

# 一种新的低秩分块稀疏表示的人脸识别算法

胡昭华<sup>1,2</sup> 赵孝磊<sup>1</sup> 徐玉伟<sup>1</sup> 何 军<sup>1</sup>

(1. 南京信息工程大学电子与信息工程学院, 南京, 210044; 2. 南京信息工程大学江苏省大气环境与装备技术协同创新中心, 南京, 210044)

**摘要:** 针对人脸图像中表情变化、遮挡、光照的问题, 本文提出了一种新颖的基于低秩分块稀疏表示的人脸识别算法。该算法采用了一种新的结构不相关的低秩矩阵恢复方法, 同时采用离散余弦变换方法联合处理人脸图像中遮挡、掩饰和光照的问题, 对处理过的图片采用一种独特的重叠分块方法, 利用冗余信息有效地提高了算法的识别率。在分类阶段, 利用 Alignment pooling 的方法, 有效地提高了识别速度。该算法在标准人脸数据库上进行了多次实验, 实验结果表明: 与现有人脸识别算法相比, 算法的识别准确率和计算效率都得到了一个提高。

**关键词:** 人脸识别; 低秩恢复; 稀疏表示; 重叠分块

**中图分类号:** TN911.73      **文献标志码:** A

## Face Recognition Algorithm Based on Novel Low-Rank and Block-Based Sparse Representation

Hu Zhaohua<sup>1,2</sup>, Zhao Xiaolei<sup>1</sup>, Xu Yuwei<sup>1</sup>, He Jun<sup>1</sup>

(1. School of Electronic & Information Engineering, Nanjing University of Information Science & Technology, Nanjing, 210044, China; 2. Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, Nanjing University of Information Science & Technology, Nanjing, 210044, China)

**Abstract:** Aiming at the problem of human faces with varying expression and illumination, as well as occlusion and disguise, a face recognition algorithm is proposed based on local structural sparse representation. This algorithm combines low-rank matrix recovery with structural incoherence and discrete cosine transform (DCT) method to remove occlusion, disguise and illumination variations in face image. Meanwhile, the partial information is fully utilized by using sparse codes of local image patches with spatial layout. In the classification stage, the algorithm effectively improves the recognition rate based on a novel alignment pooling method. Extensive experiments are conducted on publicly available face databases. Compared with the related state-of-the-art methods, the experimental results demonstrate the accuracy and efficiency of the proposed method.

**Key words:** face recognition; low-rank recovery; sparse representation; overlapped block

## 引 言

人脸识别是计算机视觉领域里一项热门的研究课题,它属于生物特征识别技术,利用生物体(一般特指人)本身的生物特征来区分生物体个体。人脸识别技术广泛采用特征分析算法,融合了计算机图像处理技术与生物学统计技术,利用生物统计学原理进行分析并建立相应的数学模型,具有广阔的发展前景。

对于给定的人脸图像训练数据,设计一个人脸识别系统,大多人都着眼于人脸特征的提取和分类模型的学习,却忽视了用来评估系统性能的测试数据。为了进一步评估所设计的人脸识别算法的鲁棒性,遮挡和伪装等现象将会出现在测试数据中。同时本文注意到尽管测试数据是损坏的,但训练数据往往从一些比较理想的人脸图像中取得,这些人脸图像训练数据没有遮挡、伪装、光照以及表情的变化等等。当把现有的人脸识别系统应用到实际的案例中时,不得不丢弃那些损坏的训练数据,同时可能会遇到小样本和过度拟合的问题。另外,丢弃损坏的人脸图像训练数据可能会丢失对识别有用的信息。在人脸识别中,许多经典的技术如 Eigenfaces, Fisherfaces, Laplacianfaces<sup>[1]</sup>,它们用来降低人脸图片数据的维数。然而,当这些人脸识别技术遇到人脸图像的遮挡、伪装<sup>[2]</sup>和光照变化时鲁棒性会降低。近期一些关于低秩矩阵恢复的工作有效地解决了训练数据中人脸遮挡、伪装的问题<sup>[3-5]</sup>。但这类方法也有不足之处:首先,低秩矩阵恢复不能处理人脸图像中光照变化的问题;其次,低秩矩阵恢复是对训练数据的预处理,对测试数据中人脸遮挡、掩饰却无能为力;最后,它不能保证求完低秩后训练数据识别效果的好坏。针对以上 3 点不足,本文分别提出了相应的解决方案。

在人脸识别领域,光照变化仍然是一个极具挑战性的问题。现有的解决光照变化问题的方法有许多。例如,直方图均衡化<sup>[6]</sup>、图像灰度校正和对数变换等,它们在光照归一化中得到广泛的应用<sup>[7-8]</sup>。然而,这些全局处理技术难以处理人脸图像中非均匀光照变化的问题。还有一些方法试图提取面部特征不变量来应对光照变化,典型的方法有边缘映射、灰度级图像扩展和 Gabor 滤波<sup>[9]</sup>等等,但实验研究表明,当改变人脸照明方向时,这些方法的效果就会很差。光照变化主要是由于三维人脸模型在不同的照明方向下出现不同的明暗现象<sup>[10]</sup>。最近,一些研究人员试图构建一个 3D 人脸模型来解决图像中的光照补偿问题<sup>[11-13]</sup>。但这种基于人脸模型的处理方法有一个缺点:大量不同光照条件下的图片的三维信息需要在训练阶段获取,这大大降低了识别速度。伴随着离散余弦变换的提出与其在图像处理中的运用,使得图像光照变化问题得到有效地解决。本文通过运用离散余弦变换与低秩分块共同处理人脸图像中光照的问题。

稀疏编码是最近比较流行的人脸识别算法<sup>[14]</sup>。稀疏编码利用  $l_1$  范数和  $l_0$  范数达到了令人满意的识别精度,类似的算法还有线性回归分类<sup>[15]</sup>。这些方法在某些控制条件下能够表现很好,然而当测试图像或训练图像中有连续性遮挡、伪装的时候表现却很差。因此,模块化方法<sup>[16]</sup>才得以应用到稀疏表示编码和线性回归分类之中。然而,模块化稀疏表示编码和模块化线性回归分类有一个共同的缺点:它们的每一个模块都独立处理,失去了模块与模块之间的关联信息。正是基于上述问题,本文提出一种新的人脸图像重叠分块方法。人脸图像重叠分块利用冗余信息可以有效地处理图像中部分遮挡、伪装问题。同时在识别阶段,为了利用模块之间的关联信息,本文提出了一个基于 Alignment pooling<sup>[17]</sup>的分类方法,通过 Pooling 同一类的稀疏系数,不仅有利于分类识别,而且提高了识别速率。同时,为了增强求取低秩<sup>[18-19]</sup>后训练数据的识别效果,本文提出了一种新颖的基于结构不相关的低秩矩阵恢复算法。该算法引进了一个能够提高类与类之间识别能力的参考项,对进一步提升人脸识别水平有着积极的作用。

## 1 低秩矩阵恢复

主成分分析(Principal component analysis, PCA)<sup>[2]</sup>作为一种经典的算法,广泛应用于数据分析和

数据降维。然而 PCA 却对大量的稀疏误差很敏感。为了有效地解决稀疏噪声的问题,大量相关工作被提出,如:交替最小化技术、影响函数的引入、低秩矩阵恢复<sup>[3]</sup>等。在这些方法中,低秩矩阵恢复能更有效地控制和去除稀疏噪声,本文算法正是基于低秩矩阵恢复技术。

低秩矩阵恢复的目的为将数据矩阵  $D$  分解成  $A + E$ , 其中  $A$  代表低秩矩阵,  $E$  代表分解后的稀疏误差。更准确地说,对于给定的输入数据矩阵  $D$ , 低秩矩阵恢复通过最小化矩阵  $A$  的秩, 同时减小  $\|E\|_0$  的值来达到输入数据矩阵  $D$  的最佳低秩逼近。但是上述的问题是 NP-hard 问题, Candes<sup>[3]</sup> 通过求解如下公式使传统的低秩矩阵恢复变得易于处理

$$\min_{A, E} \|A\|_* + \lambda \|E\|_1 \quad \text{s. t.} \quad D = A + E \quad (1)$$

式中:核范数  $\|A\|_*$  为矩阵  $A$  秩的近似;零范数  $\|E\|_0$  被一范数  $\|E\|_1$  所替代。解决这个凸优化问题就相当于解决了原始的低秩矩阵近似问题,只要保证矩阵  $A$  的秩不要太大,同时矩阵  $E$  足够稀疏。由于增广拉格朗日乘法(Augmented Lagrange multiplier, ALM)<sup>[3,20]</sup> 的高效性,在求解低秩矩阵近似问题时得到广泛应用。

## 2 离散余弦变换

离散余弦变换(Discrete cosine transform, DCT)<sup>[21]</sup>对相关的图像数据具有非常好的能量聚焦性,它能够将图像从空间域变换到频域,经过变换,图像信号能量的绝大部分被集中到变换域的少数系数上。因此,对于受光照影响的图像,只需要修改很少的频域系数,就可以对图像的光照做出较好的调整,避免了需要调节多个参数以适合不同图像的问题。

DCT 有 3 个很重要的特点:(1)在 DCT 所有系数中,绝大多数系数都接近 0;(2)DCT 的高频系数很小,接近于 0,比较大的系数一般集中在低频部分;(3)以人脸图像为例,低频系数包含的是人脸图像的绝大部分非细节特征信息,而中高频系数包含的则是人脸细节信息。根据 DCT 的这些特点,本文采用此算法来构建一个更具有鲁棒性的系统。DCT 算法的效果如图 1 所示。

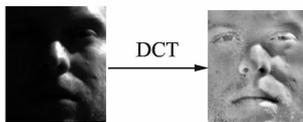


图 1 基于 DCT 变换的人脸光照处理

Fig. 1 Human face illumination processing based on DCT transform

## 3 基于结构不相关的低秩矩阵恢复

基于结构不相关的低秩矩阵恢复是算法的重要组成部分。算法通过对低秩矩阵分解的观察与分析,引入一个能够提高字典辨别能力的参考项。但参考项的引入会引起目标函数的变化,必然要求采取新的算法解决相应的优化问题。

### 3.1 低秩矩阵分解

低秩矩阵恢复将输入数据矩阵分解成两个不同的部分:一个代表低秩矩阵部分,另一个代表稀疏误差部分<sup>[18]</sup>。低秩矩阵分解既要保证输入数据的结构,如图像的纹理;又要保证误差矩阵足够稀疏。当将低秩矩阵恢复应用于一个有  $c$  个主体的人脸识别系统中时,相应的训练矩阵为  $D = [D_1, D_2, \dots, D_c]$ , 其中  $D_i$  为主体  $i$  的数据矩阵,包括人脸图像的遮挡,伪装等,如图 2(a)所示。当执行低秩矩阵分解时,输入数据矩阵  $D = [D_1, D_2, \dots, D_c]$  将会分解成低秩矩阵  $A = [A_1, A_2, \dots, A_c]$  和误差矩阵  $E = [E_1, E_2, \dots, E_c]$  之和。如图 2(b)所示,低秩矩阵  $A$  中图像可以认为是原始的人脸图像去除稀疏噪声以后效果。从

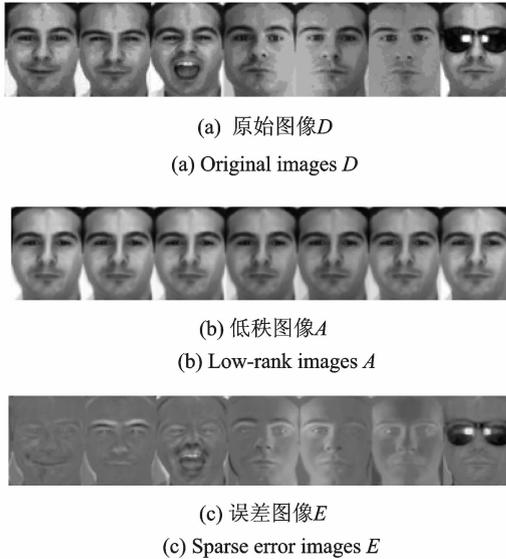
(a) 原始图像  $D$ (a) Original images  $D$ (b) 低秩图像  $A$ (b) Low-rank images  $A$ (c) 误差图像  $E$ (c) Sparse error images  $E$ 

图 2 低秩矩阵恢复

Fig. 2 Low-rank matrix recovery

而也可以看出,低秩矩阵  $A$  比原始的数据矩阵  $D$  更具有代表性,更加有利于人脸识别。

尽管低秩矩阵  $A$  比原始数据矩阵  $D$  有更强的代表能力,但不同主体的人脸图片却有着一些共同的特征,如眼睛、鼻子的位置等,这也可能导致低秩矩阵  $A$  不能够包含足够的区分信息。因此,本文决定提高低秩矩阵之间的非相关性<sup>[18,22]</sup>,尽量保持不同低秩矩阵之间的独立性。根据低秩矩阵恢复式(1),本文对目标函数增加了一个能够增强低秩矩阵之间非相关性的参考项。新的目标函数表示如下

$$\min_{A,E} \sum_{i=1}^c \{ \|A_i\|_* + \lambda_1 \|E_i\|_1 \} + \lambda_2 \Psi(A_1, A_2, \dots, A_c) \quad \text{s. t.} \quad D_i = A_i + E_i \quad (2)$$

式中:  $i=1,2,\dots,c$ ,  $c$  为训练矩阵中主体的个数;  $D_i$  为第  $i$  个训练样本数据矩阵;  $A_i$  为第  $i$  个训练样本的低秩分解矩阵;  $E_i$  为稀疏误差矩阵;  $\Psi(A_1, A_2, \dots, A_c)$  为提高低秩矩阵区分能力的参考项;  $\lambda_2$  为常数 ( $\lambda_2 \geq 0$ )。图 3 为人脸识别算法。算法分为两个阶段:训练阶段输入数据矩阵经过低秩分解、光照归一化、重叠分块组成一个新的字典;测试阶段测试图片经过重叠分块、稀疏编码,然后通过 Alignment pooling 分类达到人脸识别分类的效果。DCT 光照归一化方法是对经过低秩矩阵分解的人脸图像采用 DCT 变换,同时应用非局部均值去噪算法(Non-local means, NLM)<sup>[23]</sup>使图像光照归一化。NLM 通过采用非局部均值去噪的方法计算光照度函数,估算出反射系数。图 4 为不同条件下的人脸图像首先通过低秩矩阵分解,然后经过 DCT 光照归一化后的效果图。以图 4(a)为例,最左侧的人脸图像为原始图像,中间的为经过低秩矩阵分解后的人脸图像,最右侧的为经过 DCT 光照归一化后的人脸图像。

### 3.2 参考项的引入

参考项  $\Psi(A_1, A_2, \dots, A_c)$  的设计不仅有利于字典学习,还能够尽可能地提高字典的辨别能力。线性判别分析(Linear discriminant analysis, LDA)<sup>[22,24]</sup>的目的是从高维特征空间里提取出最具有判别能力的低维特征,这些特征能将同一个类别的所有样本聚集在一起,不同类别的样本尽量分开。根据 LDA 的准则,定义以下矩阵向量。

各类样本的均值向量

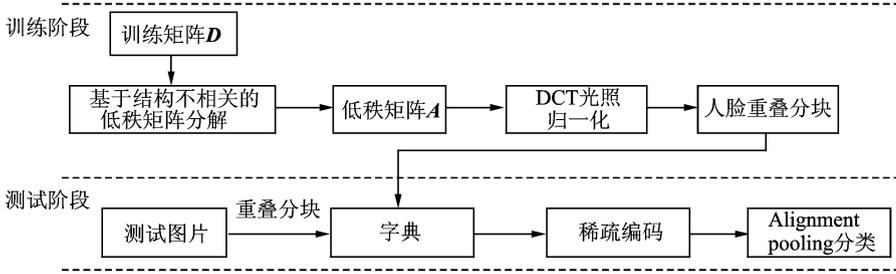


图 3 本文算法流程图

Fig. 3 Flow diagram of proposed algorithm



图 4 经过低秩分解和 DCT 光照归一化后的图像

Fig. 4 Recovered images based on low-rank recovery and DCT transform

$$m_i = \frac{1}{n_i} \sum_{a \in A_i} a \tag{3}$$

样本的总体均值向量

$$m = \frac{1}{n} \sum_{i=1}^c n_i \cdot m_i \tag{4}$$

样本的类内离散度矩阵

$$S_w = \sum_{i=1}^c \sum_{a \in A_i} (a - m_i)(a - m_i)^T \tag{5}$$

样本的类间离散度矩阵

$$S_b = \sum_{j=1}^c n_j (m_j - m)(m_j - m)^T \tag{6}$$

式(3~6)中,  $i=1, 2, \dots, c$ ;  $n_i$  为第  $i$  个主体中样本的个数;  $n$  为训练样本的总数  $n = n_1 + n_2 + \dots + n_c$ ;  $A_i$  为第  $i$  个主体的训练数据矩阵;  $a$  为  $A_i$  中的列向量。

本文根据 LDA 的原理和样本类间离散矩阵与类内离散矩阵的特性将参考项定义为

$$\Psi(A_1, A_2, \dots, A_c) = \text{tr}(S_w) - \text{tr}(S_b) \tag{7}$$

式中:  $\text{tr}(\cdot)$  为求矩阵的迹。将式(7)代入式(2), 可得

$$\min_{A, E} \sum_{i=1}^c \{ \|A_i\|_* + \lambda_1 \|E_i\|_1 \} + \lambda_2 (\text{tr}(S_w) - \text{tr}(S_b)) \quad \text{s. t.} \quad D_i = A_i + E_i \tag{8}$$

由于算法是对不同主体的数据矩阵分别求低秩矩阵分解, 因此为了避免对式(8)的直接求解, 根据文献[18, 22]可将目标函数可以转化为

$$\min_{A, E} \|A_i\|_* + \lambda_1 \|E_i\|_1 + \lambda_2 \psi(A_i) \quad \text{s. t.} \quad D_i = A_i + E_i \tag{9}$$

式中:  $\psi(A_i) = \|A_i - [m_i, m_i, \dots, m_i]_{d \times n_i}\|_F^2 - \sum_{j=1}^c \|[m_j, m_j, \dots, m_j]_{d \times n_j} - [m, m, \dots, m]_{d \times n_j}\|_F^2$ 。由于在

对第  $i$  个主体  $\mathbf{A}_i$  求低秩分解时,其他主体  $\mathbf{A}_j (i \neq j)$  已经固定,因此  $\zeta = \sum_{j=1}^c \|[m_j, m_j, \dots, m_j]_{d \times n_i} - [m, m, \dots, m]_{d \times n_i}\|_F^2$  为一常量,令  $\mathbf{M}_i = [m_i, m_i, \dots, m_i]_{d \times n_i}$ ,  $d$  为人脸图像化成列向量后的行数。因此式(9)可以化为

$$\min_{\mathbf{A}_i, \mathbf{E}_i} \|\mathbf{A}_i\|_* + \lambda_1 \|\mathbf{E}_i\|_1 + \lambda_2 \|\mathbf{A}_i - \mathbf{M}_i\|_F^2 - \lambda_2 \zeta \quad \text{s. t.} \quad \mathbf{D}_i = \mathbf{A}_i + \mathbf{E}_i \quad (10)$$

由于  $\lambda_2 \zeta$  是常量,目标函数最终整理得

$$\min_{\mathbf{A}_i, \mathbf{E}_i} \|\mathbf{A}_i\|_* + \lambda_1 \|\mathbf{E}_i\|_1 + \lambda_2 \|\mathbf{A}_i - \mathbf{M}_i\|_F^2 \quad \text{s. t.} \quad \mathbf{D}_i = \mathbf{A}_i + \mathbf{E}_i \quad (11)$$

对于式(11),本文利用 ALM 对每一个主体数据矩阵做低秩分解。

### 3.3 基于 ALM 的算法优化

ALM 作为一种标准的算法广泛应用于求解低秩矩阵恢复的问题。当约束条件  $h(\mathbf{X}) = 0$  时,求解  $f(\mathbf{X})$  的最小值,ALM 的目标函数定义为

$$L(\mathbf{X}, \mathbf{Y}, u) = f(\mathbf{X}) + \langle \mathbf{Y}, h(\mathbf{X}) \rangle + \frac{u}{2} \|h(\mathbf{X})\|_F^2 \quad (12)$$

式中:  $\mathbf{Y}$  代表拉格朗日乘数矩阵。令

$$f(\mathbf{X}) = \|\mathbf{A}_i\|_* + \lambda_1 \|\mathbf{E}_i\|_1 + \lambda_2 \|\mathbf{A}_i - \mathbf{M}_i\|_F^2 \quad (13)$$

$$h(\mathbf{X}) = \mathbf{D}_i - \mathbf{A}_i - \mathbf{E}_i \quad (14)$$

则式(11)在利用 ALM 算法时可表示如下

$$\begin{aligned} L(\mathbf{A}_i, \mathbf{E}_i, \mathbf{Y}, \mu, \lambda_2) = & \|\mathbf{A}_i\|_* + \lambda_1 \|\mathbf{E}_i\|_1 + \lambda_2 \|\mathbf{A}'_i\|_F^2 + \\ & \langle \mathbf{Y}_i, \mathbf{D}_i - \mathbf{A}_i - \mathbf{E}_i \rangle + \frac{\mu}{2} \|\mathbf{D}_i - \mathbf{A}_i - \mathbf{E}_i\|_F^2 \end{aligned} \quad (15)$$

式中:  $\mathbf{A}'_i = \mathbf{A}_i - \mathbf{M}_i$ , 为了求解式(15),算法通过迭代求解  $\mathbf{A}_i$ ,  $\mathbf{E}_i$  和  $\mathbf{Y}$ , 该算法的流程图如图 5 所示。

(1)  $\mathbf{A}_i$  更新

$$\mathbf{A}_i^{k+1} = \underset{\mathbf{A}_i}{\operatorname{argmin}} \epsilon \|\mathbf{A}_i\|_1 + \frac{1}{2} \|\mathbf{X}_a - \mathbf{A}_i\|_F^2 \quad (16)$$

式中:  $\mathbf{X}_a = \mu^k \epsilon (\mathbf{D}_i - \mathbf{E}_i + \frac{1}{\mu^k} \mathbf{Y}_i) + \mathbf{M}_i$ ;  $\epsilon = (2\lambda_2 + \mu^k)^{-1}$ 。

通过矩阵奇异值分解

$$\mathbf{A}_i^{k+1} = \mathbf{USV}^T \quad \text{s. t.} \quad (\mathbf{U}, \mathbf{S}, \mathbf{V}^T) = \operatorname{SVD}(\mathbf{X}_a)$$

(2)  $\mathbf{E}_i$  更新

$$\mathbf{E}_i^{k+1} = \underset{\mathbf{E}_i}{\operatorname{argmin}} \epsilon' \|\mathbf{E}_i\|_1 + \frac{1}{2} \|\mathbf{X}_e - \mathbf{E}_i\|_F^2 \quad (17)$$

式中:  $\epsilon' = \frac{1}{\mu^k}$ ;  $\mathbf{X}_e = \mathbf{D}_i - \mathbf{A}_i^{k+1} + \frac{1}{\mu^k} \mathbf{Y}_i^k$ 。式(16,17)可以根据式(13-15)推导出来。当  $\mathbf{A}_i$ ,  $\mathbf{E}_i$  都被迭代出来时,低秩矩阵分解也就完成。

上述凸优化问题及其收敛性可以通过  $l_1$  最小化技术<sup>[25]</sup>求解和证明。

## 4 人脸重叠分块与识别

### 4.1 人脸重叠分块

给定一个由  $c$  个主体  $n$  个训练图片组成的数据集  $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ , 其中  $\mathbf{a}_i$  代表第  $i$  个训练图片的特征向量。每一个训练图片首先重新采样到实验要求的尺寸,然后重叠分成 12 个同样大小的图像块。图 6 是一个  $80 \times 64$  像素的训练图片分块的情况,每一个图像块的大小为  $32 \times 32$  像素,图像块 2 由图像块 1 所在的矩形框向右平移 16 个像素得到,同样地,图像块 4 由图像块 1 所在的矩形框向下平移 16 个

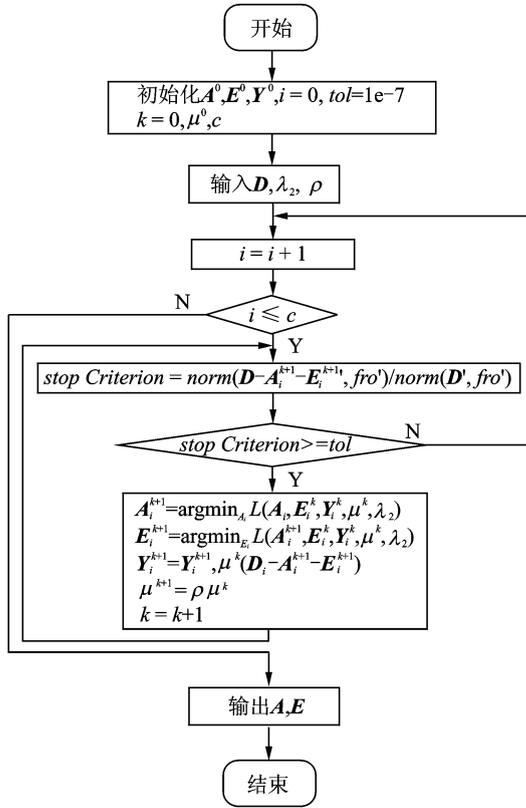


图 5 基于结构不相关的低秩矩阵恢复算法

Fig. 5 Low-rank matrix recovery algorithm with structural incoherence

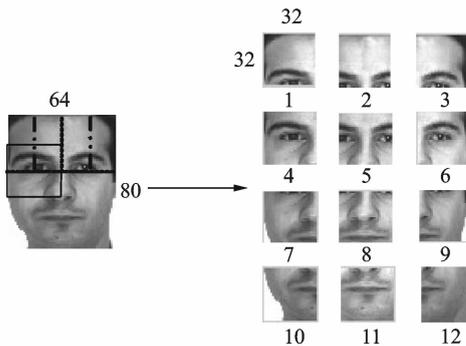


图 6 人脸图片重叠分块

Fig. 6 Face image overlapping partition

像素得到,其他的图像块依次类推。相应地,  $A$  也被分为  $A = [A_1^T, A_2^T, \dots, A_{12}^T]^T$ , 其中  $A_k \in \mathbf{R}^{d \times n}$ ,  $A_k$  的第  $l$  列代表第  $l$  个训练图片第  $k$  个模块的特征向量。按同样的分块方式,测试图片也相应地分成 12 个模块  $y_1, y_2, \dots, y_{12}$ 。

文献[4]提出了一种基于稀疏表示的人脸识别算法(Sparse representation classification, SRC)<sup>[14]</sup>。SRC 认为每一个测试图片是训练图片集的线性组合,可以通过  $l_1$  最小化来求解其中的稀疏编码问题。

根据 SRC 相关的理论,对于任意一个测试模块  $y_k$ , 有

$$y_k = A_k x_k \tag{18}$$

由于这是一个欠定方程,式(18)有很多解,需要添加一个约束条件。因为  $l_2$  范数最小化比较简单方便,使其成为最普遍的约束条件。然而,它产生的稀疏系数比较稠密,对测试图像识别没有太大益处。为了获得一个比较理想的稀疏解,本文利用了  $l_1$  范数最小化。其最优化问题

$$x_k = \operatorname{argmin}_x \{ \| A_k x - y_k \|_2^2 + \lambda \| x \|_1 \} \quad 1 \leq k \leq 12 \tag{19}$$

式中:  $x_k$  为第  $k$  个模块相应的稀疏系数。

### 4.2 识别方法

在基于稀疏表示的人脸识别中,一般都通过比较残差的大小来判断测试图片属于哪一个训练样本,从而实现正确地分类

$$i = \operatorname{argmin}_j \| A_i \delta_j(x_k) - y_k \|_2 \tag{20}$$

式中:  $\delta_j(x_k)$  代表在  $x_k$  中与第  $j$  个类相关联的特征函数<sup>[14]</sup>。

为了避免单独对模块进行处理而忽略模块之间的相关性,本文算法充分利用不同模块的稀疏系数,借鉴 Alignment pooling<sup>[17]</sup>的方法,先将不同模块的属于同一类的稀疏系数叠加在一起,然后对同一类稀疏系数从大到小排序,这样既能利用模块之间的相关信息,又能突出权重比较高的模块在分类过程中的重要性。如图 7 所示。对于一个有  $c$  个主体每个主体有  $N$  个训练样本的训练集,它的测试图片第  $k$  个模块所对应的稀疏系数为  $x_k \in \mathbf{R}^{c \times 1}$  ( $1 \leq k \leq 12$ )。其中  $v_{kj}$  代表测试图片第  $k$  个模块第  $j$  个主体所对应的同一类的稀疏系数绝对值之和,即

$$v_{kj} = \| \delta_j(x_k) \|_1 \tag{21}$$

因此  $v_k = [v_{k1}, v_{k2}, \dots, v_{kc}]^T$ 。利用 pooling 的方法,将各个模块的稀疏系数联系起来有

$$f = \frac{1}{12} \sum_{k=1}^{12} v_k \tag{22}$$

式中:  $f \in \mathbf{R}^{c \times 1}$ , 通过求解  $f$  中最大值的坐标即可确定测试图片属于哪一类,从而实现正确分类与识别。

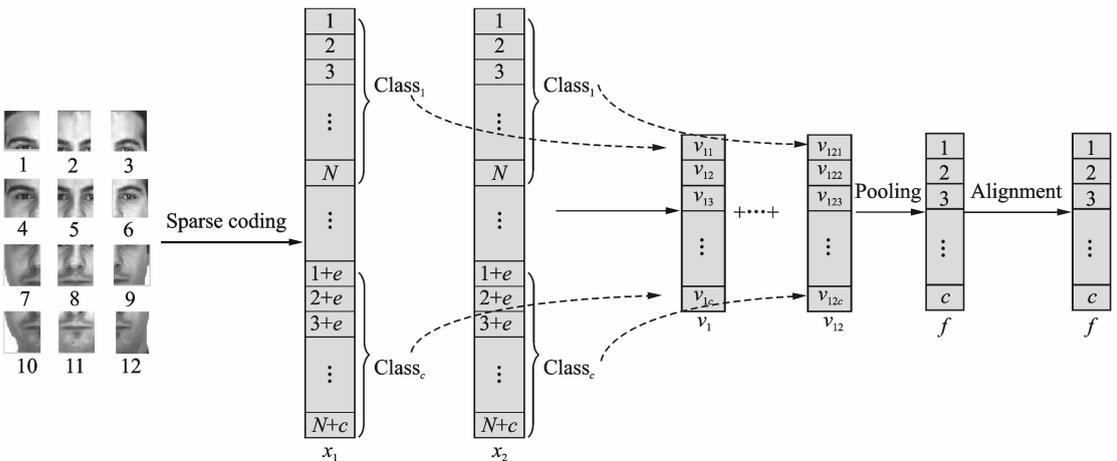


图 7 本文基于 Alignment pooling 的分类方法  
Fig. 7 Classification method based on alignment pooling

## 5 实验结果与分析

实验过程中所用到的数据库与文献[14]中所用到的数据库相同,分别为:Extended Yale B 数据库、

AR 数据库和 Multi PIE 数据库。实验过程中,将本文算法与现有的算法:Global SRC<sup>[14]</sup>, Modular SRC<sup>[14]</sup>, Global LRC<sup>[26]</sup>, Modular LRC<sup>[26]</sup>, NN<sup>[27]</sup> 和 CRC\_RLS<sup>[28]</sup> 进行比较。同时,为了弄清 DCT 光照处理与 LDA 参考项对本实验算法的影响,分别做了不使用 DCT 变换(Without DCT)与不加 LDA 参考项(Without LDA)的实验。

### 5.1 Extended Yale B 数据库

在 Extended Yale B 数据库中包含 38 个主体,共 2 280 幅图片,每幅图片的大小为 168×192 像素,每类包含 60 种不同光照条件下的人脸图片。本文随机从每一类中抽取 55 幅图片作为训练样本,另外 5 幅作为测试样本。同时,实验中将每幅图片重新采样为 20×16 像素,40×32 像素,80×64 像素,目的是比较在不同尺寸下不同算法之间的识别效果。实验识别率如表 1 所示。

表 1 不同图像尺寸下算法的识别率

Table 1 Algorithm of recognition rate under different image sizes

图像大小 / 像素	Global SRC	Modular SRC	Global LRC	Modular LRC	CRC_RLS	Without DCT	Without LDA	本文算法
20×16	95.1	67.3	95.2	60.2	95.0	90.7	93.4	94.7
40×32	97.1	97.0	96.3	73.4	96.3	95.0	97.4	98.9
80×64	97.7	98.4	96.6	87.5	98.9	97.5	98.7	100.0

对于尺寸为 20×16 像素的小图像来说,由于缺乏足够的人脸特征,模块化的方法比全局化的方法表现的更差。然而,本文方法通过利用人脸图像中重叠部分的冗余信息,在识别率上接近全局化的方法,而且远远高于模块化的方法。当图像尺寸增加到 40×32 像素时,模块化方法的表现得到了很大提高。然而,模块化方法的优点无法完全克服它自身的缺点,与模块化和全局化的方法相比,本文方法仍然能够表现最好。当处理大图像尺寸 80×64 像素时,模块化、全局化的方法虽然有所提高,但本文算法依然能够获得最高的识别率。Extended Yale B 数据库主要由不同光照条件下的人脸图片构成,从 Without DCT 实验中的识别率可以看出,DCT 光照处理对本文算法识别具有重要作用。Without LDA 实验中的识别率虽然比本文算法要低一点,但相比其他几种算法还是有所提高。同时,本文在不同的维度的特征空间下比较各种算法的识别效果,如图 8 所示。从图 8 可看出,本文算法与其他几种算法相比,识别率有相应的提高。

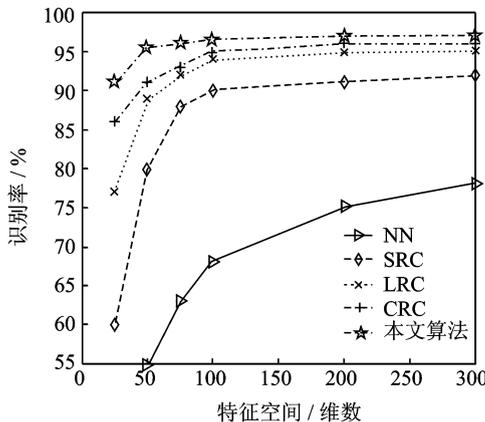


图 8 在 Extended Yale B 数据库中实验效果图

Fig. 8 Recognition performance on the extended Yale B database

### 5.2 AR 数据库

AR 数据库包含 126 个主体,共 4 000 幅人脸图片。实验中,本文从男性图片中选择 50 个主体,从每个主体中随机选择 20 幅作为训练图片,另外 6 幅作为测试图片,图像的尺寸为  $80 \times 64$  像素。实验识别率如表 2 所示。

表 2 不同遮挡下算法的识别率  
Table 2 Algorithm of recognition rate under different occlusion

Scenarios	Global SRC	Modular SRC	Global LRC	Modular LRC	CRC_RLS	Without DCT	Without LDA	本文算法
Sunglass	87.00	97.50	65.50	93.50	97.90	97.60	96.00	98.33
Scarf	59.50	93.50	12.50	73.50	97.60	97.50	96.30	98.90

当处理人脸图像中有眼镜、围巾等这些遮挡时,算法 Global SRC, Global LRC, Modular LRC 表现很差,无法处理这些遮挡。Modular SRC 虽然表现良好,但和本文算法还是有些差距,从而充分体现本算法的优越性。由于 DCT 变换对人脸图像的非连续遮挡不能进行有效处理,从实验识别率中也可以看出,在 AR 数据库中 DCT 变换对本文算法的影响比 LDA 参考项的影响要小得多。实验中有 50 个类共 300 幅测试图像,每个类有 6 幅测试图像。300 个测试图像的识别效果如图 9 所示,图中横坐标代表 300 幅测试图像,纵坐标代表测试图像的识别标签,图中波动比较大的地方出现了错误识别。同时,针对不同的场景,实验分别在不同维度的特征空间下对上述几种算法作出性能的比较。如图 10(a, b)所示,从图中可以看出,在不同的场景下,本文都能获得比其他几种算法更高的识别率。

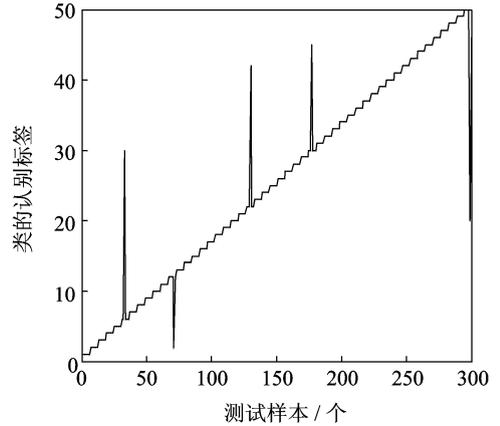
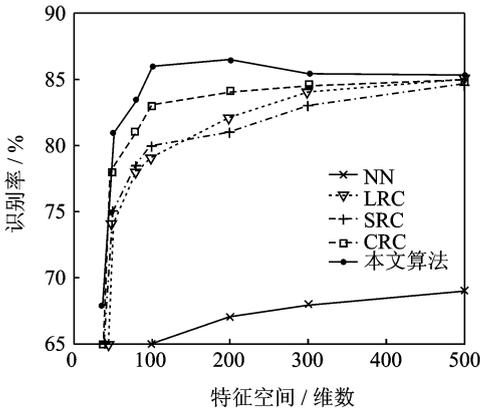
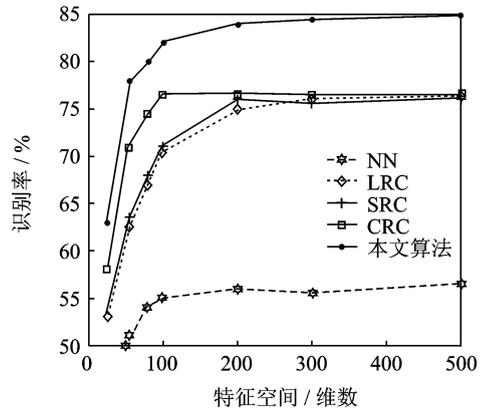


图 9 测试图像分类识别效果图  
Fig. 9 Test image classification results



(a) 眼镜  
(a) Sunglasses



(b) 眼镜围巾  
(b) Scarf

图 10 在 AR 数据库中不同场景下实验对比图

Fig. 10 Performance comparisons on AR database for different scenarios

### 5.3 Multi PIE 数据库

Multi PIE 数据库包含 337 个主体,由 750 000 多幅人脸图片组成。每个主体的图片集分别在 4 种不同的场景下拍摄完成,同时伴随着有姿势、表情和光照的变化。实验中,随机选择 250 个主体,从每个主体中选择 15 幅正面的图片和 15 幅光照变化的图片作为训练图片,随机选择 10 幅作为测试图片。实验在场景 2、场景 3 和场景 4 下进行,每幅图片尺寸为  $80 \times 64$  像素。实验识别率如表 3 所示。与表中几种算法相比,本文算法都有相应的提高。

表 3 不同场景下算法的识别率

场景	LRC	SRC	CRC_RLS	本文方法
场景 2	87.1	93.9	94.1	95.1
场景 3	81.9	90.0	89.3	90.5
场景 4	84.3	94.0	93.3	95.0

## 6 结束语

本文利用低秩矩阵非相关性,DCT 光照归一化与人脸重叠分块设计一个更加具有鲁棒性的人脸识别算法。该算法在不同条件下进行了多次实验,包括光照变化、表情变化、眼镜遮挡、围巾遮挡等,实验结果清楚地表明本文算法比模块化、全局化的方法更加优越,更加具有鲁棒性。然而本文算法也有不足之处,对于一个人脸数据库,预先要经过低秩、光照归一化、重叠分块等一系列复杂的数据处理,使系统变的更加繁琐。接下来的工作将会在不同类型、不同人种的人脸数据库中进行试验和算法改进,使算法更加简洁高效。

### 参考文献:

- [1] He Xiaofei, Yan Shuicheng, Hu Yuxiao, et al. Face recognition using Laplacianfaces[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2005, 27(3): 328-340.
- [2] Dela T F, Black M. Robust principal component analysis for computer vision[C]//Proceeding of International Conference on Computer Vision. Vancouver, Canada: IEEE, 2001: 362-369.
- [3] Candes E, Li Xiaodong, Ma Yi, et al. Robust principal component analysis[J]. Journal of the ACM, 2011, 58(3):1-37.
- [4] Tzimiropoulos G, Zafeiriou S, Pantic M. Subspace learning from image gradient orientations[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2012, 34(12): 2454-2466.
- [5] Qifa K, Kanade T. Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming[C]//Proceeding of Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE, 2005, 1:739-746.
- [6] Pizer S, Amburn E. Adaptive histogram equalization and its variations[J]. Computer Vision Graphics and Image Processing, 1987, 39(3):355-368.
- [7] Shan Shiguang, Gao Wen, Cao Bo. Illumination normalization for robust face recognition against varying lighting conditions [J]. Analysis and Modeling of Faces and Gestures, IEEE International Workshop on, 2003,10(3): 157-164.
- [8] Savvides M, Kumar V. Illumination normalization using logarithm transforms for face authentication[J]. Lecture Notes in Computer Science, 2003,2688:549-556.
- [9] 黄兵,郭继昌. 基于 Gabor 小波与 LBP 直方图序列的人脸年龄估计[J]. 数据采集与处理,2012,27(3):340-345. Huang Bing, Guo Jichang. Age estimation of facial images based on Gabor and histogram sequence of LBP[J]. Journal of Data Acquisition and Processing, 2012, 27(3): 340-345.
- [10] Shashua A, Riklin R T. The quotient image:Class-based rendering and recognition with varying illuminations[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2002, 23(2):129-139.
- [11] Liu Zhonghua, Zhao Haixia, Pu Jiexin. Face recognition under varying illumination[J]. Neural Computing and Applications, 2013, 23(1): 133-139.

- [12] Georgiades A, Belhumeur P, Jacobs D. From few to many: Illumination cone models for face recognition under variable lighting and pose[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2001, 23(6): 630-660.
- [13] Chen Chiaping, Chen Chusong. Intrinsic illumination subspace for lighting insensitive face recognition[J]. *Systems Man & Cybernetics Part B Cybernetics, IEEE Transactions on*, 2012, 42(2): 422-433.
- [14] Wright J, Yang A, Ganesh A. Robust face recognition via sparse representation[J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2009, 31(2): 210-227.
- [15] Pentland A, Moghaddam B, Starner T. View-based and modular eigenspaces for face recognition[C]//*Proceeding of Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE, 1994:84-91.
- [16] Lai Jian, Jiang Xudong. Modular weighted global sparse representation for robust face recognition[J]. *IEEE Signal Processing Letters*, 2012, 19(9):571-574.
- [17] Jia Xu, Lu Huchuan. Visual tracking via adaptive structural local sparse appearance model[C]//*Proceeding of Conference on Computer Vision and Pattern Recognition*. Providence, USA: IEEE, 2012:1822-1829.
- [18] Chen C F, Wei C P, Wang Y C F. Low-rank matrix recovery with structural Incoherence for robust face recognition[C]//*Proceeding of Conference on Computer Vision and Pattern Recognition*. Providence, USA: IEEE, 2012: 2618-2625.
- [19] Ramirez I, Sprechmann P, Sapiro G. Classification and clustering via dictionary learning with structured incoherence and shared features[C]//*Proceeding of Conference on Computer Vision and Pattern Recognition*. San Francisco, USA: IEEE, 2010:3501-3508.
- [20] Lin Zhouchen, Chen Minming. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices [EB/OL]. <http://arxiv.org/abs/1009.5055>, 2010.
- [21] Kang J, Gabbouj M, Kuo C. Sparse/DCT (S/DCT) two-layered representation of prediction residuals for video coding[J]. *IEEE Transactions on Image Processing*, 2013, 22(7): 2711-2722.
- [22] Zhou Ning, Shen Yi, Peng J. Learning inter-related visual dictionary for object recognition[C]//*Proceeding of Conference on Computer Vision and Pattern Recognition*. Providence, USA: IEEE, 2012: 3490-3497.
- [23] Zhao Yitian, Liu Yonghuai, Song Ran. Extended non-local means filter for surface saliency detection[J]. *International Conference on Image Processing*, 2012, 8556: 633-636.
- [24] Duda R, Hart P, Stork D. *Pattern classification*[M]. 2nd ed. New York: Wiley & Interscience, 2001.
- [25] Hale E, Yin W T, Zhang Y. Fixed-point continuation for l1-minimization: Methodology and convergence[J]. *SIAM Journal on Optimization*, 2008, 19(3):1107-1130.
- [26] Jaiswal A, Kumar N, Agrawal R. Local linear regression on hybrid eigenfaces for pose invariant face recognition[J]. *International Journal of Computer Vision & Image Processing*, 2012, 2(2):48-58.
- [27] Hicham G, Elmongui, M F. Continuous aggregate nearest neighbor queries[J]. *GeoInformatica*, 2013, 17(1):63-69.
- [28] Zhang Lei, Yang Meng, Feng Xiangchu. Sparse representation or collaborative representation: Which helps face recognition [C]//*Proceeding of International Conference on Computer Vision*. Barcelona, Spain: IEEE, 2011, 6(13):471-478.

#### 作者简介:



胡昭华 (1981-), 女, 副教授, 研究方向: 视频目标跟踪、模式识别和粒子滤波, E-mail: zhaohua\_hu @ 163.com。



赵孝磊 (1988-), 男, 硕士研究生, 研究方向: 人脸识别。



徐玉伟 (1987-), 男, 硕士研究生, 研究方向: 视频目标跟踪。



何军 (1978-), 男, 讲师, 研究方向: 计算机视觉、高维数据分析和压缩感知。

