

改进投影梯度非负矩阵分解的单训练样本特征提取研究

高涛 何明一

(西北工业大学电子信息学院 西安 710072)

摘要: 人脸识别是当前人工智能和模式识别的研究热点。非负矩阵分解(NMF)能够反映样本的局部的内在的联系,可用于单样本特征提取,但时间复杂度较高。投影梯度(Projected Gradient, PG)优化方法大幅降低了NMF约束优化迭代问题的时间复杂度,但是单训练样本存在对本类信息量描述不足的缺点。为此,该文提出了一种基于改进的投影梯度非负矩阵分解(Improved Projected Gradient Non-negative Matrix Factorization, IPGNMF)的单训练样本特征提取方法。在进行PGNMF算子之前,先将训练样本作Gabor分解,分解后的Gabor子图像在各个方向上可以更加丰富的描述样本特征,最后将各个Gabor子图像的PGNMF特征进行融合,作为最终的识别特征。在对人脸库ORL, YEL与FERET的识别实验中,与经典的特征提取方法比较,证明了可以有效地解决单训练样本人脸识别的问题。

关键词: 人脸识别; 非负矩阵分解; 投影梯度非负矩阵分解; 径向基网络

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2010)05-1121-05

DOI: 10.3724/SP.J.1146.2009.00622

Using Improved Non-negative Matrix Factorization with Projected Gradient for Single-Trial Feature Extraction

Gao Tao He Ming-yi

(Electronic and Information School, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Face recognition is an active research area in the artificial intelligence. A face recognition algorithm using improved Non-negative Matrix Factorization(NMF) with Projected Gradient(PG) for single-trial feature extraction is proposed based on this problem. NMF is a matrix factorization method, which can reflect the inherent partial contact and effectively express single sample information. However, NMF iteration time complexity of the gradient projection optimization method significantly reduces the NMF iteration time complexity of the problem. But the single training sample information has inadequate description of disadvantage, for this disadvantage, before the NMF operator, training sample is filtered by multi-orientation Gabor filters with multi-scale to extract their corresponding local Gabor magnitude map, the PGNMF feature of which were constructed to higher dimensional feature vectors. Experimental results on the ORL face database, YALE face database and FERET face database show that the proposed method is feasible and has higher recognition performance compared with GREY, PCA, ICA, NMF, PGNMF and other algorithms where only one sample image per person is available for training.

Key words: Face recognition; Non-negative Matrix Factorization(NMF); Projected gradient NMF; RBF network

1 引言

近几年以来,提出了许多人脸识别方法。这些方法当有充分数量并且有代表性的训练样本时,能够取得较好的识别效果。但是在现实生活中,往往很难获取具有代表性特征的多幅训练样本,而仅仅只有单样本可供训练和学习,所以提出一种基于单训练样本的人脸识别方法是非常必要的。

针对单样本的人脸识别的问题, Zhu^[1], Gao^[2], Li^[3]提出几种单样本人脸识别的方法。本文的方法如图1所示,首先对样本图像块进行多方向、多分辨率

Gabor小波滤波,并提取其对应不同方向、不同尺度的多个Gabor幅值域图谱GMM(Gabor Magnitude Map),然后使用投影梯度非负矩阵分解的方法提取每个图谱的NMF特征,再将每个图谱的特征融合,最后采用RBF神经网络的方法对融合后的特征识别分类,在Yale, ORL和FERET人脸库上的实验表明,当训练集中为单样本图像时,本文方法比传统的分析等方法取得了更高的识别率。本文提出的方法有以下的特点:

(1)基于投影梯度非负矩阵分解有效地解决了非负矩阵分解中迭代时间复杂度高的问题,使得整个识别过程更接近于实时性的要求。

(2)由于单样本的限制,很难提取丰富的类内特

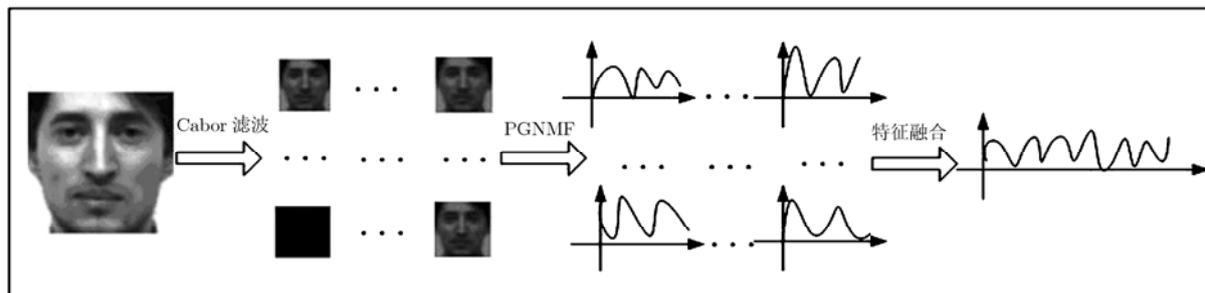


图1 GPGNMF 框图

征, 所以不直接使用PGNMF作用于样本, 而是利用多通道的Gabor 滤波器组获得丰富、规范的纹理测量, 最后将各个Gabor滤波后子图像的PGNMF特征进行融合, 作为最终的识别特征。

2 局部多通道 Gabor 滤波器特征分析

2 维 Gabor 函数是经高斯函数调制的正弦栅格, 在频域上表现为一种带通滤波器, 且主频和方向可调。Marcelja 首先使用 Gabor 滤波器来模拟视觉细胞的接收场, 指出了其对信号空间频率的局部分析能力。Daugman^[4]则进一步分析了信号在空间、空间频率和方向上的不确定联系, 得出 Gabor 滤波器是唯一达到时频测不准关系下界的函数。用多通道 Gabor 滤波进行纹理分析有其生理学等方面的基础。假设每一通道的数学模型为

$$\left. \begin{aligned} \varphi(x, y) &= \sqrt{\varphi_e^2(x, y) + \varphi_o^2(x, y)} \\ \varphi_e^2(x, y) &= \phi_e(x, y) \otimes f(x, y) \\ \varphi_o^2(x, y) &= \phi_o(x, y) \otimes f(x, y) \end{aligned} \right\} \quad (1)$$

其中 $f(x, y)$ 为输入图像, $\phi_e(x, y)$ 和 $\phi_o(x, y)$ 分别为偶对称和奇对称的Gabor滤波器。为了简化模型, 本文使用各向同性的Gabor滤波器。

$$\left. \begin{aligned} \phi_e(x, y, f, \theta, \sigma) &= g(x, y, \sigma) \cdot \cos[(2\pi f(x \cos \theta + y \sin \theta))] \\ \phi_o(x, y, f, \theta, \sigma) &= g(x, y, \sigma) \cdot \sin[(2\pi f(x \cos \theta + y \sin \theta))] \end{aligned} \right\} \quad (2)$$

其中 $g(x, y, \sigma)$ 是高斯函数。

$$g(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right] \quad (3)$$

其中 f 为空间频率, θ 为空间相位, σ 为空间常数, 它们分别为Gabor滤波器中的重要参数。

每一个通道的Gabor滤波器对应于一个特定的空间频率和方向。特征提取同时抽取频率信息和方向信息; 对于图像的纹理识别, 没有必要选择覆盖整个频域的滤波器参数空间^[4]。滤波器的中心频率越小, 提取的纹理特征的尺度越大。通常, 中心频率选择为2的幂次方。本文算法选择中心频率为2, 4,

8, 16, 32和64 Hz。对于每个中心频率, 选择4个相位角为 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 。这样, 一共有24个Gabor滤波通道。对每个通道的滤波结果, 提取其幅值作为代表该通道的特征。因此, 对于每一幅输入图像, 多通道Gabor滤波共抽取48个特征。

3 改进投影梯度非负矩阵分解的特征描述

3.1 NMF

非负矩阵分解 (Non-negative Matrix Factorization, NMF)是由 Lee 和 Seung 于 1999 年提出的^[5]。Lee^[6]和 Gabriele^[7]分别使用 NMF 对文本和图像进行特征提取, 表明效果优于传统的算法。

非负矩阵分解可以定义为一个 $m \times n$ 的非负矩阵 $V_{n \times m}$ 可以分解为两个非负矩阵 $W_{n \times r}$ 和 $H_{r \times m}$ 的乘积, 关系如下式所示:

$$V_{n \times m} = W_{n \times r} H_{r \times m} \quad (4)$$

其中 W 和 H 为 V 的非负矩阵分解, 其中 $W_{n \times r}$ 称为基矩阵, $H_{r \times m}$ 称为系数矩阵。可以将 V 中的第 i 列表示成 $v_i = Wh_i$, 则数据 v_i 则是 W 矩阵列的正线性组合, 系数为 H 中的元素值。

3.2 IPGNMF

我们取滤波器组 $\phi_e^{F(i), \theta(j)}(x, y)$ 和 $\phi_o^{F(i), \theta(j)}(x, y)$, 其中 $i = 1, 2, 3, 4, 5, 6$, $j = 1, 2, 3, 4$, 对 $f(x, y)$ 进行系统滤波, 最终可以一组滤波图像的幅值 $A_e^{F(i), \theta(j)}(x, y)$ 和 $A_o^{F(i), \theta(j)}(x, y)$ 输出。

$$A_e^{F(i), \theta(j)}(x, y) = \text{abs}[f(x, y) \otimes \phi_e^{F(i), \theta(j)}(x, y)] \quad (5)$$

$$A_o^{F(i), \theta(j)}(x, y) = \text{abs}[f(x, y) \otimes \phi_o^{F(i), \theta(j)}(x, y)] \quad (6)$$

为了更方便地提取特征, 将每一组滤波后的幅值图像按行展开, 形成一个行向量 $A_o(i, j)$ 和 $A_e(i, j)$, 最终就可得到该图像的纹理特征 A 。

$$A = [A_e(1, 1), A_e(1, 2), \dots, A_e(6, 4), A_o(1, 1), A_o(1, 2), \dots, A_o(6, 4)] \quad (7)$$

其中 $A_e(i, j)$ 为 $A_e^{F(i), \theta(j)}(x, y)$ 对应的的行向量, $A_o(i, j)$ 为 $A_o^{F(i), \theta(j)}(x, y)$ 对应的的行向量。 A 中包含了描述 $f(x, y)$ 纹理的重要信息(包含 6 个方向, 4 个尺度)。

NMF算法的时间复杂度很高, 算法的总体性能也因此显得优势非常有限。所以引入快速的迭代规则优化NMF方法, 对于提高算法的性能至关重要。投影梯度(Projected Gradient, PG)优化方法通过大幅降低NMF迭代的时间复杂度。采用PG优化的NMF方法兼具物理意义明确、数据描述稀疏有效、分类精度提升以及时间消耗大幅降低等特点^[8]。Lin^[8]提出了一种新的投影梯度的计算方法, 大大减少迭代的复杂度。算法的思想如下:

$$\left. \begin{aligned} \min_{\mathbf{H}} \bar{f}(\mathbf{H}) &= \frac{1}{2} \|\mathbf{V} - \mathbf{WH}\|_F^2 = \sum_i \sum_j (\mathbf{V}_{ij} - (\mathbf{WH})_{ij})^2 \\ \text{s.t. } \mathbf{H}_{bj} &\geq 0, \forall b, j \end{aligned} \right\} \quad (8)$$

$$\left. \begin{aligned} \min_{\mathbf{W}} \bar{f}(\mathbf{W}) &\equiv \frac{1}{2} \|\mathbf{V}^T - \mathbf{H}^T \mathbf{W}^T\|_F^2 = \sum_i \sum_j (\mathbf{V}_{ji} - (\mathbf{WH})_{ji})^2 \\ \text{s.t. } \mathbf{W}_{ib} &\geq 0, \forall i, b \end{aligned} \right\} \quad (9)$$

对式(8), 如果 $\bar{\mathbf{H}}$ 是得到的结果, 就使用投影梯度 $\tilde{\mathbf{H}} \equiv P(\bar{\mathbf{H}} - \alpha \nabla \bar{f}(\bar{\mathbf{H}}))$ 更新当前结果 $\bar{\mathbf{H}}$ 为 $\tilde{\mathbf{H}}$ 。

投影梯度 $P(x)$ 定义为 $P(x) = \max(x, 0)$, 取 x 和 0 之间的最大值。投影梯度中的 α 为步长, 步长的检测是投影梯度法中最耗时的部分, 更新 α 算法见文献^[8]。

验证 $\tilde{\mathbf{H}} \equiv P[\bar{\mathbf{H}} - \alpha \nabla \bar{f}(\bar{\mathbf{H}})]$ 所需要的时间为 $O(nmr)$ 。如有 t 个 $\bar{\mathbf{H}}$, 那么时间复杂度则为 $O(tnmr)$ 。

为了再次降低时间复杂度, 将 α 的迭代条件设为

$$(1 - \sigma) \langle \bar{f}(\bar{\mathbf{H}}), \tilde{\mathbf{H}} - \bar{\mathbf{H}} \rangle + \frac{1}{2} \langle \tilde{\mathbf{H}} - \bar{\mathbf{H}}, (\mathbf{W}^T \mathbf{W})(\tilde{\mathbf{H}} - \bar{\mathbf{H}}) \rangle \leq 0 \quad (10)$$

式(10)的计算量主要在于 $(\mathbf{W}^T \mathbf{W})(\tilde{\mathbf{H}} - \bar{\mathbf{H}})$, 其时间复杂度只有 $O(mr^2)$ 。

4 分类识别

RBF 神经网络是一种性能良好的前向神经网络模型。已经证明它具有全局逼近的性质, 且不存在局部最小问题。RBF 网络不仅具有良好的推广能力, 而且计算量少, 学习速度也比其它一般算法快得多。径向基函数网络包括三层: 输入层, 隐层和线性输出层。图 2 为径向基函数网络的结构图。

5 实验设置及其结果

5.1 实验的人脸库

为了验证算法的有效性, 本文采用 ORL 库、YALE 库和 FERET 库进行测试。

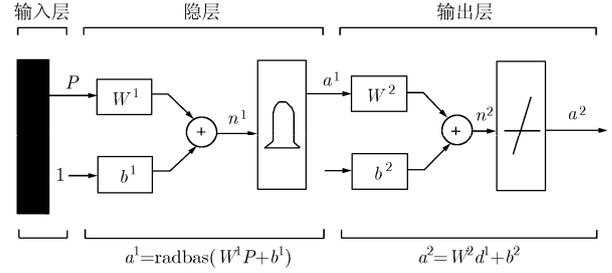


图 2 RBF 结构图

实验环境: 程序环境: MATLAB 7.0; CPU: P4 3.6 G; 内存: 512 M。

5.2 样本数对识别率的影响

为了验证训练样本多少对识别率的影响, 对 ORL 数据库, 使用传统的 PCA 算法进行了测试, 分别取训练样本为 8 个, 7 个, 依次递减, 直到单样本, 测试样本的个数为全样本 10 个。从试验结果可以看出, 识别率随着训练样本数目的递减, 呈降低趋势, 当训练样本大于 4 个时, 识别率保持在 90% 以上, 当训练样本数目为 3 个以下时, 识别率骤减, 训练为单样本时, 识别率降到了 50%, 比最高识别率降低高达 40%, 所以一种有效的单样本人脸识别算法是非常有必要的。

5.3 不同的频率和相对对识别结果的影响

本实验主要针对 Gabor 滤波器的尺度和频率的设置对识别率的影响, 对于 Gabor 滤波器组设 f 为空间频率, θ 为相位, 试验设置了 f 可以取 2 组(2,4), 4 组(2,4,8,16), 6 组(2,4,8,16,32,64), 8 组(2,4,8,16,32,64,128,256), 分别标示为 F_2, F_4, F_6, F_8 , θ 可以取 2 组(0, π), 4 组(0, $\pi/4, \pi/2, 3\pi/4$), 6 组(0, $\pi/6, \pi/3, \pi/2, 2\pi/3, 5\pi/6$), 8 组(0, $\pi/8, \pi/4, 3\pi/8, \pi/2, 5\pi/8, 3\pi/4, 7\pi/8$), 分别标示为 $\theta_2, \theta_4, \theta_6, \theta_8$, 共有 16 种组合, 标示为 $F_2\theta_2, F_2\theta_4, \dots, F_8\theta_8$ 。

试验的数据库选择 FERET 库, 该库中选择 50 个人随机选取 2 幅图像作为训练样本, 6 幅图像作为测试本。实验结果如图 3 所示, Gabor 的分辨率指频率和相位取值的组合。

通过试验结果可以看出, 当 f 取 6 组, θ 取 4 组时识别率最高, 这是由于频率和相位分得太少, 就不能提供不同角度和不同频率丰富的纹理特征, 但是当纹理特征过于细致后, 会陷入局部细节, 对全局表现不佳, 导致类间的可分性降低。

5.4 关于特征提取时间的实验

首先实验主要针对 NMF 运算速度比较慢的情况, 分别对 NMF, PGNMF, IPGNMF 的运算时间作以比较, 人脸库为 ORL 库, 实验结果如图 4

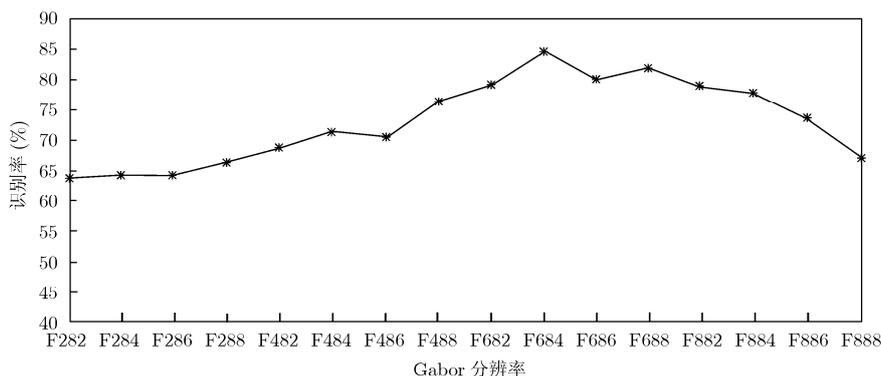


图 3 不同的采样频率和尺度对识别结果的影响

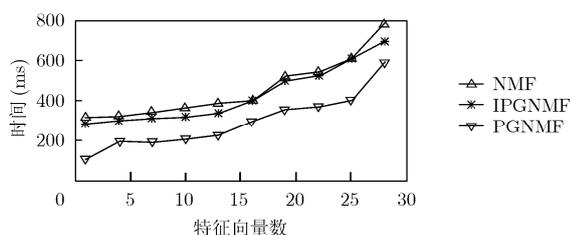


图 4 特征提取时间随特征个数的变化

所示。从图中可以看出, PGNMF 的运算时间最短, NMF 的运算时间最长, IPGNMF 的运算时间处于两者之间, 但是其识别率比 NMF 和 PGNMF 都要高, 见表 1。

表 1 不同识别策略的识别率

识别方法	ORL 识别率	YEL 识别率	FERET 识别率	平均识别率
GABOR	0.572	0.520	0.556	0.549
GABOR+PCA	0.644	0.640	0.623	0.636
GABOR+ICA	0.733	0.727	0.653	0.704
GABOR+2DPCA	0.797	0.820	0.676	0.764
GABOR+NMF	0.721	0.785	0.747	0.751
GABOR+PGNMF	0.704	0.816	0.787	0.769
GABOR+IPGNMF	0.824	0.833	0.8433	0.833

其次由于在实验中引入了 Gabor 变换, 对于一副输入的人脸图像, 提取 Gabor 特征, 特征融合的时间的讨论主要针对本文中的几个人脸库中图像进行, 结果如表 2 所示。通过结果可以看出, Gabor 特征提取和融合的时间平均是 188 ms, 这对识别的实时性是有影响的, 在以后的研究工作中要进一步的缩短这个时间。

5.5 单幅人脸不同特征提取方法的识别率比较

本实验设置主要针对常见的识别方法和本文方法的识别率做以比较, 实验的结果如表 1 所示。通

表 2 Gabor 特征提取和融合的时间

人脸库	时间
ORL	183 ms
YALE	175 ms
FERET	206 ms

过试验结果可以看出, 本文的方法比直接使用 GABOR 变换, 主成分分析(PCA), 独立分量分析(ICA), NMF(非负矩阵分解), PGNMF(梯度投影的非负矩阵分解)等算法识别效果好。

6 结束语

本文针对传统的人脸识别方法对单训练样本识别的问题和缺陷, 提出了一种基于改进的投影梯度非负矩阵分解的单训练样本特征提取新方法。该方法是在 IPGNMF 运算之前, 首先对样本图像块进行多方向、多分辨率 Gabor 小波滤波, 并提取其对应不同方向、不同尺度的多个 Gabor 幅值域图谱 GMM(Gabor Magnitude Map), 然后使用投影梯度非负矩阵分解的方法提取每个图谱的 NMF 特征, 再将每个图谱的特征融合, 最后采用 RBF 神经网络的方法对融合后的特征识别分类, 最后采用神经网络分类识别。在对人脸库 ORL、YEL、FERET 的识别实验中, 与 GREY, PCA, ICA, NMF, PGNMF 等方法比较, 证明了可以有效地解决单训练样本人脸识别的问题。对于在复杂条件下的训练样本典型特征提取, 有待于深入的研究。

参考文献

- [1] Zhu Yulian, Liu Jun, and Chen Songcan. Semi-random subspace method for face recognition[J]. *Image and Vision Computing*, 2009, 9(26): 1-13.
- [2] Gao Quan-xue, Zhang Lei, and Zhang David. Face recognition using FLDA with single training image per person[J]. *Applied Mathematics and Computation*, 2008,

- 205(2): 726-734.
- [3] 李瑞东, 祝磊, 余党军, 陈偕雄. 基于判别公共向量的单训练样本人脸识别[J]. 浙江大学学报, 2008, 35(2): 181-184.
Li Rui-dong, Zhu Lei, Yu Dang-jun, and Chen Xie-xiong. Making discriminative common vectors applicable to face recognition with one training image per person[J]. *Journal of ZheJiang University*, 2008, 35(2): 181-184.
- [4] Daugman J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two2dimensional visual cortical filters [J]. *Journal of the Optical Society of America (A)*, 1985, 2(7): 1160-1169.
- [5] 李乐, 章毓晋. 非负矩阵算法综述[J]. 电子学报, 2008, 36(4): 737-743.
Li Le and Zhang Yu-jin. A survey on algorithms of non-negative matrix factorization[J]. *Acta Electronica Sinica*, 2008, 36(4): 737-743.
- [6] Lee Ju-Hong, Park Sun, Ahna Chan-Min, and Kim Daeho. Automatic generic document summarization based on non-negative matrix factorization[J]. *Information Processing and Management*, 2008, 6(2): 20-34.
- [7] Lohmann G, Volz K G, and Ullsperger M. Using non-negative matrix factorization for single-trial analysis of fMRI data[J]. *NeuroImage*, 2007, 37(4): 1148-1160.
- [8] Lin C J. Projected gradient methods for non-negative matrix factorization[J]. *Neural Computation*, 2007, 19(10): 2756-2779.
- 高 涛: 男, 1980年生, 博士生, 研究方向为智能图像处理、神经网络.
- 何明一: 男, 1958年生, 教授, 博士生导师, 陕西省信息获取与处理重点实验室主任, 研究方向为图像处理、智能信息处理、遥感图像处理.