

基于局部特征的二维白化重构

田甲略¹, 朱玉莲², 陈飞玥¹, 刘佳慧¹

(1. 南京航空航天大学计算机科学与技术学院/人工智能学院, 南京 211106; 2. 南京航空航天大学公共实验教学部, 南京 211106)

摘要: 白化是一种能够去除数据各属性间相关性的数据预处理方法。最近提出的二维白化重构方法 (Two-dimensional whitening reconstruction, TWR) 是一种针对单张图片的白化方法, 阐述了 TWR 方法等价于基于图像列的 ZCA 白化, 即 TWR 具有去除图像列内相关性的作用; 但是图像局部块内的相关性往往远大于列内, 因此本文从去除图像局部块内相关性的角度出发, 提出了两种 TWR 的改进方法: 基于重组的 TWR (Reshaped-based TWR, RTWR) 方法和基于块的 TWR (Patch-based TWR, PTWR) 方法。RTWR 首先将图像进行重新组合使得每个列向量对应着原始图像的子块, 然后将 TWR 预处理作用在重组后的图像上; 而 PTWR 方法则将 TWR 直接作用在图像的每个子块上。在 ORL、CMU PIE、AR 三个人脸数据集上的实验结果表明, RTWR 和 PTWR 预处理比 TWR 预处理更有利于后续分类性能的提高。

关键词: 二维白化重构; ZCA 白化; 子图像; 图像处理; 相关性

中图分类号: TP391 **文献标志码:** A

Local-Feature-Based Two-Dimensional Whitening Reconstruction

TIAN Jialue¹, ZHU Yulian², CHEN Feiyue¹, LIU Jiahui¹

(1. College of Computer Science and Technology/College of Artificial Intelligence, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China; 2. Fundamental Experimental Teaching Department, Nanjing University of Aeronautics & Astronautics, Nanjing 211106, China)

Abstract: Whitening is a preprocessing method that can remove the correlation between variables of data. Two-dimensional whitening reconstruction (TWR) is a new whitening method for a single image. In this paper, we will elaborate the equivalence between TWR and column-based ZCA whitening, that is, TWR can remove the correlation in image column. However, the correlation within the local block of the image is often much greater than that within the column. From the perspective of removing the correlation within the local block of the image, this paper proposes two improved TWR methods: reshaped-based TWR (RTWR) and patch-based TWR (PTWR). RTWR firstly reshapes an image to form a new matrix of which each column vector corresponds to the sub-block of the original image, and then performs the TWR on the reshaped matrix. In PTWR method, TWR is directly applied to each sub-block of the image. The

experimental results on ORL, CMU PIE and AR face datasets show that RTWR and PTWR are more beneficial to improving the subsequent classification performance than TWR.

Key words: two-dimensional whitening reconstruction (TWR); ZCA-whitening; sub-image; image processing; correlation

引 言

白化作为一种数据预处理方法,由于可以在标准化数据的同时去除输入数据各属性间的相关性^[1],从而从计算和统计的角度简化多元数据分析的复杂度^[2],因此在图像处理^[3]、语音处理^[4]、三维物体识别^[5]等方面得到广泛应用。

白化,就是一种将 d 维随机向量 x 经过白化矩阵 W 变换到新向量 y 的线性转换过程^[2](即 $y = Wx$),其核心是寻找白化矩阵 W ,使得经过 W 变换后各属性间不相关且各属性的方差都相等。在各种白化方法中,PCA 白化(Principle component analysis)^[6]和 ZCA 白化(Zero-phase component analysis)^[7]是两种最常见的方法。PCA 白化是在 PCA^[8]的基础上做了尺度上缩放,使得白化后各属性的方差为 1。设协方差矩阵为 Σ ,且其奇异值分解为 $\Sigma = U\Lambda U^T$,则 PCA 的白化矩阵可表示为 $W_{\text{PCA}} = \Lambda^{-1/2}U^T$ 。PCA 白化可以使 y 能够尽可能多地表示原始 x 中的信息;而 ZCA 白化则是将 PCA 白化的结果经过 U 旋转回到原始空间的操作,其白化矩阵为 $W_{\text{ZCA}} = U\Lambda^{-1/2}U^T$,因此 ZCA 白化后的结果更接近原始数据。值得说明的是,由于 PCA 白化和 ZCA 白化均是在 PCA 基础上进行的转换,因此在求解白化矩阵时通常需要使用整个已知数据集,并且该数据集中的所有样本均共享一个白化矩阵 W 。

白化在图像处理方面有着广泛的应用,例如,文献[3]使用 ZCA 白化去除车牌图像邻域像素间的冗余信息;文献[9]使用 ZCA 白化作为深度神经网络输入前的预处理步骤;文献[10]则说明了白化操作有助于提高传统机器学习算法对光照、噪声等鲁棒性。Shi 等^[11]提出了一种二维白化重构(Two-dimensional whitening reconstruction, TWR)算法。作为一种图像预处理方法,不同于 PCA 白化和 ZCA 白化方法, TWR 白化矩阵的求解只依赖于要进行白化的单张图像而不需要使用其他图像。也就是说,每个图像在进行 TWR 白化变换时都对应各自专属的白化矩阵。文献[11]在人脸识别上的实验结果验证了经过 TWR 处理后的图像能够有效提高 PCA 及其变体算法的分类和聚类效果。

本文将阐述 TWR 等价于基于列的 ZCA 白化。亦即,对单张图像进行 TWR 操作等价于以图像列向量作为样本所进行的 ZCA 白化操作。由于 ZCA 白化能有效去除随机变量各属性间的相关性,因此当把图像的每一列作为随机变量时, ZCA 白化就相当于去除图像列内各属性的相关性。考虑到图像局部块内的冗余信息要远大于列内的冗余信息,基于上述等价性的发现,提出了重组的 TWR(Reshaped TWR, RTWR)算法。RTWR 首先将图像进行重新组合,使得重组后矩阵的每一列对应着原始图像的一个局部分块。随后基于这个重组后的矩阵做 TWR 操作以期去除图像局部块内的相关性。

另一方面,由于在图像处理中局部信息相对于全局信息而言对光照变换、目标遮挡等具有更好的鲁棒性^[12-13],因此有效获取局部特征成为必要。子图像方法^[14-15]是一种非常有效的获取图像局部特征的方法。该方法通过对原始图像进行划分,然后分别在每个子图像中提取特征来获取局部特征。鉴于子图像方法的成功,从局部的角度出发,提出了基于块的 TWR(Patch TWR, PTWR)算法。不同于 RTWR 对图像重组后做 TWR, PTWR 对每个分块分别进行 TWR,然后把每块的 TWR 结果按位置拼接成新的矩阵,最后再对新的矩阵进行降维或分类。为了验证 RTWR 和 PTWR 算法的有效性,在 ORL、CMU PIE 和 AR 人脸数据集上进行实验。实验结果表明 RTWR 和 PTWR 均优于 TWR,尤其是

PTWR算法,对人脸尺度变换、角度变换、光照及遮挡变换均具有非常好的鲁棒性,并且在训练样本较少的情况下这种优势更为突出。此外,对PTWR算法中的参数选择进行了分析。

1 TWR与ZCA白化的等价性

1.1 二维白化重构算法

TWR^[11]是针对单张图像进行的白化方法。它对图像进行白化的同时还可以使图像的像素分布接近高斯分布^[11]。文献[11]的实验结果表明,图片经过TWR预处理后再进行PCA及其变体算法的降维操作能够显著提升后续的分类和聚类效果,并且对光照变换具有较强的鲁棒性。

给定一张 p 行 n 列的图像 $X=[x_1, x_2, \dots, x_n] \in \mathbf{R}^{p \times n}$,其中, $x_i(i=1, 2, \dots, n)$ 是图像 X 的列向量。TWR首先将图像的每一列看作一个处理单元进行列内的去均值化操作^[16],得到列内中心化后的图像 $\tilde{X}=[\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n]$,然后对 \tilde{X} 进行奇异值分解^[17],即 $\tilde{X}=UDV^T$,则图像 \tilde{X} 经过TWR后的结果可表示为

$$Y_{\text{TWR}} = \sqrt{n} U(:, 1:m) V(:, 1:m)^T \quad (1)$$

式中 m 表示要保留的主分量的个数。

1.2 TWR和ZCA白化的等价关系

本节将说明对图像的TWR操作等价于基于图像列向量的ZCA白化。

同1.1节,给定图像 X ,经过列中心化处理后得到的图像表示为 $\tilde{X}=[\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n] \in \mathbf{R}^{p \times n}$,其奇异值分解为

$$\tilde{X} = UDV^T \quad (2)$$

如果将 \tilde{X} 的每一列 \tilde{x}_i 看作一个样本,那么 \tilde{X} 就是一个垂直中心化的列数据矩阵^[16],因此基于所有 $\tilde{x}_i(i=1, 2, \dots, n)$ 的类协方差矩阵可以表示为

$$\Sigma = \frac{1}{n} \sum_{i=1}^n \tilde{x}_i \tilde{x}_i^T = \frac{1}{n} \tilde{X} \tilde{X}^T \quad (3)$$

将式(2)代入式(3),可以得到 $\Sigma = UD^2U^T/n$ 。因此PCA白化矩阵可表示为 $W_{\text{PCA}} = \Sigma^{-1/2}U^T = \sqrt{n} D^{-1}U^T$,而ZCA白化矩阵则表示为 $W_{\text{ZCA}} = UW_{\text{PCA}} = \sqrt{n} UD^{-1}U^T$ 。在实际应用中,为了降低随机噪声的影响,ZCA白化通常只选择 D 中前 m 个非零奇异值对应的主成分方向,此时,ZCA白化矩阵为

$$W_{\text{ZCA}} = \sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T \quad (4)$$

式中: $\tilde{U} = U(:, 1:m)$, $\tilde{D} = \text{diag}\{D(1,1), \dots, D(m,m)\}$ 。因此, \tilde{X} 所有列 $\tilde{x}_i(i=1, 2, \dots, n)$ 的ZCA白化结果则可表示为

$$Y_{\text{ZCA}} = \left[\sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T \tilde{x}_1, \sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T \tilde{x}_2, \dots, \sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T \tilde{x}_n \right] = \sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T \tilde{X} = \sqrt{n} \tilde{U} \tilde{D}^{-1} \tilde{U}^T U D V^T = \sqrt{n} U(:, 1:m) V(:, 1:m)^T \quad (5)$$

由式(1)和式(5)可得: $Y_{\text{ZCA}} = Y_{\text{TWR}}$ 。因此,TWR等价于将图像的每一列作为一个样本进行ZCA白化。

2 RTWR和PTWR

2.1 RTWR

根据1.2节的结论可知,对图像 X 整体进行TWR预处理等价于将图像 X 的每一列作为一个样本进

行ZCA白化的操作,这就是说TWR能有效去除图像列内的相关性。因为图像结构信息的存在,局部块内的冗余信息要远多于列内的冗余信息,因此从去除冗余信息的角度出发,提出了RTWR方法。RTWR首先对人脸图像进行重新组合,使得重组后的矩阵的每一列对应着原始图像的一个局部块,然后针对新的重组矩阵进行TWR操作,最后将每一列还原成局部块,并对所有块进行重组。具体的实现步骤如下:

(1) 给定一张图像 $X \in \mathbb{R}^{p \times n}$, 首先不重叠地将图像分成若干个大小为 $a \times b$ 的矩形块 $X_{ij} (i \in \{1, 2, \dots, p/a\}, j \in \{1, 2, \dots, n/b\})$ (图1(a)将 16×16 的图像划分成16个大小为 4×4 的矩形块)。

(2) 将每个分块 X_{ij} 拉成一个列向量 $x_k (k \in \{1, 2, \dots, \tilde{n}\}, \tilde{n} = p*n/\tilde{p}, \tilde{p} = a*b)$, 并将它们拼接成新的矩阵 $X_{\text{new}} = [x_1, x_2, \dots, x_{\tilde{n}}] \in \mathbb{R}^{\tilde{p} \times \tilde{n}}$ (图1(b))。

(3) 对 X_{new} 进行列内中心化。令 $Z = X_{\text{new}}^T = [z_1, z_2, \dots, z_{\tilde{p}}]$, 首先计算 Z 的列向量均值向量 $\mu = \frac{1}{\tilde{p}} \sum_{j=1}^{\tilde{p}} z_j$, 然后用 Z 的每一列减去均值向量 μ 得到 $Z_h = [z_1 - \mu, z_2 - \mu, \dots, z_{\tilde{p}} - \mu]$, 进而得到垂直中心化的列数据矩阵^[16] $\tilde{X}_{\text{new}} = Z_h^T$ 。

(4) 对 \tilde{X}_{new} 进行SVD分解, 得 $\tilde{X}_{\text{new}} = UDV^T$, 且只保留最大的前 m 个非零奇异值所对应的特征向量, 因此得到 $Y = \sqrt{n} U(:, 1:m) V(:, 1:m)^T$ 。图1(c)所示为 Y 的可视化结果。

(5) 分别将 Y 中每一列 $y_k (k \in \{1, 2, \dots, \tilde{n}\})$ 重组为 $a \times b$ 大小的块, 并将它们按原始块位置重新组成 $p \times n$ 大小的矩阵, 得到RTWR的结果(图1(d))。

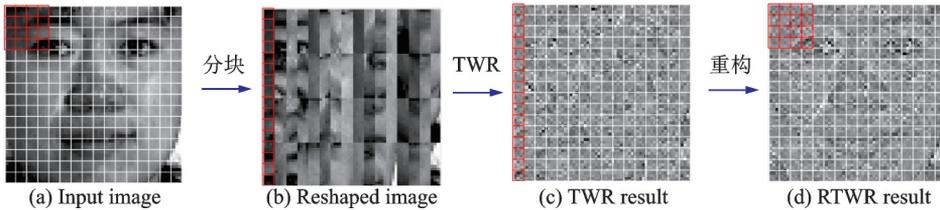


图1 RTWR具体操作流程(图像来自CMU PIE)

Fig.1 Operation process of RTWR (images from CMU PIE)

值得说明的是:①当分块大小和原始图像大小一致时,经过步骤1和2,整个图像会重组成一个列向量,即 X_{new} 为列向量;由于步骤3中所获得的 \tilde{X}_{new} 只是对 X_{new} 列内元素进行了去均值的处理,仍为列向量,因此步骤4对 X_{new} 白化处理后结果是不变的,即 $Y = X_{\text{new}}$, 所以当 Y 还原为原始图像大小的图像后该重构图像和原图像只是差了一个均值。也就是说,当分块大小和原始图像大小一致时,RTWR只是对图像元素做了一个去均值的处理。②当每个分块对应原图像的每一列时,步骤2的重组操作并不会改变原始图像的结构,即 $X_{\text{new}} = X$ 。因此当对 X_{new} 做TWR时就等同于对 X 做TWR,所以,当图像的每个分块对应于原图像的每一列时,RTWR退化为TWR,也就是说TWR是RTWR的一个特例。

2.2 PTWR

RTWR方法通过将图像的每个子块拉成列向量实现对原始图像的重组,然而这样的操作在一定程度上会破坏图像的局部信息。为了充分利用图像的局部信息,同时鉴于子图像方法在图像去噪、小目标检测^[18]等研究领域的成功应用,提出了PTWR算法。不同于RTWR将图像分块拉成列向量拼接后

整体进行TWR操作,PTWR是在每个子块内分别做TWR,以期保留更多的局部信息。

PTWR的具体流程如下:给定一张图像 $X \in \mathbb{R}^{p \times n}$,首先不重叠地将图像分成若干个大小 $a \times b$ 的矩形块 $X_{ij}(i \in \{1, 2, \dots, p/a\}, j \in \{1, 2, \dots, n/b\})$;然后在每个分块上分别进行TWR操作,最后将每块的处理结果复位重新构成大小为 $p \times n$ 的矩阵,该矩阵即为 X 经过PTWR后的结果。图2展示了图像经过PTWR的处理过程。在图2(a)中,16×16大小的人脸图像被划分成16个4×4的矩形块,然后每个4×4的块分别做TWR操作(图2(c)),最后所有TWR处理后的块按位重组成为16×16的矩阵(图2(d))。值得说明的是,当把整张图像看作一个分块时,此时PTWR就退化为TWR。也就是说,TWR也是PTWR的一个特例。

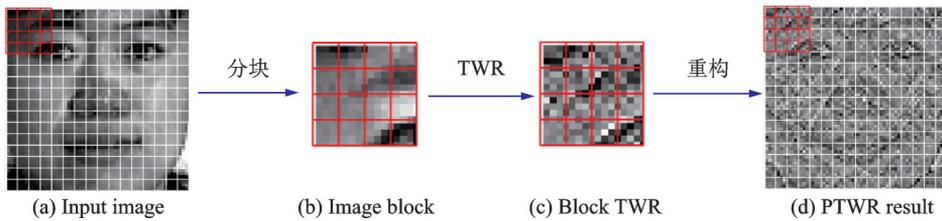


图2 PTWR具体操作流程

Fig.2 Operation process of PTWR

图3分别给出了原始图像和经过RTWR与PTWR处理后像素点的分布情况。从图中可以看出,严重受光线影响的图像其像素点的分布主要集中在低值区域(图3(a));但图像经过RTWR处理后

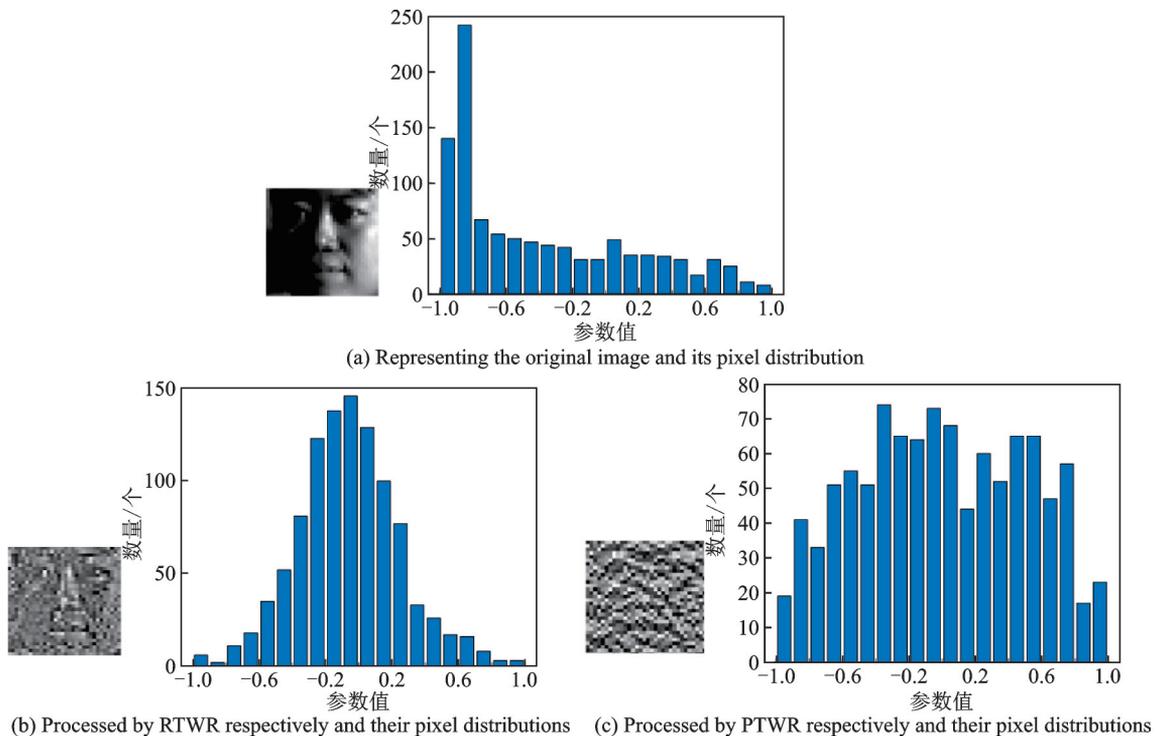


图3 像素分布对比

Fig.3 Comparison of pixel distribution

(图 3(b)),像素分布接近高斯分布;而图像经过 PTWR 处理后(图 3(c)),像素分布更趋于平均化。因此,经过 RTWR 和 PTWR 预处理的图像,其像素分布更接近高斯分布或均匀分布。

图 4 给出了一组图像分别经过 3 种预处理后结果。从图 4 可以很明显地看出,RTWR($a \times b = 8 \times 4, m = 28$)和 PTWR($a \times b = 8 \times 4, m = 2$)预处理后的图像轮廓比 TWR 更清晰,同时结构化特征也更突出。

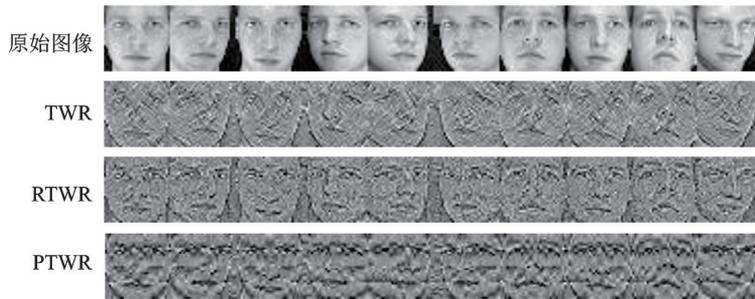


图 4 TWR、RTWR 和 PTWR 处理效果对比(图片来自 ORL)

Fig.4 Processing results of TWR, RTWR and PTWR (images from ORL)

3 实验结果与分析

为了验证 RTWR 和 PTWR 方法的有效性,本文在 ORL、CMU PIE 和 AR 人脸数据集上对 RTWR、PTWR 和 TWR 进行对比实验。其中,ORL 数据集主要用于测试算法对人脸轻微角度旋转、图像尺度缩放的鲁棒性;CMU PIE 数据集主要用于测试算法对光照变换的鲁棒性;而 AR 数据集则是用于测试算法对各种面部遮挡的鲁棒性。另外,还分析了算法中图像分块大小 $a \times b$ 及主分量的个数 (m) 对算法性能的影响。

3.1 数据集

(1) ORL 人脸数据集^[19]

ORL 人脸数据集包括 40 个人的共 400 张图像(每人 10 张图像),这些图像包含人脸转动角度的变换(不超过 20°)、面部表情和配饰(眼镜等)的变换以及图像尺度的变换(放缩比例约为 1.2)。原始图像大小为 112×92 ,本文在使用时将其缩放为 32×32 大小。

(2) CMU PIE 人脸数据集

CMU PIE 人脸数据集包括 68 个人在 13 种姿势变换、43 种光照变换以及 4 种表情变换下的 41 368 张图像。本文中所使用的图像为文献[20]处理后的子集(C27)。该子集共包括 3 329 张正面图像,每类人脸有 49 张图像(除第 38 类是 46 张图像),每张图像大小为 64×64 。

(3) AR 人脸数据集^[21]

AR 数据集包含 126 个人(70 个男性和 56 个女性)的 4 000 余张包括不同的表情、光照和遮挡变换的彩色人脸正面照。本文中,选用其中的一个黑白图像子集,该子集包括 100 个人(50 个男性和 50 个女性,每人 26 张)的 2 600 张图像,每张图像大小为 64×48 。每人的 26 张图像分别在两个场景拍摄得到,且每个场景有 13 张图像。对于每个场景的 13 张图像而言,1st 为中立图像、2nd~4th 为表情变换图像、5th~7th 张为光照变换图像、8th~10th 张为眼部遮挡图像(眼镜),而 11th~13th 张为嘴部遮挡图像(围巾),如图 5 所示。

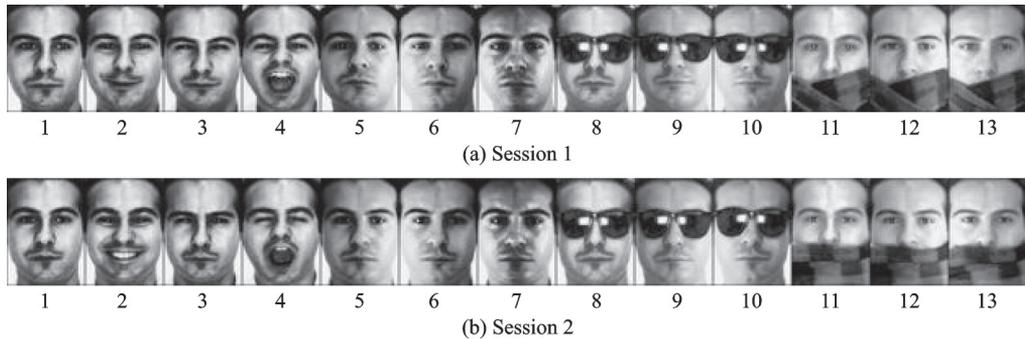


图5 AR数据库样例

Fig.5 AR database samples

3.2 实验设置与结果

在实验中,首先分别使用RTWR、PTWR和TWR三种方法对每张图像进行预处理,然后使用PCA或2DPCA^[22]对处理后的结果进行降维,最后使用最近邻分类器进行分类(欧式距离),并将分类结果作为分析的依据。现对实验设置进行说明。对于ORL数据集,从每类图像中随机选取 q ($q=2,4,6,8$)个样本组成训练集,其余样本用于测试。在实验中,TWR采用与文献[11]相同的参数设置,即 m 取28;RTWR和PTWR的图像分块大小均为 8×4 , m 分别取28和2。实验重复20次,最后给出平均识别率。对于CMU PIE数据库,从每类图像中随机选取 q ($q=2,4,6,8$)个样本组成训练集,其余样本用于测试。TWR的 m 取40;RTWR和PTWR图像分块大小分别为 32×2 和 16×4 , m 分别取40和2。实验重复20次,最后给出平均识别率和标准差。而对于AR数据库,每类图像第一个场景的前7张(表情和光照变换)图像组成训练集,两个不同场景下的眼部遮挡和嘴部遮挡分成4个测试集分别用于测试。在TWR中, $m=44$;RTWR和PTWR的图像分块大小取 12×6 , m 分别取44和5。

图6和表1、2分别给出了ORL、CMU PIE和AR数据集经过TWR、RTWR和PTWR预处理后进行PCA和2DPCA降维后的分类结果。从整体的结果上可以看出,RTWR和PTWR均优于TWR。具体而言:(1)对于有轻微角度变换和尺度变换的ORL数据集,RTWR比TWR表现稍差,而PTWR则明显优于TWR,尤其是当训练样本较少时(如 $q=2$ 或4时),这种优势更为明显。(2)对于包含各种光线变换的CMU PIE数据集,RTWR并没有获得比TWR更好的性能。在使用PCA进行降维时RTWR略优于TWR,而使用2DPCA时RTWR表现略差于TWR,整体上两者性能相当;而PTWR却给出相当令人满意的结果。当训练样本数为2,使用2DPCA降维时,PTWR比TWR能高出5.7%;同时,PTWR有比TWR更低的方差,这也说明PTWR比TWR更具有稳定性。(3)对于有眼部遮挡和嘴部遮挡的AR数据集,RTWR和PTWR均明显比TWR具有更好的性能。相对于TWR而言,RTWR和PTWR最小的提升幅度分别为3.0%和0.7%;而最大的提升幅度则为6.0%和12.4%;(4)比较表2中RTWR与PTWR的实验结果发现,RTWR在处理眼部遮挡时比PTWR更具有优势,而PTWR则对嘴部遮挡等影响因素表现出很好的适应性。

3.3 参数分析

3.3.1 块大小对性能的影响

在RTWR和PTWR方法中,块大小($a \times b$)是一个非常重要参数。图7给出了采用不同块大小($a=b$)时RTWR和PTWR预处理后的结果。从图7可以看出:(1)对于RTWR,无论采用怎样的分块

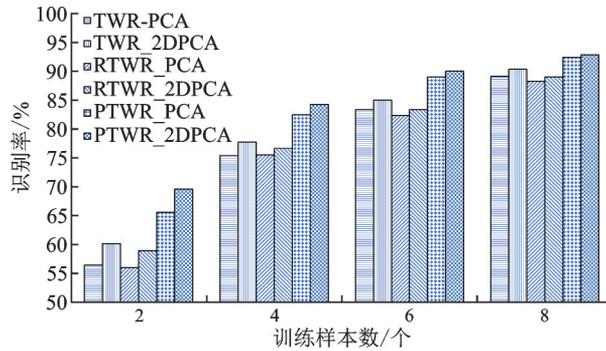


图6 ORL数据集实验结果

Fig.6 Experimental results on ORL dataset

表1 CMU PIE数据集实验结果

Table 1 Experimental results on CMU PIE

数据集	$q=2$	$q=4$	$q=6$	$q=8$
TWR+PCA	69.6 ± 2.0	86.3 ± 1.2	91.3 ± 0.6	93.4 ± 0.5
TWR+2DPCA	75.2 ± 2.0	87.4 ± 1.2	91.7 ± 0.6	93.6 ± 0.6
RTWR+PCA	72.3 ± 2.1	87.1 ± 1.0	91.7 ± 0.5	93.6 ± 0.6
RTWR+2DPCA	74.6 ± 1.9	86.5 ± 1.2	90.8 ± 0.7	92.7 ± 0.6
PTWR+PCA	79.2 ± 1.5	90.4 ± 1.1	93.5 ± 0.5	94.8 ± 0.4
PTWR+2DPCA	80.9 ± 1.6	90.8 ± 1.0	93.7 ± 0.5	94.8 ± 0.4

表2 AR数据集实验结果

Table 2 Experimental results on AR

数据集	眼部遮挡(S1)	嘴部遮挡(S1)	眼部遮挡(S2)	嘴部遮挡(S2)
TWR+PCA	83.0	81.7	63.3	59.3
TWR+2DPCA	85.0	83.0	64.3	61.3
RTWR+PCA	87.0	85.0	67.3	64.0
RTWR+2DPCA	88.0	86.3	67.7	67.3
PTWR+PCA	86.7	92.3	64.0	71.7
PTWR+2DPCA	87.3	91.7	65.0	70.3

大小都能清晰看出图像的轮廓。随着分块的增大,RTWR重组后的轮廓不仅越来越清晰而且也越来越接近于原始图像。2.1节的分析表明,当 $a=32$ 时(此时分块大小与原图像相同),RTWR重组的图像和原图像只相差一个灰度差;(2)对于PTWR,当分块大小较小时(如 $a=2$),很难看出图像的轮廓;而随着 a 的增加,图像轮廓逐渐清晰。当 $a=32$ 时(此时分块大小小于原图像相同),PTWR退化为TWR,此时轮廓最为清晰。关于RTWR重构结果和PTWR的差别,可以做如下分析:RTWR是以去除局部块内相关性的角度出发,首先将图像进行了重组(使得重组后的每一列对应着原始图像的一个局部块)然后TWR直接执行在重组的图像上。虽然RTWR将图像进行了划分,但由于TWR是作用在整体重组图像上的,因此RTWR相对更侧重于获取图像整体信息,因此,图7(a)中人脸轮廓清晰可见;而PTWR则是从保持原始图像局部特征的角度出发,分别对每个分块做TWR操作,然后通过TWR获取每个分块

的信息。因此PTWR更侧重获取图像的局部信息。所以图7(b)中所呈现的PTWR重构后图像相邻区域间变换较大(分块为 2×2 时尤为突出)。鉴于RTWR与PTWR的上述差别,进一步分析了分块大小对两种方法性能的影响。

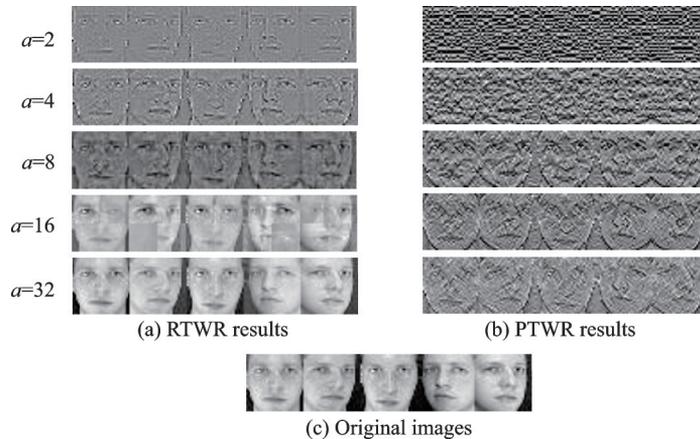


图7 不同分块的白化效果图

Fig.7 Renderings of different blocks

(1) 分块大小对RTWR的影响

在RTWR方法中,TWR保留主分量的个数(m 值)的选取与分块大小密切相关。对一张 $p \times n$ 大小的图像以 $a \times b$ 大小进行分块重组将得到大小为 $\tilde{p} \times \tilde{n}$ ($\tilde{n} = p \cdot n / \tilde{p}$, $\tilde{p} = a \cdot b$)的图像,那么 m 取值一定需要满足 $1 \leq m \leq \min(\tilde{n}, \tilde{p})$ 。而太大或太小的分块都会导致 \tilde{n} 或 \tilde{p} 很小,从而使得 m 最大取值很小,进而影响RTWR最后的性能。本文尽量保证选取的主成分的个数能与TWR方法一致,因此遵循分块重组后图像尺寸尽可能接近原始图像尺寸的原则对图像进行分块。如对于 64×64 的原始图像,采用 32×2 的分块,以使得重构后的图像仍然是 64×64 。然而重组成 64×64 的图像会存在多种分块情况(如 $4 \times 16, 16 \times 4$ 等),仍然无法选择采用哪种分块形式。鉴于集成学习方法往往能够弱化各个分类器的不足,对提高最终分类器的性能有着较好的适用性^[23-24],因此考虑采用集成的方法设计多分类器系统。即每种分块情况分别设计一个子分类器,最终将所有子分类器的结果进行简单的投票。在CMU PIE数据集上进行相关实验。实验中,从每类图像中随机选取 q ($q = 2, 4, 6, 8$)个样本用于训练,剩余的样本组成测试集。针对每组训练样本和测试样本,首先将图像按块 $a \times b$ 进行分块重组处理(如2.1节步骤1,2),然后对重组的图像进行RTWR操作(同TWR取相同的 m ,即 $m = 40$),最后使用PCA降维并用最近邻分类器进行分类。针对于每一种分块($a \times b = 1 \times 64, 2 \times 32, 4 \times 16, 8 \times 8, 16 \times 4, 32 \times 2, 64 \times 1$)均做上述处理,因此对于每个测试样本均能获得多个分类结果,最后将所有分类结果进行投票,并将此投票结果作为该测试样本的类别。重复实验20次取平均识别率。

实验结果如表3所示。从表3可以看出,①分块大小对RTWR的影响的确很大:当分块为 1×64 时识别效果最差,而在 32×2 时效果最好,两者最大相差6%;②集成分类器所得结果与单一分块相比在识别率和稳定性方面均表现良好,集成后的结果仅次于所有分块中的最优性能且总体上相差并不大(最大差距不到1%,最小差距只有0.1%)。因此在确定重构图像大小的前提下,采用这种集成方式来弱化具体分块选择是可行的。

(2) 分块大小对PTWR的影响

PTWR方法首先将图像进行分块,然后分别对每块做TWR处理,因此PTWR属于子图像方法^[15]。

对于子图像方法,分块大小是一个非常重要的参数,它严重影响着最后的性能。本文将讨论分块大小对PTWR的影响。分块大小不宜过小或过大,过小的分块会导致过多的考虑局部信息而忽略全局信息的重要性,从而可能会造成全局信息的丢失;而过大的分块则会忽略局部信息的重要性。为简单起见,仅考虑正方形分块的情形,即 $a=b$ 。在CMU PIE数据集上进行相关实验。实验中,从每类图像中随机选取 $q(q=2, 4, 6, 8)$ 个样本用于训练,剩余的样本组成测试集。针对每组训练样本和测试样本,首先将图像按块 $a \times$

$a(a=2, 4, 8, 16, 32, 64)$ 进行分块处理,然后对分块后的图像进行PTWR操作(对应分块情况下的 m 分别为1, 2, 4, 12, 28, 40),最后使用PCA降维并分类。重复实验20次取平均识别率作为最后的结果。

图8展示了在CMU PIE数据集上分块大小对PTWR识别效果的影响。从实验结果中可以看出:①当分块大小为 8×8 (最接近原图像大小的1/100)时PTWR取得了最好的识别效果,这也验证了文献[25-26]的经验结论,即子图像大小在原图像大小的1/100附近时,子图像方法往往能够获取较好效果。②整体上说,当分块大小为 2×2 时识别率较低,随着分块大小的增加,识别性能逐渐增加;当分块为 8×8 时达到最好的效果;而后随着分块的增加识别率又逐渐减低。值得说明的是,虽然PTWR的分块对性能有较大影响,但从整体上来看,PTWR在大多数分块情况下,性能都优于TWR。③在每类训练样本较小时($q=2$),PTWR算法对应的实线和TWR算法对应的虚线所夹面积相差较大;而随着训练样本数的增加,所夹面积逐渐变小,这也说明了训练样本较少的情况下PTWR算法相对于TWR更具有优势。值得说明的是,当分块大小和原始图像大小一致时,PTWR就转换为TWR,即TWR为PTWR的一个特例。

3.3.2 主分量个数(m)对性能的影响

(1) 主分量个数对RTWR的影响

从表3的数据可以看出,集成学习能够有效弱化分块大小对RTWR识别率的影响,进而解决分块大小难以选择的问题。本节将在集成分块的基础上,讨论 m 值对RTWR的影响。以CMU PIE数据集为例,以RTWR重组后的图像大小和原图像相同为分块原则,考察了 m 值从4开始以4为步长变化至64对集成RTWR(所有可能分块的集成)的影响。图9(a)给出了 m 对集成RTWR和TWR两种方法的性能影响,其中 q 表示每类有 q 个样本做训练,其他样本做测试。从图9可以看出:①随着 m 的增加,两种方法的识别率都是先上升后逐渐趋于平稳再下降的走势。这主要是因为当 m 较小时只有很少的

表3 CMU PIE数据集RTWR分块大小实验结果

Table 3 Experimental results of patch size analysis of RTWR on CMU PIE %

$a \times b$	$q=2$	$q=4$	$q=6$	$q=8$
1×64	66.1 ± 2.2	84.3 ± 1.2	90.2 ± 0.8	92.6 ± 0.6
2×32	69.8 ± 2.4	85.5 ± 1.2	90.4 ± 0.6	92.6 ± 0.7
4×16	70.1 ± 2.2	85.8 ± 1.1	90.7 ± 0.7	92.7 ± 0.7
8×8	70.4 ± 2.5	86.1 ± 1.2	91.0 ± 0.6	92.9 ± 0.6
16×4	71.4 ± 2.3	86.8 ± 1.2	91.5 ± 0.6	93.3 ± 0.6
32×2	72.3 ± 2.1	87.1 ± 1.0	91.7 ± 0.5	93.6 ± 0.6
64×1	69.6 ± 2.0	86.3 ± 1.2	91.3 ± 0.6	93.4 ± 0.5
集成	71.4 ± 2.2	86.9 ± 1.1	91.6 ± 0.6	93.5 ± 0.5

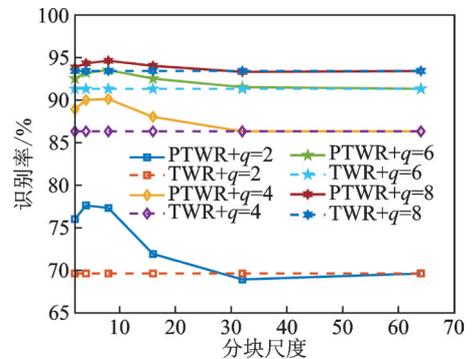


图8 分块大小对PTWR的影响(CMU PIE)
Fig.8 Effect of patch size on PTWR (CMU PIE)

有用信息被使用,而大量的有用信息丢失,所以性能较低;反之,当 m 较大时,一些噪声信息将会被利用,从而导致识别性能降低。②随着训练样本数目 q 的增加, m 对集成的RTWR和TWR的影响逐渐减小,尤其是对集成的RTWR影响更小。图9(a)显示,除 $q=2$ 外,两种方法的 m 在块尺寸的50%~80%之间时均取得较好的识别率。

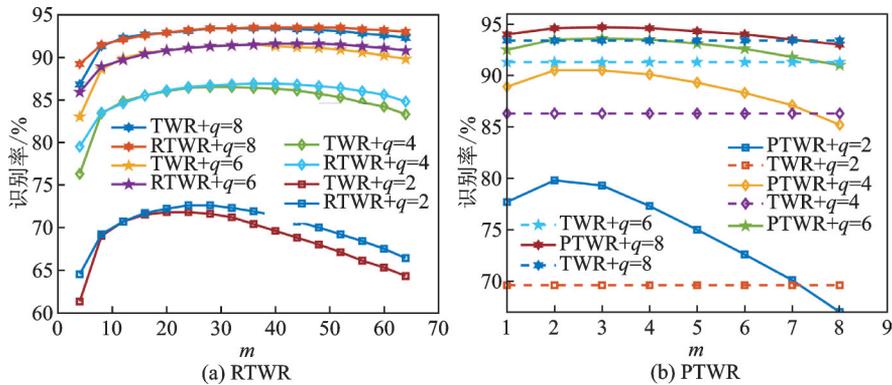


图9 参数 m 对识别率的影响(CMU PIE)

Fig.9 Effect of parameter m on recognition rate (CMU PIE)

(2) 主分量个数对PTWR的影响

图8的实验结果显示,在CMU PIE数据集上,当分块大小为 8×8 时PTWR获得了最好的识别性能,因此,将在 8×8 的分块情况下考察 m 对PTWR识别效果的影响。实验仍然以CMU PIE为例,主分量个数 m 从1开始变化至8(对于 8×8 的分块,最大的 m 为8),得到不同 m 值对应的PTWR的识别结果。为了进行性能比较,在图9(b)中也给出了TWR在 $m=40$ 时的性能。从实验结果可以看出,当每块保留主成分信息过少时($m=1$),因为大量的信息被丢弃,所以对应的识别效果并不理想;而当保留所有主成分信息时($m=8$),由于可能使用了包含随机噪声在内的信息,从而影响了后续的分类操作。在 8×8 的分块情况下, $m=2$ 时达到最优的识别效果,并且在几乎所有取值下(除 $m=8$)PTWR均获得了比TWR更好的性能。

4 结束语

本文首先证明了基于单张图像的TWR预处理与对单张图像按列做ZCA白化操作的等价性关系;然后在两者等价性的基础上,从去除图像局部块内相关性的角度提出了基于重组的RTWR算法;从尽可能提取原始图像结构信息和局部特征的角度,提出了基于块的PTWR算法。在ORL、CMU PIE和AR数据集上做了对比实验,实验结果表明,本文提出的RTWR和PTWR相对于TWR对表情、光照、遮挡等影响因素具有更好的鲁棒性。另外,还讨论了子块大小以及主分量个数对算法性能的影响。

参考文献:

- [1] PAN Xingang, ZHAN Xiaohang, SHI Jianping, et al. Switchable whitening for deep representation learning[C]//Proceedings of the IEEE International Conference on Computer Vision. [S.l.]: IEEE, 2019: 1863-1871.
- [2] KESSY A, LEWIN A, STRIMMER K. Optimal whitening and decorrelation[J]. The American Statistician, 2018, 72(4): 309-323.

- [3] SOON F C, KHAW H Y, CHUAH J H, et al. Vehicle logo recognition using whitening transformation and deep learning[J]. *Signal, Image and Video Processing*, 2019, 13(1): 111-119.
- [4] SHON S, MUN S, KO H. Recursive whitening transformation for speaker recognition on language mismatched condition[J]. *arXiv eprints*, 2017. DOI: arXiv: 1708.01232.
- [5] ZHENG Jun, WEI Haiyong. Three-dimensional object recognition and posture calculations based on the whitening transformation and curvature characteristics[J]. *Journal of Tsinghua University(Science and Technology)*, 2016, 56(10): 1025-1030.
- [6] JÉGOU H, CHUM O. Negative evidences and co-occurrences in image retrieval: The benefit of PCA and whitening[C]// *Proceedings of European Conference on Computer Vision*. [S.l.]: Springer, 2012: 774-787.
- [7] BELL A J, SEJNOWSKI T J. The “independent components” of natural scenes are edge filters[J]. *Vision Research*, 1997, 37(23): 3327-3338.
- [8] WOLD S, ESBENSEN K, GELADI P. Principal component analysis[J]. *Chemometrics and Intelligent Laboratory Systems*, 1987, 2(1/2/3): 37-52.
- [9] TUN N M, GAVRILOV A I, PAING P P. Human face recognition using combination of ZCA feature extraction method and deep neural network[C]// *Proceedings of 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. [S.l.]: IEEE, 2019: 1-5.
- [10] VINAY A, KAILASAM S, BHARADWAJ A, et al. Effective utilization of whitening for person identification[C]// *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM)*. [S.l.]: Amity University Rajasthan, 2019: 26-28.
- [11] SHI Xiaoshuang, GUO Zhenhua, NIE Feiping, et al. Two-dimensional whitening reconstruction for enhancing robustness of principal component analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 38(10): 2130-2136.
- [12] 吴辉, 王思玮. 基于图像灰度变化的局部特征研究综述[C]//第 37 届中国控制会议论文集. 武汉:[s.n.], 2018: 816-821.
WU Hui, WANG Siwei. A review of local features based on the change of gray value of image pixel[C]// *Proceedings of the 37th Chinese Control Conference*. Wuhan:[s.n.], 2018: 816-821.
- [13] 陈凯, 陈秀宏. 基于 ELM 的局部空间信息的模糊 C 均值聚类图像分割算法 [J]. *数据采集与处理*, 2019, 34(1): 100-116.
CHEN Kai, CHEN Xiuhong. Fuzzy C-means clustering image segmentation algorithm with local spatial information based on ELM[J]. *Journal of Data Acquisition and Processing*, 2019, 34(1): 100-116.
- [14] GOTTUMUKKAL R, ASARI V K. An improved face recognition technique based on modular PCA approach[J]. *Pattern Recognition Letters*, 2004, 25(4): 429-436.
- [15] 谢飞, 郭宇, 张红蕾, 等. 基于图和子图同构算法的制造特征识别方法[J]. *南京航空航天大学学报*, 2018, 50(3): 390-396.
XIE Fei, GUO Yu, ZHANG Honglei, et al. Manufacturing feature recognition based on graph and subgraph isomorphism algorithm[J]. *Journal of Nanjing University of Aeronautics & Astronautics*, 2018, 50(3): 390-396.
- [16] YANG J, ZHANG D, YANG J Y. Is ICA significantly better than PCA for face recognition?[C]// *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05)*. [S.l.]: IEEE, 2005: 198-203.
- [17] KLEMA V, LAUB A. The singular value decomposition: Its computation and some applications[J]. *IEEE Transactions on Automatic Control*, 1980, 25(2): 164-176.
- [18] JI Zexuan, XIA Yong, CHEN Qiang, et al. Fuzzy C-means clustering with weighted image patch for image segmentation[J]. *Applied Soft Computing*, 2012, 12(6): 1659-1667.
- [19] SAMARIA F S, HARTER A C. Parameterisation of a stochastic model for human face identification[C]// *Proceedings of 1994 IEEE Workshop on Applications of Computer Vision*. [S.l.]: IEEE, 1994: 138-142.
- [20] SIM T, BAKER S, BSAT M. The CMU pose, illumination, and expression (PIE) database[C]// *Proceedings of the Fifth*

- IEEE International Conference on Automatic Face Gesture Recognition. [S.l.]: IEEE, 2002: 53-58.
- [21] MARTINEZ A, BENAVENTE R. The AR face database[R]. Barcelona, Spain: The Computer Vision Center, 1998.
- [22] YANG Jian, ZHANG David, FRANGI A F, et al. Two-dimensional PCA: A new approach to appearance-based face representation and recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(1): 131-137.
- [23] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2014: 1867-1874.
- [24] DIETTERICH T G. Ensemble methods in machine learning[C]//Proceedings of International Workshop on Multiple Classifier Systems. [S.l.]: [s.n.], 2000: 1-15.
- [25] 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.
- [26] 洪泉, 陈松灿, 倪雪蕾. 子模式典型相关分析及其在人脸识别中的应用[J]. 自动化学报, 2008, 34(1): 21-30.
HONG Quan, CHEN Songcan, NI Xuelei. Sub-pattern canonical correlation analysis with application in face recognition[J]. ACTA Automatica Sinica, 2008, 34(1): 21-30.

作者简介:



田甲略(1998-),男,硕士研究生,研究方向:模式识别, E-mail: jltian@nuaa.edu.cn.



朱玉莲(1978-),通信作者,女,副教授,硕士生导师,研究方向:图像处理、机器学习,人工智能, E-mail: lianyi_1999@nuaa.edu.cn.



陈飞玥(1995-),女,硕士研究生,研究方向:模式识别。



刘佳慧(1999-),女,硕士研究生,研究方向:图形学与图像处理技术。

(编辑:夏道家)