建图像。显然,当重建图像与原图像一致时, p 取极大值 1。PCA 基图像和重建图像的计算方法见文献[13], NMF 重建图像计算方法见文献[14]。结果见表 1,可以看出本文算法得到的基图像具有最好的重建结果。

表 1 重建图像与测试图像吻合程度比较(p)

算 法	$r=56$	$r=81$
PCA	87.22%	88.10%
NMF(Lee's method)	88.50%	89.39%
NMF(Our method)	90.21%	91.57%

5 结论

本文提出了一种基于混合算法(可行下降方向算法与模拟退火算法相结合)的 NMF 子空间特征提取方法。该方法有

二个特点值得强调:(1)可行下降方向算法没有采用归一化或罚函数的方法来保证等式约束。所以,在训练的最后阶段,本文算法优于 Lee 等人的算法;(2)以基于梯度的算法进行局部搜索,以 SA 作为全局搜索的手段,本文算法在保证收敛速度的同时,更有可能获得全局最优解。实验表明,由于在随机搜索的同时,通过点运算增强了基图像的对比度,本文算法得到的基图像比 Lee 等人的算法得到的基图像更加地空间局部化。

下一步的工作方向包括:(1)研究能更好地提取人脸局部信息的 NMF 模型;(2)研究具有更快收敛速度的学习算法。

附录 A 修正公式(4)和(5)的推导和说明

设 $A = WH$, 不难看出 W 中的元素 w_{ki} 仅影响 A 中的元素 a_{ki} , $i = 1, 2, \dots, m$ 。所以,定义关于 w_{ki} 的目标函数:

$$f_w = \sum_{i=1}^m \left[v_{ki} \log \frac{v_{ki}}{a_{ki}} - v_{ki} + a_{ki} \right]$$

则 $\frac{5f_w}{5w_{ki}} = \sum_{i=1}^m \left[-\frac{v_{ki}}{a_{ki}} h_{ki} + h_{ki} \right]$, 而负梯度方向对 f_w 而言是一个下降方向。修正公式可以定义成:

$$w_{ki} \leftarrow w_{ki} - G \frac{5f_w}{5w_{ki}}$$

为了保证非负性,取 $G = \frac{w_{ki}}{\sum_i h_{ki}}$, 可得 $w_{ki} \leftarrow \frac{w_{ki}}{\sum_i h_{ki}} \sum_{i=1}^m \frac{v_{ki}}{a_{ki}} h_{ki}$ 。对

W 的各个元素, $\sum_i h_{ki}$ 可以认为是常量,且为了保证式(3)的等式约束,需要进行后续的归一化操作。所以忽略分母中的 $\sum_i h_{ki}$, 可以得到式(4)的迭代公式。

H 中的元素 h_{li} 仅影响 A 中的元素 a_{li} , $i = 1, 2, \dots, n$ 。按相似的方法可以得到式(5)修正公式。

参考文献:

- [1] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. Fisher2 faces: recognition using class specific linear projection[J]. IEEE Trans PAMI 1997, 19(7): 711- 720.
- [2] Moghaddam B. Principal manifolds and probabilistic subspaces for visual recognition[J]. IEEE Trans PAMI, 2002, 24(6): 780- 788.
- [3] Hyvarinen A, Oja E. Independent component analysis: algorithms and applications[J]. Neural Networks, 2000, 13(4, 5): 411- 430.
- [4] Bartlett M S, Movellan J R, Sejnowski T J. Face recognition by independent component analysis[J]. IEEE Trans Neural Networks, 2002, 13(6): 1450- 1464.

- [5] Li S Z, Hou X W, Zhang H J, Cheng Q S. Learning spatially localized, part2 based representation[A]. Proceedings of the 2001 IEEE Computer Society Conference on CVPR[C]. IEEE, 2001. 207- 212.
- [6] Lee D D, Seung H S. Learning the parts of objects by non2negative matrix factorization[J]. Nature, 1999, 401(10): 788- 791.
- [7] Lee D D, Seung H S. Algorithms for non2negative matrix factorization [A]. Advances in Neural Information Processing[C]. MIT Press, 2001. 556- 562.
- [8] Guillemet D, Vitria J. Evaluation of distance metrics for recognition based on nonnegative matrix factorization[J]. Pattern Recognition Letters, 2003, 24(9, 10): 1599- 1605.
- [9] Bazzara M S, Sherali H D, Shetty C M. Nonlinear Programming[M]. John Wiley & Sons Inc. 1993.
- [10] Castlman K R. Digital Image Processing[M]. Prentice2Hall Inc. 1996.
- [11] Phillips P J, Wechsler H, Huang J, Rauss P. The FERET database and evaluation procedure for face recognition algorithms[J]. Image and Vision Computing, 1998, 16(5): 295- 297.
- [12] Martinez A M, Benavente R. The AR Face Database[R]. CVC Technical Report # 24, 6, 1998.
- [13] 边肇祺, 张学工, 等. 模式识别[M]. 北京: 清华大学出版社, 2000.
- [14] Feng T, Li S Z, Shum H Y, Zhang H J. Local non2negative matrix factorization as a visual representation[A]. Proceedings of the 2nd International Conference on Development and Learning[C]. IEEE, 2002. 178- 183.

作者简介:



陈卫刚 男, 1967 年生于浙江省, 现为上海交通大学博士研究生, 主要研究方向为数字视频处理、数字图像处理、目标跟踪和识别。



戚飞虎 男, 1938 年生于上海, 现为上海交通大学教授、博导, 主要研究方向为数字视频处理和神经计算, 已在国内外公开发表论文 100 余篇。