



稀疏综合字典学习的小样本人脸识别

狄岚, 矫慧文, 梁久祯

引用本文:

狄岚, 矫慧文, 梁久祯. 稀疏综合字典学习的小样本人脸识别[J]. 智能系统学报, 2021, 16(2): 218–227.

DI Lan, JIAO Huiwen, LIANG Jiuzhen. Sparse comprehensive dictionary learning for small-sample face recognition[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(2): 218–227.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.201910028>

您可能感兴趣的其他文章

基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects

智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

一种恢复图像自然色彩的重构方法

Reconstruction method for restoring the natural color of an image

智能系统学报. 2019, 14(5): 877–881 <https://dx.doi.org/10.11992/tis.201805011>

结构化加权稀疏低秩恢复算法在人脸识别中的应用

A low rank recovery algorithm for face recognition with structured and weighted sparse constraint

智能系统学报. 2019, 14(3): 455–463 <https://dx.doi.org/10.11992/tis.201711026>

基于正交Log-Gabor滤波二值模式的人脸识别算法

Face recognition based on orthogonal Log-Gabor binary pattern

智能系统学报. 2019, 14(2): 330–337 <https://dx.doi.org/10.11992/tis.201708015>

图正则化字典对学习的轻度认知功能障碍预测

Dictionary pair learning with graph regularization for mild cognitive impairment prediction

智能系统学报. 2019, 14(2): 369–377 <https://dx.doi.org/10.11992/tis.201709033>

鲁棒的正则化编码随机遮挡表情识别

Recognition of facial expression in case of random shielding based on ro-bust regularized coding

智能系统学报. 2018, 13(2): 261–268 <https://dx.doi.org/10.11992/tis.201609002>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.201910028

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20200728.1555.008.html>

稀疏综合字典学习的小样本人脸识别

狄岚^{1,2}, 矫慧文¹, 梁久祯³

(1. 江南大学 人工智能与计算机学院, 江苏 无锡 214122; 2. 道路交通安全公安部重点实验室, 江苏 无锡 214151; 3. 常州大学 信息科学与工程学院, 江苏 常州 213164)

摘 要: 传统以字典学习为基础的小样本人脸识别方法存在字典低辨别性、弱鲁棒性等缺点, 对此, 本文提出稀疏综合字典学习模型。该模型有效利用和生成人脸变化, 以镜像原理及 Fisher 准则扩充训练样本多样性, 通过构造混合特色字典、扩充干扰字典以及低秩字典原子, 提取不同类别数据之间的共性、特殊性和异常情况, 从而提高算法识别率以及对表情变化、姿态变化、遮挡等异常情况的处理能力。在 AR、YALEB、LFW 等人脸数据库进行仿真实验, 实验结果验证了算法的有效性和可行性。

关键词: 综合字典学习; 人脸识别; 类别特色字典; Fisher 准则; 小样本; 图像扩充; 镜像准则; 扩充干扰字典; 混合特色字典; 低秩字典

中图分类号: TP394 文献标志码: A 文章编号: 1673-4785(2021)02-0218-10

中文引用格式: 狄岚, 矫慧文, 梁久祯. 稀疏综合字典学习的小样本人脸识别 [J]. 智能系统学报, 2021, 16(2): 218-227.

英文引用格式: DI Lan, JIAO Huiwen, LIANG Jiuzhen. Sparse comprehensive dictionary learning for small-sample face recognition[J]. CAAI transactions on intelligent systems, 2021, 16(2): 218-227.

Sparse comprehensive dictionary learning for small-sample face recognition

DI Lan^{1,2}, JIAO Huiwen¹, LIANG Jiuzhen³

(1. School of Artificial Intelligence and Computer Science, Jiangnan University, Wuxi 214122, China; 2. Laboratory of Ministry of Public Security for Road Traffic Safety, Wuxi 214151, China; 3. School of Information Science and Engineering, Changzhou University, Changzhou 213164, China)

Abstract: Traditional small-sample face recognition methods based on dictionary learning have disadvantages such as poor dictionary discrimination and lack of robustness. In this paper, we propose a sparse comprehensive dictionary learning model. This model effectively utilizes and generates facial changes, expands the diversity of training samples by the mirror principle and Fisher's criterion, and extracts the commonalities, specialties, and anomalies between different categories of data by constructing a hybrid feature dictionary, extended interference dictionary, and low-rank dictionary atoms. This strategy improves the recognition rate of the algorithm and its ability to handle abnormal situations such as expression changes, pose changes, and occlusions. The results of simulation experiments performed on the face databases AR, YALEB, and LFW verify the effectiveness and feasibility of the proposed algorithm.

Keywords: comprehensive dictionary learning; face recognition; class-specific dictionary learning; Fisher discrimination criterion; small sample; image expansion; mirror principle; extended interference dictionary; hybrid feature dictionary; low-rank dictionary

收稿日期: 2019-10-23. 网络出版日期: 2020-07-28.

基金项目: 江苏省研究生科研与实践创新计划项目 (KYCX19_1895); 道路交通安全公安部重点实验室开放课题 (2020ZDSYSKFKT03-2, A 类).

通信作者: 狄岚. E-mail: dilan@jiangnan.edu.cn.

近年来, 深度学习进入蓬勃发展时代, 以深度学习为基础的图像识别^[1-3]虽然识别准确率高, 但往往对硬件设备要求严格, 具有训练时间长达数周、样本量需求过大等不足之处。与之相比, 基

于稀疏表示的图像识别训练简单、对噪声有强鲁棒性,逐渐应用在社会安全、经济工程等领域。

字典学习的目的是通过训练样本图片,习得可以表示给定信号的字典和编码,共分为有监督和无监督两种情况。经典的无监督字典学习如KSVD^[4]算法,可以有效解决高维矩阵求解问题,广泛应用于图像重建及图像压缩领域。而有监督字典学习多应用于图像识别领域,如人脸识别、表情识别等。根据字典原子和训练标签之间的相关性,有监督字典学习可分为类共享字典学习、类别特色字典学习、混合字典学习3类。

类共享字典学习如文献[5-7]等,字典和稀疏编码由所有训练样本共同构造,单个字典原子可表示全体数据类别。Zhang等^[8]以KSVD算法为基础,在保证字典识别能力的基础上引入分类误差,提高线性分类器的分类能力。Jiang等^[9]在其基础上,增加标签一致性术语,提出标签一致性KSVD(LCKSVD)算法。Mairal等^[10]提出了任务驱动的字典学习框架。类共享字典在保留表示数据共性的同时丢失字典原子和类标签之间的关系,忽视了类别之间的差异性。

类别特色字典学习如文献[11-13]等,每个字典原子对应一类标签信息,通过类特定误差进行分类。Wright等^[14]提出基于稀疏表示的分类框架(sparse representation based classification, SRC),自此,类别特色字典广泛应用于人脸分类器设计中。Sprechmann等^[15]为每一类分别学习稀疏表示字典,Ramirez等^[16]提出了一种结构非相干字典学习模型,通过最小化字典的相干项来提高字典的学习效率。Yang等^[17]提出了Fisher判别字典学习(fisher discrimination dictionary learning, FDDL),其中表示残差和表示系数都实现了判别信息。2018年,Wang等^[18]结合字典学习和局部约束思想,提出FDDLLCSRC方法。同年,Li等^[19]在FDDL基础上提出IKCFDDL算法,在字典学习基础上增添K均值聚类思想,提升算法运行效率。

混合字典即结合类共享字典与类别特色字典。Deng等^[20]提出了基于扩展稀疏表示的分类方法(extended SRC: undersampled face recognition via intraclass variant dictionary, ESRC),在学习类别特色字典的同时构造类内差异字典,即共享字典。受ESRC算法启发,2015年Li等^[21]提出SCSDL算法,2017年提出CSICVDL算法^[22],提取不同类别样本中数据共性提升字典辨别力。这些混

合字典学习方法在人脸识别方面表现良好,然而,在对于噪声、异常值和遮挡等干扰情况时处理能力较差。

除上述问题以外,由于样本数量受限,传统字典学习习得的字典鲁棒性差,不能完全表示数据组成的共性、特殊性、干扰性。针对此种情况,本文提出一种新的稀疏综合字典学习的小样本人脸识别模型(SCDL)。SCDL模型包括混合特色字典、扩充干扰字典以及低秩字典3项。将类共享字典与类别特色字典以及Fisher准则整合至混合特色字典模型中,在提取不同数据特殊性的同时捕捉数据之间的共性,利用扩充干扰字典和低秩字典增强模型对异常情况(如遮挡、噪声等)的处理能力。最后,本文针对模型提出一种新的分类策略,并在受限AR人脸数据库、YaleB人脸数据库和非受限LFW人脸数据库进行实验,结果表明本文算法不仅可以取得较高分类率,并且在面对遮挡、噪声等异常情况时较其余算法有更好的识别力和鲁棒性。

1 相关工作

设定训练数据集 \mathbf{A} :

$$\mathbf{A} = \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_K\} \in \mathbf{R}^{m \times N}$$

测试数据集合 \mathbf{Y} :

$$\mathbf{Y} = \{\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_K\} \in \mathbf{R}^{m \times N_1}$$

式中: \mathbf{A} 的每列表示一个 m 维向量; $\mathbf{A}_i \in \mathbf{R}^{m \times n_i}$ 为第 i 类训练数据; $N = \sum_{i=1}^K n_i$ 为训练数据集样本总数; y 表示单个测试样本。

1.1 扩展稀疏表示

稀疏表示(SRC)针对测试集 \mathbf{Y} , 以学习合适的字典 \mathbf{D} 为手段, 寻找能表达 \mathbf{Y} 的稀疏编码 \mathbf{X} , ESRC算法在此基础上另学习一个类内差异字典表示训练集 \mathbf{A} 和测试集 \mathbf{Y} 之间的变化, 算法模型如下:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \arg \min_{a,b} \left\{ \left\| \mathbf{y} - [\mathbf{A}, \mathbf{D}] \begin{bmatrix} a \\ b \end{bmatrix} \right\|_2^2 + \lambda \left\| \begin{bmatrix} a \\ b \end{bmatrix} \right\|_1 \right\}$$

ESRC模型分为重构误差项与正则项两部分, 正则化参数 $\lambda > 0$ 。稀疏编码 a, b 的快速求解可参考贪婪算法(如压缩感知匹配追踪^[23]、稀疏自适应匹配追踪^[24])或凸松弛算法(如最小绝对值收缩选择算法^[25]、最小角度回归算法^[26])。

1.2 类别特色字典学习

类别字典学习为每类数据单独学习一个子字典, 即习得的字典 $\mathbf{D} = \{\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_K\}$ 由 K 个子字典组成, 算法模型如下:

$$\min_{\{D_i, X_i\}} \left\{ \|A_i - D_i X_i\|_F^2 + \lambda \|X_i\|_1 \right\} \text{ s.t. } \|d_j\|_2 = 1, \forall j$$

对于单个测试样本 y , 求得的稀疏系数为 $x = \{x_1, x_2, \dots, x_K\}$ 。其中, D_i 表示第 i 类训练样本的子字典, x_i 由 D_i 重构得到, 重构误差为 $e_i = \|y - D_i x_i\|_2$ 。

1.3 Fisher 判别准则

Fisher 判别准则以投影思想为核心, 目的是使同类样本尽可能集中的同时, 异类样本尽可能疏散。

假定 u_0 为稀疏编码 X 的中心, u_i 表示各类稀疏编码均值向量:

$$u_i = \frac{1}{n_i} \sum_{x_k \in X_i} x_k$$

可定义类内散度矩阵 $S_w(X)$ 、类间散度矩阵 $S_B(X)$:

$$S_w = \sum_{i=1}^K \sum_{x_k \in X_i} (x_k - u_i)(x_k - u_i)^T$$

$$S_B = \sum_{i=1}^K n_i (u_i - u_0)(u_i - u_0)^T$$

2 稀疏综合字典学习的小样本人脸识别算法

稀疏综合字典学习分类方法包括训练和测试两大步骤, 分为扩充干扰字典、混合特色字典、低秩字典三大部分, 采用分步优化方法求解, 充分提取数据之间的共性、特殊性、干扰性。

2.1 扩充干扰特色字典

2.1.1 扩充干扰字典模型

对于一张图片 a , 以图片中线为轴, 如图 1 所示, 分别得出左镜像图片 a_1 、右镜像图片 a_2 、全镜像图片 a_3 。

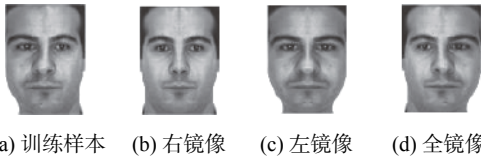


图 1 镜像图片
Fig. 1 Mirror image

以 a 为例, 对于训练集 A , 求出左镜像集 A_1 , 右镜像集 A_2 , 全镜像集 A_3 , 以 Fisher 准则为基础, 分别计算出 3 个集合的类间散布系数与类内散布系数之比作为权重 $w = [w_1, w_2, w_3]$ 。

集合 $A' = w_1 A_1 + w_2 A_2 + w_3 A_3$ 。以 A' 为训练集, 提出扩充干扰字典模型如下:

$$J_{(D_b, B)} = \arg \min_{(D_b, B)} \{g(A', D_b, B) + \lambda_1 \|B\|_1 + \lambda_2 f(B, D_b)\} \quad (1)$$

模型分为重构误差项、稀疏保证项和判别系

数项 3 部分, 其中, $\lambda_1 > 0$ 为正则项参数, $\lambda_2 > 0$ 为判别系数项参数。下面依次论述模型每一项原理。

$g(A', D_b, B)$ 表示重构误差项, 该项保证字典的识别力, 公式表示为

$$g(A', D_b, B) = \|A'_i - D_b B_i\|_F^2 + \|A'_i - D_{bi} B_i^i\|_F^2$$

式中: B_i 是 A'_i 被 D_b 重构的稀疏编码; B_i^i 是 A'_i 被 D_{bi} 重构的稀疏编码; $\|A'_i - D_b B_i\|_F^2$ 表示整体字典针对 A'_i 的重构误差; $\|A'_i - D_{bi} B_i^i\|_F^2$ 表示该类字典针对 A'_i 单独重构误差。

为了提高编码识别力, 增添编码约束项 $\text{tr}(S_w(B) - S_B(B))$, 将 Fisher 准则整合到编码约束项, 扩大类间距, 减小类内距, 进而增强辨别性。为了确保子字典不连贯, 提高空间多样性, 增添字典约束项 $\|D_{bi}^T D_{bi-}\|_F^2$ 。其中,

$$D_{bi-} = [D_{b1} \ D_{b2} \ \dots \ D_{bi-1} \ D_{bi+1} \ \dots \ D_{bK}]$$

表示字典中除了 D_{bi} 项之外剩余集合。

最终, $f(B, D_b)$ 公式如下:

$$f(B, D_b) = \text{tr}(S_w(B) - S_B(B)) + \|D_{bi}^T D_{bi-}\|_F^2$$

式中: $\|B\|_1$ 表示稀疏保证项, 确保编码的稀疏项。

目标函数 (1) 的完整形式为

$$J_{(D_b, B)} = \arg \min_{(D_b, B)} \left\{ \|A'_i - D_b B_i\|_F^2 + \|A'_i - D_{bi} B_i^i\|_F^2 + \lambda_1 \|B\|_1 + \lambda_2 (\text{tr}(S_w(B) - S_B(B)) + \|D_{bi}^T D_{bi-}\|_F^2) \right\}$$

2.1.2 扩充干扰字典优化

式 (1) 为非凸函数, 其优化过程如下所示:

1) 初始化字典 D_b 。

将字典原子初始化为训练数据 $A' = \{A'_1, A'_2, \dots, A'_K\}$ 的特征向量, 对字典 D_b 的每一类归一化, 使其 l_2 范数为 1。

2) 保证字典 D_b 固定, 更新稀疏编码 B 。

目标函数转化为

$$Q(B_i) = \arg \min_{B_i} \{\theta(A'_i, D_{bi}, B_i) + 2\gamma \|B_i\|_1\} \quad (2)$$

使用文献 [27] 中的方法求解稀疏编码 $B_i \in \mathbf{R}^{N \times n_i}$, $\gamma = \lambda_1/2$ 。

$$\theta(A'_i, D_{bi}, B_i) = \|A'_i - D_b B_i\|_F^2 + \|A'_i - D_{bi} B_i^i\|_F^2 + \lambda_2 f(B, D_b)$$

3) 保证稀疏编码 B 固定, 更新字典 D_b 。

本文逐个更新 D_b 中的子字典, 即当更新第 j 个子字典时, 保证其他子字典 $D_j (i \neq j)$ 不动, 默认更新完成。

目标函数可转化为

$$D_{bi} = \arg \min_{D_{bi}} \left\{ \|\hat{A} - D_{bi} X_i\|_F^2 + \|A'_i - D_{bi} B_i^i\|_F^2 + \|D_{bi}^T D_{bi-}\|_F^2 \right. \\ \left. \text{s.t. } \|D_{bi}(:, j)\|_2 = 1, \forall i, j \right.$$

其中 $\hat{A} = A' - \sum_{j=1, j \neq i}^k D_{bj} B^j$, B^j 是所有数据 A' 由字典 D_{bj} 表示的稀疏编码, 因此 $Q(D_{bi})$ 可转换为

$$D_{bi} = \arg \min_{D_{bi}} \|A_i - D_{bi} Z_i\|_F^2 + \|D_{bi}^T D_{bi-}\|_F^2$$

$$\text{s.t. } \|D_{bi}(:, j)\|_2 = 1, \forall i, j$$

其中 $A_i = [\hat{A}, A'_i]$, $Z_i = [B_i, B'_i]$, 同理, 逐个更新 $D_{bi} = [d_{bi}^1, d_{bi}^2, \dots, d_{bi}^K]$ 中的原子, 当更新 d_{bi}^k 时, 保证其他原子不动, 默认已更新完成。

设 $Z_i = [z_{(1)} \ z_{(2)} \ \dots \ z_{(m)}]$, 其中 $z_{(k)}$ 为 Z_i 的第 k 行, 令 $\hat{A}_i = A_i - \sum_{j \neq k} d_{bi}^j z_{(j)}$, 于是:

$$d_{bi}^k = \arg \min_{d_{bi}^k} \left\{ g(d_{bi}^k) = \|\hat{A}_i - d_{bi}^k z_{(k)}\|_F^2 + \|d_{bi}^{kT} D_{bi-}\|_F^2 \right\} \quad (3)$$

令 $\partial \frac{g(d_{bi}^k)}{d_{bi}^k} = 0$, 得到:

$$d_{bi}^k = \left((D_{bi-})(D_{bi-})^T + \|z_{(k)}\|_2^2 I \right)^{-1} \hat{A}_i z_{(k)}^T$$

将得到的 d_{bi}^k 规范化: $d_{bi}^k = d_{bi}^k / \|d_{bi}^k\|_2$

4) 重复 2) 和 3), 直到前后两次的函数 Q 的值满足判断条件为止。

扩充干扰字典算法总体实现步骤如下:

输入 训练样本 A' , 规范化参数 γ ;

输出 字典 D_b 和稀疏编码 B 及相应的标签。

1) 初始化字典 D_b 。

2) 固定字典 D_b , 更新稀疏编码 B 。初始化字典后, 利用式 (2) 依次求解。

3) 固定稀疏编码 B , 更新字典 D_b 。利用式 (3) 依次更新。

4) 重复 2) 和 3), 直到前后两次的函数的值满足判断条件为止。

2.2 混合特色字典

2.2.1 混合特色字典模型

对于训练集 A 和测试集 Y , 习得类共享字典 D_c 和由 K 个子字典组成的类别特色字典 D : $D = \{D_1, D_2, \dots, D_K\}$ 。根据 Fisher 判别准则, 混合特色字典模型如式 (4):

$$J(D_c, D, C, X) = \arg \min_{D_c, D, C, X} \left\{ \|A_i - D_c C_i - D_i X_i\|_F^2 + \lambda_1 (\|C\|_1 + \|X\|_1) \right\} + \lambda_2 (\text{tr}(S_w(X)) - S_B(X)) \quad (4)$$

$\|A_i - D_c C_i - D_i X_i\|_F^2$ 可分为 $\|A_i - D_c C_i\|_F^2$ 和 $\|A_i - D_i X_i\|_F^2$ 两项理解, 其中, C_i 是 A_i 被 D_c 重构的稀疏编码, X_i^i 是 A_i 被 D_i 重构的稀疏编码。

$\|A_i - D_c C_i\|_F^2$ 提取数据之间的共性, $\|A_i - D_i X_i\|_F^2$ 根据类标签信息, 提取数据之间的特殊性, 确保每个样本可以由该类字典单独重构, 以此提高子字典的判别力。其中, $\lambda_1 > 0$ 为正则项参数, $\lambda_2 > 0$ 稀疏编码参数。

2.2.2 混合特色字典优化

式 (4) 为非凸函数, 其优化过程如下所示:

1) 初始化字典 D 和 D_c 。

将训练数据 $A = \{A_1, A_2, \dots, A_K\}$ 的特征向量初始化为字典的原子, 对字典 D_c 归一化, 使其 l_2 范数为 1; 分别对字典 D 的每一类归一化, 使其 l_2 范数为 1。

2) 保证字典 D 固定, 更新稀疏编码 X ; 保证字典 D_c 固定, 更新稀疏编码 C 。

更新 X 时, 目标函数转化为

$$Q(X_i^i) = \arg \min_{X_i^i} \left\{ \theta_1(A_i, D_i, X_i) + 2\gamma \|X_i^i\|_1 \right\} \quad (5)$$

其中,

$$\theta_1(X_i^i) = \|A_i - D_c C_i - D_i X_i^i\|_F^2 + \lambda_2 \text{tr}(S_w(X)) - S_B(X)$$

更新 C 时, 目标函数转化为

$$Q(C_i) = \arg \min_{C_i} \left\{ \|A_i - D_c C_i - D_i X_i^i\|_F^2 + 2\gamma \|C_i\|_1 \right\} \quad (6)$$

3) 以章节 2.1.2 为例, 固定稀疏编码 X , 更新字典 D ; 固定稀疏编码 C , 更新字典 D_c 。

4) 重复 2) 和 3), 直到前后两次的函数 Q 的值满足判断条件为止。

混合特色字典算法总体实现步骤如下:

输入 训练样本 A , 规范化参数 γ ;

输出 字典 D 和稀疏编码 X , 字典 D_c 和稀疏编码 C 及相应的标签。

1) 初始化字典 D 和 D_c 。

2) 固定字典 D , 更新稀疏编码 X ; 固定字典 D_c , 更新稀疏编码 C ; 初始化字典后, 利用式 (5)、(6) 依次求解。

3) 固定稀疏编码 X , 更新字典 D ; 固定稀疏编码 C , 更新字典 D_c 。

4) 重复 2) 和 3), 直到前后两次的函数的值满足判断条件为止。

2.3 低秩字典

2.3.1 低秩字典模型

在人脸识别中, 假设扰动分量只占图像特征的一小部分, 即图像的稀疏分量, 使用矩阵低秩分解^[28]方法从图像中提取干扰成分(如噪声、孤立点和遮挡)。

$$\min_{\Gamma_j, E_j} \text{rank}(\Gamma_j) + r \|E_j\|_0 \text{ s.t. } G_j = \Gamma_j + E_j$$

以如图 2 所示, 取图片低秩分量 E_j 作为训练集。其中, $\lambda_1 > 0$ 为正则项参数, 提出低秩字典模型如下:

$$J(D_p, P) = \arg \min \left\{ \|E - D_p P_i\| + \lambda_1 \|P_i\|_1 \right\} \quad (7)$$

2.3.2 低秩字典优化

式 (7) 为非凸函数, 其优化过程如下所示:

1) 初始化字典 D_p 。

将字典原子初始化为训练数据 $E = \{E_1, E_2, \dots, E_H\}$ 的特征向量, 对字典 D_p 的每一类归一化, 使其 l_2 范数为 1。

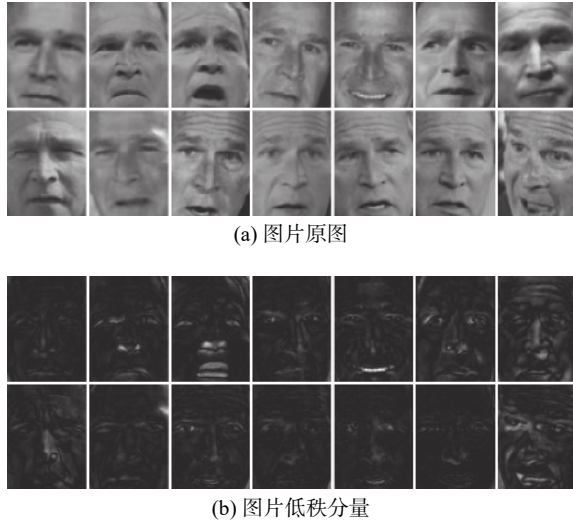


图 2 原图及低秩分量

Fig. 2 Original graph and low rank components

2) 保证字典 D_p 固定, 更新稀疏编码 P ; 更新 X 时, 目标函数转化为

$$Q(P_i) = \arg \min_{P_i} \{ \|E_i - D_p P_i\|_F^2 + 2\gamma \|P_i\|_1 \}$$

3) 以章节 2.1.2 为例, 固定稀疏编码 P , 更新字典 D_p 。

4) 重复 2) 和 3), 直到前后两次的函数的值满足判断条件为止。

低秩字典优化算法总体实现步骤如下:

输入 训练样本 E , 规范化参数 γ ;

输出 字典和稀疏系数 P , 及相应的标签。

1) 初始化字典 D_p ;

2) 固定字典 D_p , 更新稀疏编码 P ;

3) 固定稀疏编码 P , 更新字典 D_p ;

4) 重复 2) 和 3), 直到前后两次的函数的值满足判断条件为止。

2.4 分类策略

本文提出全局分类策略如下:

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} = \arg \min_{x_1, x_2, x_3, x_4} \left\{ \left\| y - [D_c, D, D_b, D_p] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \right\|_2^2 + \lambda_1 \left\| \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \right\|_1 \right\}$$

测试样本 y 的重构误差为

$$e_i = \left\| y - [D_c, D, D_b, D_p] \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} \right\|_2^2 \quad (8)$$

于是 y 的标签为

$$\text{identify}(y) = \arg \min_i \{e_i\} \quad (9)$$

2.5 本文整体算法步骤及流程

稀疏综合字典学习的小样本人脸识别算法分为 3 个子算法, 利用训练数据 A 、辅助数据 C 构造扩充干扰字典、混合特色字典、低秩字典。算法

流程图如图 3 所示。

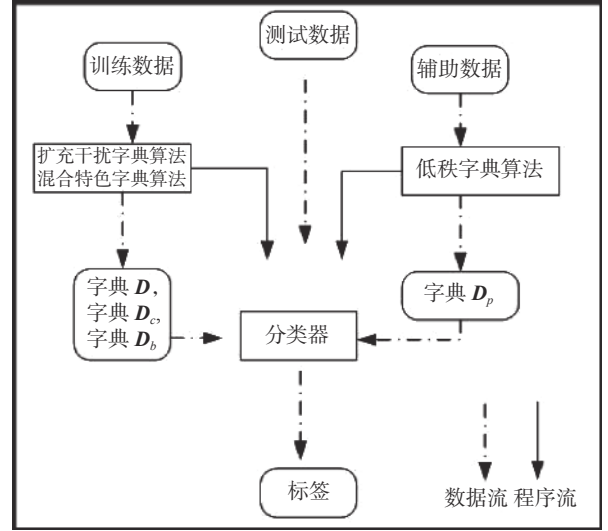


图 3 算法流程

Fig. 3 Algorithm flowchart

稀疏综合字典学习的小样本人脸识别算法总体实现步骤如下:

输入 训练样本 A , 辅助数据 C , 测试数据 Y , 参数 λ_1, λ_2 。

输出 分类标签。

1) 利用训练数据 A 构造数据集 A' , 算法 1 习得扩充干扰字典 D_b ;

2) 利用训练数据 A 及算法 2 习得混合特色字典 D, D_c ;

3) 利用辅助数据 C 及算法 3 习得混合特色字典 D_p ;

4) 利用式 (8)、(9) 得到样本标签。

3 实验结果及分析

3.1 实验平台和参数设置

本文实验环境为 64 位 Windows 10 操作系统, 内存 32 GB, Intel(R) Xeon(R) CPU E5-2 620 v4 @ 2.10 GHz, 并用 MatlabR2016b 软件编程实现。

实验选取 AR 人脸数据库、YaleB 人脸数据库、LFW 人脸数据库进行实验, 多次实验取平均值。图像都经过标准化处理, 比较算法包括 SRC、FDDL、CRC、ESRC、SVGDL 和 CSICVDL、SCSDL、FDDLCSRC、LKCFDDL。

3.2 AR 数据库实验

本文在 AR 人脸数据库上设计两个实验方案, 实验 1 随机选取 100 人, 每人 26 张图片分为 5 个集合, 训练集合选取两张标准人脸, 其余按特点分为 4 个集合, 作为不同的测试集。如图 4 所示, 集合 S_1 为样本内所有表情变化图片; 集合 S_2

为所有光照变化图片;集合 S_3 为所有眼镜遮挡图片;集合 S_4 为所有围巾遮挡图片。

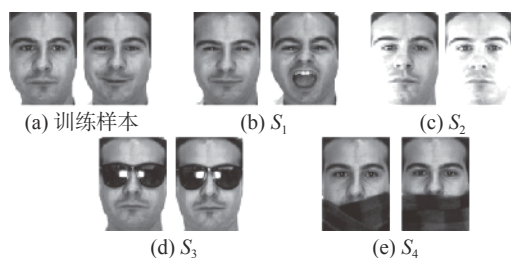


图4 AR人脸数据库样本(1)

Fig. 4 Cropped face samples of AR database (1)

在具体实现过程中,随机选取80人用于训练扩充干扰字典和混合特色字典,其余20个人用于训练低秩字典。首先将数据集下采样为 60×80 , 并采用PCA降至100维。各算法在AR数据库的识别率如表1所示。

表1 算法在AR库上的实验结果

Table 1 Accuracy of different methods on Experiment1 of AR database

算法	S_1	S_2	S_3	S_4
CRC	78.96	91.86	40.21	35.21
SRC	83.13	61.46	42.08	33.54
FDDL	84.79	87.70	48.75	35.83
ESRC	86.04	89.17	48.95	40.56
SVGDL	82.90	92.50	51.80	45.80
CSSVDL	86.25	99.38	79.16	55.39
SCSDL	84.46	99.38	77.34	54.29
FDDLCSRC	86.25	99.38	69.75	51.75
IKCFDDL	86.17	98.88	74.08	53.21
本文算法	86.46	99.38	84.38	67.29

从表1可知,FDDL算法识别率高于SRC、CRC算法,说明提取数据特殊性的的重要性,FDDLCSRC、IKCFDDL算法在面对光照、表情变化时实验效果良好,然而对数据存在遮挡异常时处理效果欠缺。而CSICVDL、SCSDL、本文算法识别率高于FDDL、LKCFDDL等,说明了在提取特殊性之外,捕捉数据共性的必要性。本文算法性能较稳定,混合特色字典提取了数据共性和特殊性,低秩字典、扩充干扰字典增强算法鲁棒性和容错能力,在所有集合都能达到最高识别率。

为探究本文算法对遮挡、光照、异常等情况的综合处理能力,如图5所示,实验2选取每人两张正常状态下的人脸图片作为训练集,将围巾遮挡、墨镜遮挡作为测试集合。



图5 AR人脸数据库样本(2)

Fig. 5 Cropped face samples of AR database(2)

将数据集下采样为 60×80 , 采用PCA将数据降为 {50,150,250,350,450} 维。如图6所示,本文算法在不同维度下皆取得最高识别率,在面对遮挡等异常情况时表现最佳。

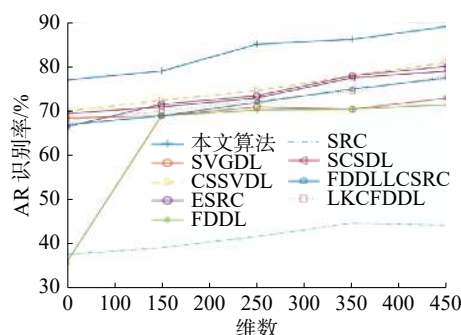


图6 AR人脸数据库识别率

Fig. 6 Accuracy of different methods on Experiment1 of AR database

3.3 YaleB 数据库实验

本文在 The extended Yale B 人脸数据库上设计两个实验。实验1中,每人随机选取5张图片训练,其余图片进行测试。图7表示其中一个训练样本和部分测试样本。



图7 YaleB人脸数据库样本

Fig. 7 Cropped face samples of YaleB database

在具体实验时,选取30个人训练扩充干扰字典和混合特色字典,其余8个人用于训练低秩字典。将数据分别降维至 {150,250,350,450,550} 维,各算法在各维度上的识别率如表2所示。

表2 算法在 Yale B database 库上的实验结果
Table 2 Accuracy of different methods on Experiment1 of Yale B database

算法	150维	250维	350维	450维	550维
SRC	52.05	57.99	59.04	58.82	58.57
CRC	51.94	57.10	59.66	61.82	62.26
FDDL	57.88	59.32	59.82	59.82	59.76
SVGDL	59.10	59.66	62.04	67.41	69.63
CSSVDL	60.63	61.59	65.82	69.73	70.57
ESRC	60.10	59.04	62.04	66.94	67.83
SCSDL	60.88	62.10	64.57	68.85	70.57
FDDLCSRC	59.70	60.66	62.58	64.10	67.89
LKCFDDL	64.46	67.10	69.38	70.53	71.46
本文算法	65.65	69.53	71.81	73.36	74.41

从表2可以看出,本文算法识别率高于其他算法,并随着维数增多而增高。

为探究各算法鲁棒性,实验2赋予算法不同的参数值,参数1与参数2分别取值{0.001, 0.005, 0.01, 0.05},参数1限制正则化项对识别率的影响,参数2限制稀疏编码对识别率的影响,实验结果如图8所示,本文算法与FDDL算法识别率随参数1、2的取值增大而增大,并最终达到稳定,本文算法识别率随参数改变的波动小,趋于平缓,鲁棒性强。

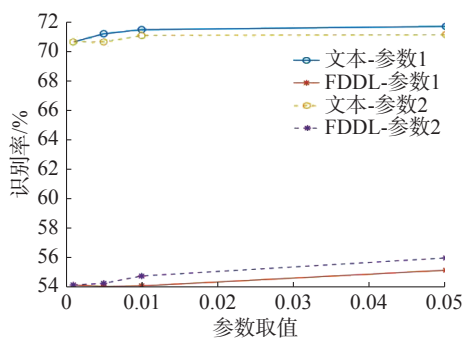


图8 参数对 YaleB 人脸数据库识别率的影响
Fig. 8 The parameter analysis on the YaleB database

3.4 LFW 数据库实验

本文在非受限人脸数据库 LFW 设计两次实验,如图9所示,利用3d校正补齐因转向、遮挡而缺失的特征信息。

实验一,选取单人图片数量大于10张的158人作为实验数据。在具体实验时,选取148个人训练扩充干扰字典和混合特色字典,其余10个人用于训练低秩字典。如图10所示,随机选取每人10张图片,5张图片作为训练集,其

余为测试集。将数据分别将至{50, 100, 150, 200, 250}维,各算法在各维度上的识别率如表3所示。



图9 校正 LFW 人脸数据库
Fig. 9 Correction of LFW Face Database



图10 LFW 人脸数据库样本
Fig. 10 Cropped face samples of LFW database

表3 算法在 LFW database 库上的实验结果
Table 3 Accuracy of different methods on Experiment1 of LFW database

算法	50维	100维	150维	200维	250维
SRC	54.90	63.92	61.18	42.74	21.18
CRC	45.88	60.78	61.96	59.61	59.21
FDDL	45.10	61.57	63.53	70.20	71.76
SVGDL	49.10	61.96	64.04	71.18	71.63
CSSVDL	50.63	62.53	66.82	70.73	70.57
ESRC	49.92	60.04	63.53	70.14	67.38
SCSDL	50.17	59.98	63.92	68.21	68.88
FDDLCSRC	49.17	60.71	62.59	65.43	69.94
LKCFDDL	53.46	68.79	70.63	71.14	72.21
本文算法	54.51	69.80	70.59	72.16	73.33

如表3所示,多数情况下算法的识别率随维数增多而升高,本文算法优于其他算法,但识别率整体不高,这可能是因为前期校正造成的信息损失。

为验证低秩字典对算法影响,实验2随机挑选19、39、59、79个人作为低秩字典训练数据及CSSVDL类内差异字典辅助数据,与基础FDDL

算法、包含辅助字典的 CSSVDL 算法进行对比,各算法在各集合上的识别率如图 11 所示。

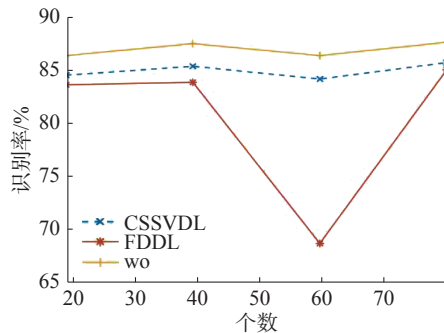


图 11 LFW 人脸数据库识别率

Fig. 11 Accuracy of different methods on experiment1 of LFW database

从图 11 可知,随着构筑类内差异字典的辅助数据增加,CSSVDL 算法的识别率大致增加。随着构筑低秩字典训练数据增加,本文算法识别率大致增加。本文算法、CSSVDL 算法、FDDL 算法在 59 人实验中识别率均受个别样本选择影响下降,本文算法和 CSSVDL 由于辅助数据捕捉数据共性,较 FDDL 算法有更好的鲁棒性。在非受限人脸数据库 LFW 上,本文算法分类效果最优。

3.5 算法评价

3.5.1 复杂度分析

本文算法复杂度分为扩充干扰字典、混合特色字典、低秩字典 3 个部分计算,每部分又分为更新稀疏编码和更新字典两步骤。

以扩充干扰字典为例,设训练样本个数为 n ,样本特征维数为 q ,更新稀疏系数的时间复杂度为 $nO(q^2n')$,其中, $r \geq 1.2$ 为常数。更新字典的时间复杂度为 $\sum_j n_j O(2nq)$,其中, n_j 表示 D_i 的原子个数。

扩充干扰字典总复杂度为

$$nO(q^2n') + \sum_j n_j O(2nq) \quad (10)$$

3.5.2 运行效率分析

为探究算法运行效率,本文设计实验。如图 12 所示,实验 1 取 AR 数据库 80 人,每人前 13 张图片训练,其余图片进行测试。



图 12 AR 实验

Fig. 12 Experiment on AR

实验 2 选取 YALE 数据库,如图 13 所示,每

人前 2 张人脸图像为训练集,剩余 9 张为测试图像。随机选取 5 人作为辅助数据,数据库其余人数作为训练和测试数据,算法运行时间及实验结果如表 4 所示。

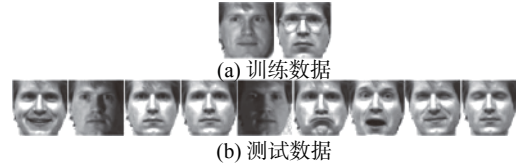


图 13 YALE 实验

Fig. 13 Experiment on YALE

表 4 AR 及 YALE 库实验

Table 4 Experiment on AR and YALE

算法	AR数据库		YALE数据库	
	时间/s	识别率/%	时间/s	识别率/%
FDDL	1801	84.42	228	85.19
SVGDL	1859	91.56	255	90.42
CSSVDL	2202	93.11	292	92.38
SCSDL	2674	92.05	343	90.42
LKCFDDL	1879	92.26	244	93.83
FDDLCSRC	1864	92.17	239	91.53
本文算法	2097	95.19	277	95.78

综合表 4 可以看出,本文算法、CSSVDL 算法、SCSDL 算法由于辅助数据的构建,虽然算法识别率提高,但算法运行时间也随之增加。两实验中本文算法识别率最高,且运行时间小于 CSSVDL 算法、SCSDL 算法,证明本文算法具有更高的性价比。

4 结束语

本文提出稀疏综合字典模型,加入 Fisher 判别准则,学习混合特色字典提取数据共性和特殊性,学习扩充干扰字典与低秩字典提取数据异常、干扰性,分别在 AR、YaleB、LFW 等人脸库上进行实验。实验表明,在小样本训练情境下,即使测试样本与训练样本存在较大差异如表情变化、遮挡等,本文仍能保持较好的性能。在实际应用中还需进一步探讨算法对训练样本的依赖性以及算法的稳定性。

参考文献:

- [1] OH B S, TOH K A, TEOH A B J, et al. An analytic Gabor feedforward network for single-sample and pose-invariant face recognition[J]. *IEEE transactions on image pro-*

- cessing, 2018, 27(6): 2791–2805.
- [2] ZHANG Ni, YAN Yan, CHEN Si, et al. Multi-label learning based deep transfer neural network for facial attribute classification[J]. *Pattern recognition*, 2018, 80: 225–240.
 - [3] 冉鹏, 王灵, 李昕, 等. 改进 Softmax 分类器的深度卷积神经网络及其在人脸识别中的应用 [J]. 上海大学学报 (自然科学版), 2018, 24(3): 352–366.
RAN Peng, WANG Ling, LI Xin, et al. Improved Softmax classifier for deep convolution neural networks and its application in face recognition[J]. *Journal of Shanghai University (Natural Science)*, 2018, 24(3): 352–366.
 - [4] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. *IEEE transactions on signal processing*, 2006, 54(11): 4311–4322.
 - [5] YANG Meng, CHANG Heyou, LUO Weixin. Discriminative analysis-synthesis dictionary learning for image classification[J]. *Neurocomputing*, 2017, 219: 404–411.
 - [6] SHEKHAR S, PATEL V M, CHELLAPPA R. Analysis sparse coding models for image-based classification[C]// *Proceedings of 2014 IEEE International Conference on Image Processing*. Paris, France, 2015.
 - [7] WU Fei, JING Xiaoyuan, YOU Xinge, et al. Multi-view low-rank dictionary learning for image classification[J]. *Pattern recognition*, 2016, 50: 143–154.
 - [8] ZHANG Qiang, LI Baoxin. Discriminative K-SVD for dictionary learning in face recognition[C]// *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA, 2010.
 - [9] JIANG Zhuolin, LIN Zhe, DAVIS L S. Label consistent K-SVD: learning a discriminative dictionary for recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(11): 2651–2664.
 - [10] MAIRAL J, BACH F, PONCE J, et al. Discriminative learned dictionaries for local image analysis[C]// *Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, AK, USA, 2008.
 - [11] CASTRODAD A, SAPIRO G. Sparse modeling of human actions from motion imagery[J]. *International journal of computer vision*, 2012, 100(1): 1–15.
 - [12] 王念兵, 吴秦, 梁久祯, 等. 变化字典学习与显著特征提取的单样本人脸识别 [J]. 小型微型计算机系统, 2017(9).
WANG Nian-bing, WU Qin, LIANG Jiu-zhen, XU Jie, ZHANG Huai. Single Sample Face Recognition via Variant Dictionary Learning and Salient Feature Extraction[J]. *Journal of Chinese Computer Systems*, 2017(9).
 - [13] YANG Meng, ZHANG Lei, FENG Xiangchu, et al. Sparse representation based fisher discrimination dictionary learning for image classification[J]. *International journal of computer vision*, 2014, 109(3): 209–232.
 - [14] WRIGHT J, YANG A Y, GANESH A, et al. Robust face recognition via sparse representation[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2009, 31(2): 210–227.
 - [15] SPRECHMANN P, SAPIRO G. Dictionary learning and sparse coding for unsupervised clustering[C]// *Proceedings of 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*. Dallas, TX, USA, 2010: 2042–2045.
 - [16] RAMIREZ I, SPRECHMANN P, SAPIRO G. Classification and clustering via dictionary learning with structured incoherence and shared features[C]// *Proceedings of 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA, 2010.
 - [17] YANG Meng, ZHANG Lei, FENG Xiangchu, et al. Fisher discrimination dictionary learning for sparse representation[C]// *Proceedings of 2011 IEEE International Conference on Computer Vision (ICCV)*. Barcelona, Spain, 2011: 543–550.
 - [18] 王威, 朱宗玖, 陆俊. 基于字典学习和局部约束的稀疏表示人脸识别 [J]. 电脑知识与技术, 2018, 14(5): 200–202.
WANG Wei, ZHU Zongjiu, LU Jun. Sparse representation for face recognition based on dictionary learning and locality constraint[J]. *Computer knowledge and technology*, 2018, 14(5): 200–202.
 - [19] 黎明, 彭秀姣, 王艳. 基于改进的字典学习与稀疏表示的人脸表情识别 [J]. 系统仿真学报, 2018, 30(1): 28–35, 44.
LI Ming, PENG Xiujiao, WANG Yan. Facial expression recognition based on improved dictionary learning and sparse representation[J]. *Journal of system simulation*, 2018, 30(1): 28–35, 44.
 - [20] DENG Weihong, HU Jiani, GUO Jun. Extended SRC: undersampled face recognition via intraclass variant dictionary[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(9): 1864–1870.
 - [21] LI Wenjing, LIANG Jiuzhen, WU Qin, et al. An efficient face classification method based on shared and class-specific dictionary learning[C]// *Proceedings of 2015 IEEE*

- International Conference on Image Processing (ICIP). Quebec City, QC, Canada, 2015.
- [22] LI Wenjing, LIANG Jiuzhen. Adaptive face representation via class-specific and intra-class variation dictionaries for recognition[J]. *Multimedia tools and applications*, 2018, 77(12): 14783–14802.
- [23] NEEDELL D, TROPP J A. CoSaMP: iterative signal recovery from incomplete and inaccurate samples[J]. *Applied and computational harmonic analysis*, 2009, 26(3): 301–321.
- [24] DO T T, GAN Lu, NGUYEN N, et al. Sparsity adaptive matching pursuit algorithm for practical compressed sensing[C]//Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers. Pacific Grove, CA, USA, 2009.
- [25] TIBSHIRANI R. Regression shrinkage and selection via the lasso[J]. *Journal of the royal statistical society: series B (Methodological)*, 1996, 58(1): 267–288.
- [26] EFRON B, HASTIE T, JOHNSTONE I, et al. Least angle regression[J]. *Annals of statistics*, 2004, 32(2): 407–499.
- [27] MAIRAL J, BACH F, PONCE J. Task-driven dictionary learning[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(4): 791–804.
- [28] PENG Yigang, GANESH A, WRIGHT J, et al. RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(11): 2233–2246.

作者简介:



狄岚, 副教授, 中国人工智能学会粒计算与知识发现专业委员会委员, 中国计算机学会会员, 江苏省“六大人才高峰”资助对象, 主要研究方向为数字图像处理和计算机仿真。近几年先后主持及参与国家级、省部级科研项目 7 项, 主持校级科研项目 4 项、企业

合作项目近 20 项, 获省级自然科学学术奖 1 次, 行业联合会技术奖 3 次。发表学术论文 50 余篇。



矫慧文, 硕士研究生, 中国计算机学会会员, 主要研究方向为图像处理和人脸识别。



梁久祯, 教授, 博士, 中国计算机学会会员, 主要研究方向为计算机视觉和数字图像处理。主持项目 10 余项, 获得浙江省青年英才奖等。发表教材与专著 4 部, 专利成果 57 项, 发表学术论文 160 余篇。