

# 多核学习融合局部和全局特征的人脸识别算法

杨 赛<sup>1</sup>, 赵春霞<sup>2</sup>, 刘 凡<sup>3</sup>

(1. 南通大学电气工程学院, 江苏南通 226019; 2. 南京理工大学计算机科学与工程学院, 江苏南京 210094;  
3. 河海大学计算机与信息学院, 江苏南京 210098)

**摘要:** 提出一种基于词袋模型的新的人脸识别算法. 该方法将词袋模型和词袋模型的全局模式分别作为人脸图像的局部特征和全局特征描述, 最后使用多核学习方法将二者进行融合. AR、FERET、CMU PIE 以及 LFW 公开人脸数据库上的实验结果表明, 本文方法能够更好的解决小样本问题, 并且对人脸的表情变化、姿态变化以及面部遮挡具有更优良的鲁棒性.

**关键词:** 词袋模型; 全局特征; 多核学习; 人脸识别

**中图分类号:** TP391.4      **文献标识码:** A      **文章编号:** 0372-2112 (2016)10-2344-07

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2016.10.009

## Fusion of Local and Global Features Using Multiple Kernel Learning for Face Recognition

YANG Sai<sup>1</sup>, ZHAO Chun-xia<sup>2</sup>, LIU Fan<sup>3</sup>

(1. School of Electrical Engineering, Nantong University, Nantong, Jiangsu 226019, China;  
2. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China;  
3. College of Computer and Information, Hohai University, Nanjing, Jiangsu 210098, China)

**Abstract:** A new face recognition algorithm via bag-of-words (BoW) is proposed. In specific, it uses BoW and the global pattern of BoW respectively as the local feature and global feature of face images. Multiple kernel learning is adopted to fuse the local and global features. Extensive experiments were carried out on four face databases, i. e. AR, FERET, CMU PIE and LFW. The results show that our method can effectively solve the small training size problem and is more robust to expression changes, position variations and occlusion.

**Key words:** bag-of-words; global feature; multiple kernel learning; face recognition

### 1 引言

人脸识别因其在身份验证、视频监控、疑犯追踪等领域的广泛应用而成为当前计算机视觉研究中的热点. 人脸识别的核心步骤之一为特征提取, 可分为全局特征和局部特征. 前者缺乏对局部变化的鲁棒性, 而目前人脸识别中所使用的局部特征仅限于底层视觉特征, 因此本文将更具有判别能力的中层语义特征—词袋模型 (Bag-of-Words, BoW) 应用到人脸识别领域.

全局与局部特征之间并不存在严格的界限, 当局部表示中用于提取特征的图像区域变大时, 局部表示的全局性逐渐变强. 受此启发, 本文提出词袋模型的全局模式 (Global Bag-of-Words, GBoW), 即将提取底层特

征的区域扩展到整幅图像. 相关研究表明, 人类是综合利用全局和局部信息对人脸进行辨识, 因而一些学者提出融合全局和局部特征的识别算法, 通常是在决策层将二者融合, 不可避免地会损失部分判别信息, 为此本文提出一种基于多核学习 (Multiple Kernel Learning, MKL) 的局部和全局特征融合算法.

### 2 相关工作

如何对人脸模式进行有效表示是自动人脸识别的关键步骤. 基于子空间的特征表示是目前的主流方法, 主成分分析<sup>[1]</sup> (Principal Component Analysis, PCA) 是其中的典型方法之一, 该方法寻找最小均方差意义下的最优主成分分量, 通过这些主成分的线性组合重构原

始样本,从而达到最大限度去除图像特征中冗余信息的目的;而 Fisher 线性判别分析<sup>[2]</sup> (Fisher Linear Discriminative Analysis, FLDA) 寻找最优投影方向使得同类样本的离散程度和不同类样本的交叠程度在新的坐标系内同时达到最小,从而取得最佳的分类效果。

近年来,稀疏表示理论在人脸识别中的优异性能引起了学者们的广泛研究兴趣。Wright 等<sup>[3]</sup> 提出了一种稀疏表示分类 (Sparse Representation Classification, SRC) 方法,该方法通过求解带  $l_1$  范数的最小化问题,使用训练样本线性组合给定测试样本,根据系数向量的稀疏性对测试样本进行分类决策;Zhang 等<sup>[4]</sup> 提出协同表示分类 (Collaborative Representation Classification, CRC) 方法,该方法使用  $l_2$  范数代替  $l_1$  范数求解原问题,从而解决稀疏表示在人脸识别中的实时性问题;与上述方法使用所有训练样本构建冗余词典不同,Yang 等<sup>[5]</sup> 提出构建紧致稳健字典的 Metaface 学习算法,该算法在每类训练样本中学习一组字典向量,根据类别词典上的稀疏重构残差对测试人脸图像进行分类判别;Yang 等<sup>[6]</sup> 提出利用 Fisher 准则函数在训练样本中寻找一组最优字典基向量。然而此类人脸识别方法仍然以提取图像中的全局特征为基础。

由于局部特征对光照、表情和姿态变化具有很好的鲁棒性,近年来在人脸识别中得到了广泛应用。例如,Gabor 函数与人类视觉皮层感受野具有一致性,对于提取图像局部变化特征具有良好的特性,然而多尺度、多方向的 Gabor 变换会使人脸特征的维数急剧增加,为此文献[7]同时使用下采样和 PCA 降低 Gabor 特征的维数,然后使用增强 Fisher 鉴别分析完成识别;针对上述方法会造成大量分类特征信息丢失的问题,文献[8]利用 LBP 算子对 Gabor 特征进行编码,计算直方图统计特

征作为分类判别依据;为了最大化利用特征信息并得到人脸特征的紧致描述,文献[9]将图像划分为若干小区域,计算小区域内所有像素 Gabor 系数的均值和方差,然后将所有区域特征拼接起来描述人脸图像;最近,Yang 等<sup>[10]</sup> 采用多尺度 LBP 金字塔模型描述人脸图像,提出一种基于局部统计特征的鲁棒核函数 (Statistical Local Feature based Robust Kernel Representation, SLF-RKR) 度量图像之间的相似性计算重建残差,测试样本最终被判定为重建残差最小的训练样本对应的类别。

另外,子图像方法也是人脸识别中获得成功应用的局部特征。文献[11]将人脸图像分割为相等大小的子图像,然后使用基于子图像构建的分量分类器对测试样本的相应子图像进行判别,最后使用加权投票方式对分类结果进行集成形成最后决策;文献[12]利用类内散度最小类间散度最大原则训练 Volterra 核函数建立子图像与类别之间的映射,以学习得到的映射作为相似性度量建立最近邻分类器对测试样本的相应子图像进行分类,最后使用多数投票方法融合所有分类结果;最近,Zhu 等<sup>[13]</sup> 提出基于协同表示的子图像分类方法 (Multi-Scale Patch based Collaborative Representation Classification, MSPCRC),该方法提取若干尺寸不等的子图像表征尺度信息,并利用 CRC 方法对测试样本的相应子图像进行分类,最后使用多数投票方式融合分类结果,可以很好的解决小样本问题。

### 3 本文方法

本文提出的多核学习融合局部和全局特征的人脸识别算法的总体框架如图 1 所示,各个环节的具体描述如图 1。

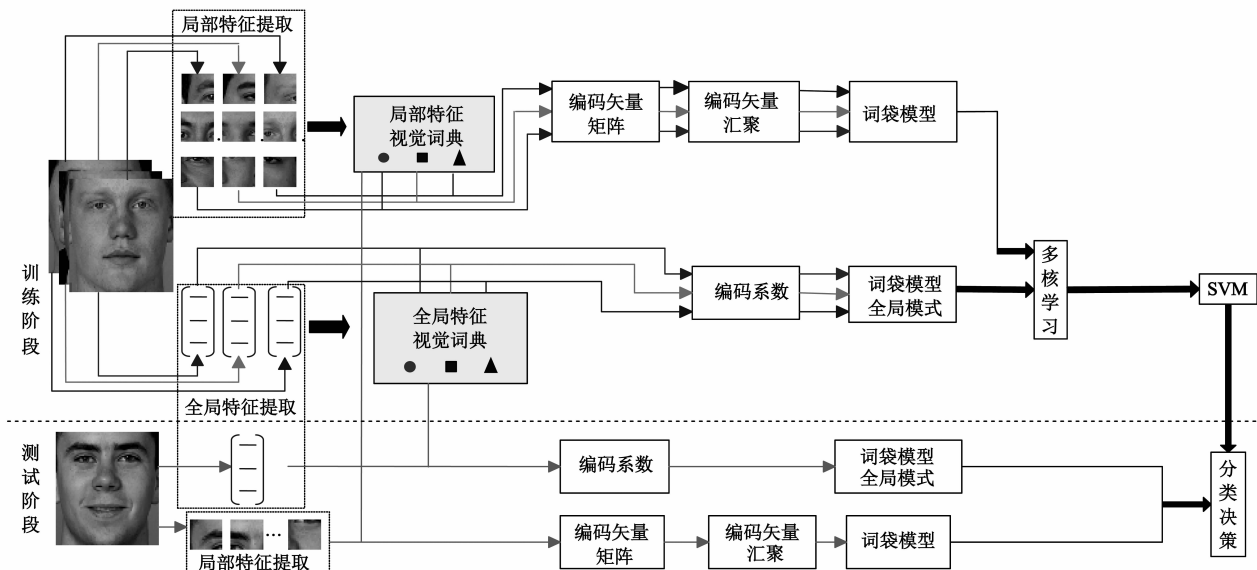


图1 本文人脸识别方法的总体框架

### 3.1 基于词袋模型的人脸特征描述

对于人脸图像数据集  $I = \{I_1, I_2, \dots, I_N\}$ , 经过局部特征提取之后, 所有训练图像被表示为  $Y = \{y_1, y_2, \dots, y_N\}$ , 使用  $K$  均值算法对此局部特征集合  $Y$  进行聚类得到局部特征视觉词典  $V_L = [v_1, v_2, \dots, v_K] \in R^{D \times K}$ ,  $v_k \in R^{D \times 1}$ ,  $k = 1, 2, \dots, K$  表示第  $k$  个视觉单词向量,  $K$  为视觉单词数目,  $D$  为局部特征的维数. 假设  $X = [x_1, x_2, \dots, x_T] \in R^{D \times T}$  表示数据集  $I$  中任意一幅人脸图像的局部特征集合, 其中  $x_t \in R^{D \times 1}$ ,  $t = 1, 2, \dots, T$  表示第  $t$  个局部特征矢量. 在使用人脸数据库中所有图像的线性组合来表示一幅给定人脸图像时, 其系数向量是稀疏的, 其中同一类别图像的组合系数不为 0, 稀疏表示能够提供人脸图像合适的描述. 同样对于人脸图像中的局部特征来说, 稀疏编码也能够保证相似局部特征的组合系数不为 0, 对局部特征进行稀疏编码的表达式为:

$$u_t = \arg \min_{u_t \in R^k} \|x_t - V_L u_t\|_2 + \lambda \|u_t\|_1 \quad (1)$$

式(1)中,  $u_t \in R^{K \times 1}$  表示第  $t$  个局部特征的编码矢量, 等式(1)右边第一项为拟合项, 表示利用  $V_L$  重构  $x_t$  所产生的误差, 第二项为稀疏惩罚项, 保证对  $x_t$  的表示是稀疏的, 稀疏正则系数  $\lambda$  控制两项的比重, 对其进行求解就得到  $x_t$  在  $V_L$  上的投影系数  $u_t$ .

完成上述编码环节后, 图像此时被表示为一个由  $T$  个  $K$  维矢量组成的集合, 需要计算编码矢量的汇聚特征才能使用传统分类器对图像进行分类. 人脸图像中的内容是按照一定的规律组织的, 图像子区域之间的空间关系对于分类来说是有用的信息, 因此使用空间金字塔模型对局部编码矢量进行汇聚. 记  $l = 0, 1, \dots, L$  为此模型的层次, 总层数为  $L + 1$ , 则第  $l$  层在水平和垂直方向都将图像分成  $2^l$  块, 图像子区域的总数为  $2^l \times 2^l = 4^l$  块. 图 2 展示的是使用划格方法得到金字塔的第 0, 1, 2 层, 此模型能够实现各子区域之间的匹配, 即使图像出现光照、表情、姿态以及遮挡等变化, 也能够保证不变区域之间的匹配, 因此能够进一步提高词袋模型对各种变化的鲁棒性. 假设第  $l$  层金字塔模型的第  $i$

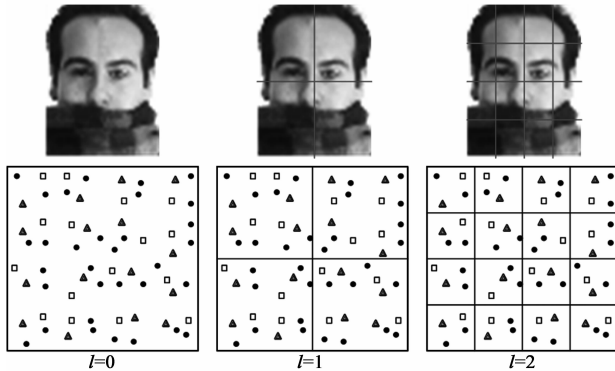


图2 空间金字塔模型(注:  $\square \triangle \circ$  表示视觉单词)

个子区域内有  $T_l$  个编码矢量, 此区域内编码矢量的最大统计值的计算公式为:

$$B_{li} = \max_{t=1, 2, \dots, T_l} u_t \quad (2)$$

将所有子区域内的统计特征拼接起来作为图像的第  $l$  层特征  $B_l = [B_{l1}, B_{l2}, \dots, B_{l4}^l]$ , 则人脸图像的局部特征表示为  $d_B = [B_1, B_2, \dots, B_{L+1}]$ .

### 3.2 基于全局特征的人脸特征描述

本文将词袋模型的全局模式作为人脸图像的全局特征描述, 即将此模型提取底层特征的区域扩展到整幅图像, 并利用人脸识别中经典的子空间特征表示方法对其进行描述, 例如 PCA、FLDA、ICA、LPP 等. 本文对上述四种经典全局特征进行了测试, 其中 FLDA 引入了鉴别信息, 性能最优, 故本文均采用 FLDA 作为整幅图像的底层特征提取方法.

对于大小为  $p \times q$  的人脸图像, 通过首尾连接每列像素将其转换为向量, 使用 FLDA 对其进行变换后, 所有训练图像被表示为  $Z = \{z_1, z_2, \dots, z_N\}$ . 与词袋模型一致, 使用  $K$  均值算法对此全局特征集合  $Z$  进行聚类得到全局视觉词典  $V_C = [v_1, v_2, \dots, v_M] \in R^{C \times M}$ ,  $M$  为视觉单词的数目,  $v_m \in R^{C \times 1}$ ,  $m = 1, 2, \dots, M$  表示第  $m$  个视觉单词向量,  $d_C \in R^{C \times 1}$  表示相应的编码矢量,  $C$  是 FLDA 特征的维数, 对特征  $z$  进行稀疏编码的表达式为:

$$d_C = \arg \min_{d_C \in R^M} \|z - V_C d_C\|_2 + \lambda \|d_C\|_1 \quad (3)$$

式(3)右边第一项为拟合项, 表示利用  $V_C$  重构  $Z$  所产生的误差, 第二项为稀疏惩罚项, 保证对  $Z$  的表示是稀疏的, 常数  $\lambda$  控制两项的比重, 对其进行求解就可以得到  $z$  在  $V_C$  上的投影系数  $d_C$ .

### 3.3 基于多核学习的局部与全局特征的融合

对于 3.1 得到人脸图像特征  $d_B$ , 假设所使用的核函数为  $K_B$ , 相应的权系数为  $h_B$ , 对于 3.2 得到人脸图像特征  $d_C$ , 假设所使用的核函数为  $K_C$ , 相应的权系数为  $h_C$ . 此时任意第  $i$  幅与第  $j$  幅人脸图像之间的相似性度量量为:

$$K(i, j) = h_B K_B(d_{Bi}, d_{Bj}) + h_C K_C(d_{Ci}, d_{Cj}) \quad (4)$$

式(4)中,  $d_{Bi}, d_{Ci}$  分别表示第  $i$  幅图像的局部特征和全局特征,  $d_{Bj}, d_{Cj}$  分别表示第  $j$  幅图像的局部特征和全局特征, 如果分类器选用支持向量机, 二分类的情况下, 多核学习方法的目标函数为:

$$\begin{aligned} & \min \frac{1}{2} \left( \frac{1}{h_B} \|w_B\|^2 + \frac{1}{h_C} \|w_C\|^2 \right) + C \sum_{i=1}^N \xi_i \\ & \text{s. t. } y_i (w_B^T \cdot d_{Bi} + w_C^T \cdot d_{Ci} + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, N \\ & h_B, h_C \geq 0, h_B + h_C = 1 \end{aligned} \quad (5)$$

式(5)中,  $y_i$  为图像的类别标号,  $C$  为惩罚系数,  $\xi_i$  为松弛变量,  $w_B, w_C$  为分类器的参数,  $N$  为图像样本数目. 采

用分块协调下降算法求解式(5),首先固定  $h_B$  和  $h_C$ ,式(5)转化为关于  $\xi_i, \mathbf{w}_B, \mathbf{w}_C$  的带线性约束的二次规划问题;然后固定参数  $\xi_i, \mathbf{w}_B, \mathbf{w}_C$ ,式(5)转化为关于  $h_B$  和  $h_C$  的带线性约束的非线性最小化问题. 交替求解上述两个问题,直到满足收敛条件就可以得到  $s$  个系数不为 0 的支持向量以及  $h_B$  和  $h_C$  的值,则对第  $j$  幅测试图像的分类判决函数为:

$$f(\mathbf{d}_{Bj}, \mathbf{h}_{Cj}) = \operatorname{sgn} \left\{ \sum_{i=1}^s \gamma_i \alpha_i [h_B K_B(\mathbf{d}_{Bj}, \mathbf{d}_{Bi}) + h_C K_C(\mathbf{h}_{Cj}, \mathbf{h}_{Ci})] + b^* \right\} \quad (6)$$

式(6)中,  $\alpha_i, i=1, 2, \dots, s$  表示第  $i$  个支持向量对应的系数,  $b^*$  为分类阈值,可以用任一个支持向量求得.

## 4 实验与分析

在公开人脸数据库将本文方法与目前主流的 PCA<sup>[1]</sup>、FLDA<sup>[2]</sup>、CRC<sup>[4]</sup>、FDDL<sup>[6]</sup>、RKR-SLF<sup>[10]</sup>、Volterrar<sup>[12]</sup>、MSPCRC<sup>[13]</sup> 方法进行对比,使用作者提供的源代码和参数设置进行仿真实验. 同时实验还给出单独使用词袋模型(BoW)和全局特征(GBoW)进行分类的性能. 实验过程中通过尝试不同的参数设置得到最优结果. 建立词袋模型过程中,实验结果显示图像小块尺寸越小,其表征性能越优,然而太小不足以提取 SIFT 特征,因此将小块的尺寸设为  $8 \times 8$ ;并且密集采样步长越小,局部特征数目越多,其表征性能越优,因此将采样步长设为 1;将视觉单词数目和空间金字塔模型的层数分别选定为 1500 和 3,这是因为视觉单词的数目过少,不能充分编码图像中的局部特征,金字塔模型的层数过少,不能充分利用图像的空间信息,然而当单词数目大于 1500 以及层数大于 3 时所带来的计算复杂度的代价将远大于其分类正确率增加所带来的性能改善;另外,参数  $\lambda$  为 0.5~1 之间,性能相差无几,因此将稀疏正则系数设为 0.5. 多核学习过程中,基核函数为线性核函数,惩罚系数  $C$  设为 10 以上,性能相差无几,将外层和内层循环的迭代次数设定为 200,足以求解多核学习模型的最优解.

### 4.1 AR 数据库上的实验结果

AR 数据库包含 126 个人不同时期拍摄的 26 幅图

表 2 各种人脸识别算法在人脸表情变化子集上的识别率(%)

识别算法	PCA	FLDA	CRC	FDDL	SLF-RKR	Volterrar	MSPCRC	BoW	GBoW	本文方法
识别率	67.22	68.61	73.06	69.17	92.22	62.31	62.78	92.50	71.66	<b>92.77</b>

使用文献[10]中的实验设置验证遮挡对识别算法的影响. 将第 1 个时期的前 7 幅无遮挡图像作为训练数据,将第 1 个时期的第 8~10 张图像(S1\_Sunglass)、第 11~13 张图像(S1\_Scarf)、第 2 个时期的第 8~10 张图像(S2\_Sunglass)、第 11~13 张图像(S2\_Scarf)分别

像,尺寸为  $50 \times 40$  像素,如图 3 所示,不同时期的图像被归为不同的组,第 1 张为无表情变化图像,第 2~4 张为表情变化图像,第 5~7 张为光照变化图像,第 8~10 张为太阳镜遮挡图像,第 11~13 张为围巾遮挡图像.

首先验证不同训练样本数目下识别算法的性能. 分别随机选取每个人的 2,3,4,5 幅图像作为训练数据,剩余的图像作为测试数据,重复选取 10 次. 由表 1 中的结果可知:(1)在不同的训练样本数目下,本文方法的识别率最高,分别高于 SLF-RKR 方法 1.80%、0.32%、0.76%、0.10%,说明本文方法可以很好的解决小样本问题;(2)本文方法的识别率高于 BoW 和 GBoW,说明多核学习能够有效地将互补的局部和全局特征进行结合.

表 1 AR 数据库上不同训练样本数目下各种方法的识别率  $\pm$  标准差 (%)

人脸识别算法	2	3	4	5
PCA	62.28 $\pm$ 4.39	63.99 $\pm$ 3.81	67.27 $\pm$ 2.52	71.35 $\pm$ 5.40
FLDA	66.84 $\pm$ 6.66	74.75 $\pm$ 8.36	79.37 $\pm$ 5.95	83.12 $\pm$ 3.20
CRC	72.54 $\pm$ 6.40	76.87 $\pm$ 7.08	81.30 $\pm$ 4.70	84.85 $\pm$ 4.33
FDDL	77.69 $\pm$ 4.69	78.47 $\pm$ 5.85	85.50 $\pm$ 4.94	90.13 $\pm$ 4.98
SLF-RKR	80.47 $\pm$ 0.55	88.94 $\pm$ 0.89	93.38 $\pm$ 0.60	96.01 $\pm$ 0.61
Volterrar	62.64 $\pm$ 7.60	73.15 $\pm$ 4.27	74.63 $\pm$ 7.23	83.52 $\pm$ 2.79
MSPCRC	71.16 $\pm$ 10.78	71.38 $\pm$ 7.88	77.64 $\pm$ 4.29	80.77 $\pm$ 4.57
BoW	67.52 $\pm$ 1.46	80.08 $\pm$ 1.62	87.94 $\pm$ 1.01	92.02 $\pm$ 0.84
GBoW	79.51 $\pm$ 1.43	83.04 $\pm$ 1.56	91.43 $\pm$ 0.63	93.98 $\pm$ 0.76
本文方法	<b>82.27 <math>\pm</math> 1.33</b>	<b>89.26 <math>\pm</math> 1.32</b>	<b>94.14 <math>\pm</math> 0.74</b>	<b>96.11 <math>\pm</math> 0.51</b>



图 3 AR 数据库的示意图

使用文献[14]中的实验设置验证表情变化对识别算法的影响. 将第 1 个时期拍摄的前 7 幅图像作为训练数据,将第 2 个时期拍摄的第 2~4 张表情变化的图像作为测试数据. 由表 2 中的结果可知:(1)本文方法的识别率最高,高于 SLF-RKR 方法 0.55%;(2)BoW 的识别率高于 GBoW,说明局部特征比全局特征在表情变化中起到更重要的作用.

作为测试数据. 由表 3 中的结果可知:(1)本文方法的识别率最高,分别高于 SLF-RKR 方法 0.28%、0.28%、2.78%、0.44%;(2)只有 BoW 的识别率得到了提升并且高于 GBoW,说明局部特征在遮挡情况下起到更重要的作用.

表 3 各种人脸识别算法在遮挡变化子集上的识别率(%)

人脸识别算法	S1_Sunglass	S1_Scarf	S2_Sunglass	S2_Scarf
PCA	95.83	40.00	62.50	23.61
FLDA	96.94	38.89	60.83	21.67
CRC	96.11	67.22	58.89	33.33
FDDL	97.22	60.83	55.50	29.72
SLF-RKR	99.72	98.33	79.44	82.33
Volterrarr	94.29	75.39	51.33	33.42
MSPCRC	83.33	75.83	45.56	33.33
BoW	100.00	98.33	82.22	82.77
GBoW	97.22	50.00	62.22	23.05
本文方法	<b>100.00</b>	<b>98.61</b>	<b>82.22</b>	<b>82.77</b>

#### 4.2 FERET 数据库上的实验结果

FERET 数据库包含 1199 个人的不同局部变化的图像. 本文选择其中的一个姿态子库作为实验数据, 该子库有 200 个人, 每个人有 7 幅标号不同的图像, 图像的尺寸为  $80 \times 80$  像素, 各种标号的含义如图 4 所示.

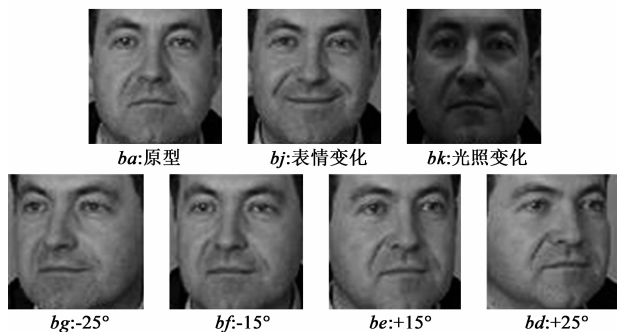


图 4 FERET数据库的示意图

首先验证不同训练样本数目下不同人脸识别算法的性能. 分别随机选取每个人的 2, 3, 4, 5 幅图像为训练数据, 剩余的图像为测试数据, 重复选取 10 次. 由表 4 中的结果可知: (1) 在不同的训练样本数目下, 本文方法的识别率最高, 分别高于 SLF-RKR 方法 12.07%、11.92%、8.71%、8.12%, 说明本文方法可以很好的解决小样本问题; (2) 本文方法的识别率高于 BoW 和 GBoW, 说明多核学习能够有效地将互补的局部和全局特征进行结合.

表 4 FERET 数据库上不同训练数目下各种方法的识别率  $\pm$  标准差 (%)

识别算法	2	3	4	5
PCA	47.64 $\pm$ 8.94	60.20 $\pm$ 8.86	64.88 $\pm$ 11.94	70.45 $\pm$ 11.88
FLDA	47.90 $\pm$ 4.25	65.91 $\pm$ 8.20	68.22 $\pm$ 15.95	78.50 $\pm$ 15.64
CRC	44.26 $\pm$ 3.73	56.93 $\pm$ 7.06	57.93 $\pm$ 15.06	67.35 $\pm$ 14.79
FDDL	50.42 $\pm$ 3.52	59.89 $\pm$ 7.51	62.89 $\pm$ 12.07	73.50 $\pm$ 15.20
SLF-RKR	51.97 $\pm$ 1.23	67.59 $\pm$ 1.55	78.72 $\pm$ 1.18	85.13 $\pm$ 2.04
Volterrarr	46.85 $\pm$ 13.53	58.94 $\pm$ 12.08	59.07 $\pm$ 13.45	74.40 $\pm$ 6.69
MSPCRC	47.60 $\pm$ 7.15	58.13 $\pm$ 10.60	63.87 $\pm$ 12.36	69.95 $\pm$ 15.65
BoW	58.85 $\pm$ 1.41	72.06 $\pm$ 1.41	82.03 $\pm$ 1.32	88.35 $\pm$ 0.94
GBoW	54.94 $\pm$ 1.36	71.55 $\pm$ 2.09	82.45 $\pm$ 1.57	87.42 $\pm$ 1.87
本文方法	<b>64.04 <math>\pm</math> 1.32</b>	<b>79.51 <math>\pm</math> 1.90</b>	<b>87.43 <math>\pm</math> 1.14</b>	<b>93.25 <math>\pm</math> 1.14</b>

使用文献[10]中的实验设置验证姿态变化对识别算法的影响. 将标号为 'ba', 'bj', 'bk' 的图像作为训练数据, 标号为 'bg', 'bf', 'be', 'bd' 的图像分别作为测试数据. 由表 5 中的结果可知: (1) 当姿态角从  $-25^\circ$  变化到  $+25^\circ$ , 本文方法的识别率最高并且非常稳定, 分别高于 SLF-RKR 方法 49.00%、0.50%、11.00%、42.00%, 表明本文方法对姿态的变化具有更好的鲁棒性; (2) BoW 的识别率高于 GBoW, 说明局部特征在姿态变化时起到更重要的作用.

表 5 各种人脸识别算法在姿态变化子集上的识别率(%)

人脸识别算法	bg	bf	be	bd
PCA	21.00	45.00	57.50	14.50
FLDA	17.00	40.00	72.00	16.00
CRC	15.00	33.50	56.50	17.50
FDDL	8.00	64.00	29.50	13.00
SLF-RKR	41.00	98.00	85.00	39.50
Volterrarr	7.22	26.33	44.65	10.00
MSPCRC	16.50	69.00	36.00	16.50
BoW	90.00	98.50	96.00	81.50
GBoW	12.50	73.50	35.00	18.50
本文方法	<b>90.00</b>	<b>98.50</b>	<b>96.00</b>	<b>81.50</b>

表 6 CMU PIE 数据库上不同训练数目下各种方法的识别率  $\pm$  标准差 (%)

人脸识别算法	2	3	4	5
PCA	19.81 $\pm$ 3.42	24.82 $\pm$ 3.50	29.14 $\pm$ 4.26	32.44 $\pm$ 3.49
FLDA	42.42 $\pm$ 9.41	58.02 $\pm$ 6.41	63.23 $\pm$ 5.05	67.25 $\pm$ 5.01
CRC	43.78 $\pm$ 5.97	56.48 $\pm$ 4.46	65.53 $\pm$ 3.99	70.57 $\pm$ 4.58
FDDL	48.68 $\pm$ 5.71	61.91 $\pm$ 6.79	72.00 $\pm$ 5.43	76.59 $\pm$ 4.95
SLF-RKR	50.38 $\pm$ 1.03	61.90 $\pm$ 2.03	69.27 $\pm$ 2.01	75.47 $\pm$ 1.11
Volterrarr	32.08 $\pm$ 6.81	50.12 $\pm$ 3.47	58.29 $\pm$ 4.83	62.20 $\pm$ 6.44
MSPCRC	51.60 $\pm$ 5.40	61.75 $\pm$ 4.53	68.27 $\pm$ 4.09	72.14 $\pm$ 4.05
BoW	45.73 $\pm$ 1.52	57.18 $\pm$ 1.87	66.48 $\pm$ 2.24	72.32 $\pm$ 1.64
GBoW	50.07 $\pm$ 2.08	60.69 $\pm$ 1.64	67.79 $\pm$ 1.25	70.68 $\pm$ 1.05
本文方法	<b>52.52 <math>\pm</math> 1.76</b>	<b>64.29 <math>\pm</math> 1.55</b>	<b>72.73 <math>\pm</math> 1.22</b>	<b>76.68 <math>\pm</math> 0.84</b>

#### 4.3 CMU PIE 数据库上的实验结果

CMU PIE 包含 68 个人的不同局部变化的图像, 本文使用文献[15]提供的五个子集进行实验, 示例图像见图 5. 在此数据库上验证不同训练样本数目下人脸识别算法的性能, 分别随机选取每个人的 2, 3, 4, 5 幅图像作为训练数据, 剩余的图像作为测试数据, 重复选取 10 次. 由表 6 中的结果可知: (1) 本文方法的识别率最高,

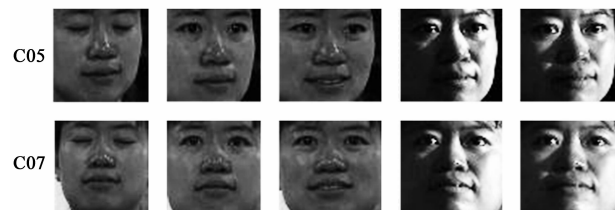


图 5 CMU PIE数据库的示意图

训练数目为 2 时,高于 MSPCRC 方法 0.92%;训练数目为 3、4、5 时,高于 FDDL 方法 2.38%、0.73%、0.09%,说明本文方法可以很好的解决小样本问题;(2)本文方法的识别率高于 BoW 和 GBoW,说明多核学习能够有效地将互补的局部和全局特征进行结合.

#### 4.4 LFW 数据库上的实验结果

LFW 数据库包含 5279 人的 13233 张图片, LFW-a 是利用商业校准软件进行校准之后的版本,与文献 [13] 一致,本文使用其中样本数目不少于 10 幅的 158 个人的图像进行实验,图像的尺寸为  $32 \times 32$  像素,示例图像见图 6. 在此数据库上验证不同训练样本数目下人脸识别算法的性能,分别随机选取每个人的 2、3、4、5 幅图像作为训练数据,剩余的图像作为测试数据,重复选取 10 次. 由表 7 中的结果可知:(1)本文方法的识别率最高,分别高于 MSPCRC 方法 0.40%、0.58%、3.46%、5.70%,说明本文方法可以很好的解决小样本问题;(2)本文方法的识别率高于 BoW 和 GBoW,说明多核学习能够有效地将互补的局部和全局特征进行结合.

表 7 LFW 数据库上不同训练样本数目下各种方法的识别率  $\pm$  标准差 (%)

人脸识别算法	2	3	4	5
PCA	9.35 $\pm$ 0.75	11.53 $\pm$ 0.69	13.36 $\pm$ 0.71	14.84 $\pm$ 0.80
FLDA	19.31 $\pm$ 0.88	28.69 $\pm$ 1.05	35.37 $\pm$ 0.96	42.51 $\pm$ 1.36
CRC	22.06 $\pm$ 0.72	28.16 $\pm$ 1.32	33.53 $\pm$ 1.52	38.23 $\pm$ 1.51
FDDL	25.69 $\pm$ 1.48	32.55 $\pm$ 2.02	38.08 $\pm$ 1.52	39.98 $\pm$ 2.04
SLF-RKR	25.98 $\pm$ 1.08	32.30 $\pm$ 1.60	36.37 $\pm$ 0.79	39.48 $\pm$ 1.14
Volterra	26.03 $\pm$ 3.04	32.01 $\pm$ 3.47	36.39 $\pm$ 3.38	40.25 $\pm$ 2.71
MSPCRC	34.95 $\pm$ 2.15	41.54 $\pm$ 1.75	45.85 $\pm$ 1.43	49.40 $\pm$ 1.95
BoW	27.37 $\pm$ 1.11	35.03 $\pm$ 1.15	40.88 $\pm$ 1.70	45.34 $\pm$ 1.60
GBoW	25.66 $\pm$ 1.29	32.69 $\pm$ 1.13	38.49 $\pm$ 1.73	45.96 $\pm$ 0.98
本文方法	<b>35.35 <math>\pm</math> 1.17</b>	<b>42.12 <math>\pm</math> 1.13</b>	<b>47.31 <math>\pm</math> 1.43</b>	<b>52.10 <math>\pm</math> 1.51</b>



图6 LFW数据库的示意图

## 5 总结

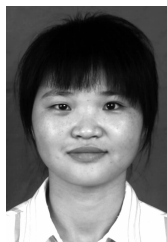
本文提出将词袋模型作为人脸图像的局部特征描述,同时提出词袋模型的全局模式,并将其作为人脸图像的全局特征描述,最后使用多核学习方法将二者进行融合. 公开人脸数据库上的实验结果表明:(1)本文方法能够很好地解决小样本问题;(2)多核学习方法能够有效地将互补的局部和全局特征进行结合;(3)本文方法对人脸的表情变化、姿态变化以及遮挡具有良好的鲁棒性.

## 参考文献

- [1] Lai Z H, Xu Y, Chen Q C, Yang J. Multilinear sparse principal component analysis[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25 (10): 1942 - 1950.
- [2] Wang H X, Lu X S, Zheng W N. Fisher discriminant analysis with  $L_1$ -norm[J]. IEEE Transactions on Cybernetics, 2013, 44(6): 828 - 842.
- [3] Wright J, Yang A Y, Ganesh A, Satri S S, Ma Y. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(2): 210 - 226.
- [4] Zhang L, Yang M, Feng X C. Sparse representation or collaborative representation: which helps face recognition [A]? Proceedings of the 13th IEEE International Conference on Computer Vision [C]. Barcelona, Spain: IEEE Computer Society, 2011. 471 - 478.
- [5] Yang M, Zhang L, Zhang D. Metaface: learning for sparse representation based face recognition [A]. Proceedings of the 17th International Conference on Image Processing [C]. Hong Kong, China: IEEE Computer Society, 2010. 1601 - 1604.
- [6] Yang M, Zhang L, Feng X C, Zhang D. Sparse representation based Fisher discrimination dictionary learning for image classification[J]. International Journal of Computer Vision, 2014, 109(3): 209 - 232.
- [7] Liu C, Wechsler H. Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition [J]. IEEE Transactions on Image Processing, 2002, 11(4): 467 - 476.
- [8] 张文超, 山世光, 张洪明, 陈杰, 陈熙林, 高文. 基于局部 Gabor 变化直方图序列的人脸描述与识别[J]. 软件学报, 2006, 17(12): 2508 - 2517.  
Zhang Wenchao, Shan Shiguang, Zhang Hongming, Chen Jie, Chen Xilin, Gao Wen. Histogram sequence of local Gabor binary pattern for face description and identification [J]. Journal of Software, 2006, 17(12): 2508 - 2517. (in Chinese)
- [9] 李宽, 殷建平, 李永, 刘发耀. 基于 Gabor 系数分块统计和自适应特征选择的人脸描述与识别[J]. 计算机研究与发展, 2012, 49(4): 777 - 784.  
Li Kuan, Yin Jianping, Li Yong, Liu Fayao. Local statistical analysis of Gabor coefficients and adaptive feature extraction for face description and recognition [J]. Journal of Computer Research and Development, 2012, 49(4): 777 - 784. (in Chinese)
- [10] Yang M, Zhang L, Shiu S, Zhang D. Robust kernel representation with statistical local features for face recognition

- [J]. IEEE Transactions on Neural Networks and Learning Systems, 2013, 24(6): 900–912.
- [11] Tan K R, Chen S C. Adaptively weighted subpattern PCA for face recognition [J]. Neurocomputing, 2004, 64: 505–511.
- [12] Kumar R, Banerjee A, Vemuri B C. Volterrafaces: discriminant analysis using volterra kernels [A]. Proceedings of the 22nd Conference on Computer Vision and Pattern Recognition [C]. Miami, USA: IEEE Computer Society, 2009. 150–155.
- [13] Zhu P F, Zhang L, Hu Q H, Shiu S. Multi-scale patch based collaborative representation for face recognition with margin distribution optimization [A]. Proceedings of the 12th European Conference on Computer Vision [C]. Florence, Italy: Springer-Verlag, 2012. 822–835.
- [14] 朱玉莲, 陈松灿. 特征采样和特征融合的子图像人脸识别方法 [J]. 软件学报, 2012, 23(12): 3209–3220.  
Zhu Yulian, Chen Songcan. Sub-image method based on feature sampling and feature fusion for face recognition [J]. Journal of Software, 2012, 23(12): 3209–3220. (in Chinese)
- [15] Cai D, He X F, Han J W, Zhang H J. Orthogonal Laplacian faces for face recognition [J]. IEEE Transactions on Image Processing, 2006, 15(11): 3608–3614.

#### 作者简介



杨赛女, 1981年出生, 山东滨州人。2015年获南京理工大学博士学位, 现为南通大学讲师, 研究方向为计算机视觉与机器学习。

E-mail: yangsai166@126.com



赵春霞女, 1964年出生, 北京人。1985年、1988年和1998年在哈尔滨工业大学分别获得工学学士、工学硕士和工学博士学位, 现为南京理工大学教授, 博士生导师。主要研究方向为地面智能机器人与复杂环境理解。