

文章编号:1001-9081(2007)07-1660-03

基于 $(2D)^2$ NMF 及其改进算法的人脸识别

高宏娟, 潘 晨

(宁夏大学 数学计算机学院, 银川 750021)

(juanzi_cherry1213@126.com)

摘 要:非负矩阵分解(NMF)是基于部分的特征提取方法,能够克服局部遮挡和光照问题,在图像识别任务中效果较好。然而传统算法中,NMF 提取的特征是非正交的,且二维图像常被向量化处理,不仅丢失一些结构信息,还导致了数据的高维,不利于提高识别精度和速度。利用图像矩阵取代传统的图像向量表示,提出新的 $(2D)^2$ NMF 方法提取二维图像特征,并通过特征正交化和图像变形等措施,改善了算法性能。人脸识别实验表明,上述措施能够有效提高识别的精度和速度。

关键词:人脸识别;非负矩阵分解;二维非负矩阵分解;对角化;正交化

中图分类号: TP391 **文献标志码:** A

Face recognition based on $(2D)^2$ NMF and its improvement

GAO Hong-juan, PAN Chen

(School of Mathematics & Computer Science, Ningxia University, Yinchuan Ningxia 750021, China)

Abstract: Non-negative matrix factorization (NMF) is an effective method for parts-based feature extraction, and it can deal with partial occlusion and some illumination problems. However, the bases learned via NMF are not orthogonal. Translating an image into a vector often loses the structure information of pixels and leads to a high dimensionality. By adopting the image matrix instead of image vector, $(2D)^2$ NMF method was presented to extract the 2D features of an image, so the performance of the method was improved by the NMF bases orthogonalization and image diagonalization. The experimental results show the improved performance compared with the traditional NMF and 2DNMF.

Key words: face recognition; Non-negative Matrix Factorization (NMF); 2DNMF; diagonalization; rothogonalization

0 引言

人脸识别的特征提取方法主要有两类,一是基于整体脸的方法,二是基于局部特征的方法。基于整体脸的方法如主成分分析(Principal Component Analysis, PCA),核主成分分析(Kernel Principal Component Analysis, KPCA)等,保留了人脸部件之间的拓扑关系与各部件本身的信息;而基于局部特征的方法如非负矩阵分解(Non-negative Matrix Factorization, NMF)等,则是通过提取局部信息来实现分类。应用中,整体脸方法对局部遮挡和光照变化问题敏感;而基于局部特征的方法则具有鲁棒性。较一些传统算法而言,NMF 具有实现上的简便性、分解形式和分解结果上的可解释性,以及占用存储空间少等诸多优点^[1]。人脸识别中的大规模人脸数据库需要通过矩阵形式进行有效处理,NMF 思想正好为此提供了一种新的途径。

非负矩阵分解虽然已经被成功地应用于人脸识别^[2],但是当图像规模较大时,图像向量化常常导致数据高维,使得 NMF 需要耗费大的计算代价。另外,NMF 提取的特征是非正交的,不便于设计后续分类器。本文在传统 NMF 方法的基础上做了几项改进:一是提出新的二维非负矩阵分解方法 $(2D)^2$ NMF;二是将图像镜像和对角化,以便提取图像结构信息;三是将非负矩阵分解得到的基矩阵正交化,以满足用欧氏距离度量样本相似性。利用人脸数据库的实验对比结果,证

明以上改进策略能够有效地提高人脸图像识别的正确率。

1 NMF 算法

给定一个大小为 $n \times m$ 的非负矩阵 V ,NMF 算法可以得到一个大小为 $n \times r$ 的非负矩阵 B 和一个大小为 $r \times m$ 的非负矩阵 C ,使得:

$$V \approx BC \quad (1)$$

式中, r 是参考维数,其值的选择应为 $(n+m)r < nm$,得出的 C 是 V 中数据的压缩形式。

NMF 不允许 B 和 C 中有负数元素。该约束使得在用部分表示整体的过程中,只有加法运算而没有减法运算。这个特性反映了由部分构成整体最直观的感受。

NMF 算法通过迭代方式分解各子矩阵。其中每次迭代 B 和 C 的新值由当前值乘以某个系数而得到,而这些系数取决于式(1)中的近似程度。按照这样的规则不断地迭代,可以确保 B 和 C 收敛到一个局部最优的矩阵分解。本文用到一种乘性迭代规则:

$$C_{ij} = C_{ij} \frac{(B^T V)_{ij}}{(B^T B C)_{ij}} \quad (2)$$

$$B_{ij} = B_{ij} \frac{(V C^T)_{ij}}{(B C C^T)_{ij}} \quad (3)$$

按照上述的规则不断地迭代,直到下式所示评价函数达

收稿日期:2007-01-22;修回日期:2007-03-28。

基金项目:国家自然科学基金资助项目(60663003);宁夏自然科学基金资助项目(NZ0610)。

作者简介:高宏娟(1980-),女,宁夏人,硕士研究生,主要研究方向:图形图像、多媒体技术;潘晨(1966-),男,浙江人,副教授,博士,主要研究方向:图像处理与分析、模式识别、机器学习。

到局部最小:

$$F = \sum_i \sum_j (V_{ij} - (BC)_{ij})^2 \quad (4)$$

该评价函数定义为 V 与 BC 间的欧式距离,并用它来评价两者近似程度。

2 (2D)²NMF 算法

传统子空间算法提取图像特征时,大都是将二维图像按像素扫描顺序排列为列向量,然后用一组基向量表示它们。这样使得二维图像变成了高维向量,丢失了图像结构信息,且图像尺寸越大,该向量维数越高,对系统性能带来不利影响。为了解决此类问题,杨键等^[3]提出 2DPCA 方法,张道强等^[4]提出了 2DNMF 方法。这些方法直接用二维图像矩阵方式来构建训练矩阵,通过对该矩阵进行行矩阵分解和列矩阵分解,提取图像特征。一般而言,二维方法能够大大减少图像矩阵维数,降低计算的时间花费,且考虑了图像结构信息,能提高分解精度。但是在实现过程中,文献[4]方法是串行的,需要更多的中间存储和运算过程。本文提出一种改进算法,可以并行对图像矩阵实施行和列分解,能够达到相同的效果。

2.1 图像矩阵行方向的非负矩阵分解

将大小为 $p \times q$ 的 m 幅图像用的矩阵 $X_{p \times qm} = [A_1, A_2, \dots, A_m]$ 来表示, A_k 代表 M 幅人脸图像中的一幅。类似于 2DNMF 算法,首先利用 NMF 分解找到一个大小为 $p \times d$ 的非负矩阵 L 和一个大小为 $d \times qm$ 的非负矩阵 H ,使得:

$$X_{p \times qm} \approx L_{p \times d} H_{d \times qm} \quad (5)$$

这里 d 是参考维数, L 是矩阵 X 在图像行方向上的分解得到的基矩阵, H 为系数矩阵。我们将计算基矩阵 L 的步骤称为行方向非负矩阵分解。如图 1(a) 示例说明。

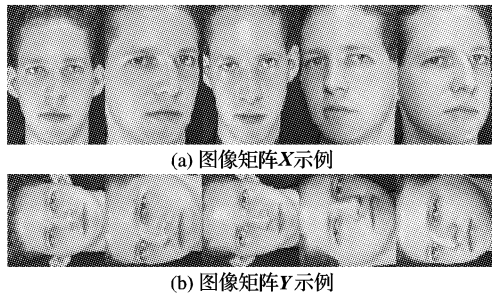


图1 图像矩阵 X 和 Y (对训练矩阵截取的一小部分)

文献[4]为了进一步提取列方向信息,继续对式(5)得到的 H 矩阵进行非负矩阵分解。而我们认为,无论行方向信息和列方向信息实际都存在于原图像中,应该直接利用原图像提取才最为简便。以下提出一种改进方法,直接对原图像进行简单处理来得到列方向信息。

2.2 图像矩阵列方向的非负矩阵分解

将大小为 $p \times q$ 的 m 幅图像用的矩阵 $Y_{q \times pm} = [B_1, B_2, \dots, B_m]$ 来表示,其中 $B_k = A_k^T$,是将原图像转置处理。类似于上节算法,利用 NMF 分解找到一个大小为 $q \times r$ 的非负矩阵 R 和一个大小为 $r \times pm$ 的非负矩阵 H ,使得:

$$Y_{q \times pm} \approx R_{q \times r} H_{r \times pm} \quad (6)$$

这里 r 是参考维数, R 是矩阵 X 在图像列方向上的分解得到的基矩阵, H 为系数矩阵。我们将计算基矩阵 R 的步骤称为列方向非负矩阵分解。如图 1(b) 示例说明。

矩阵 $X_{p \times qm}$ 和 $Y_{q \times pm}$ 分别是根据训练样本的原图及其转置图构成的,因此对它们的分解可以同时进行。为了和文献[4]串行计算的 2DNMF 方法相区别,我们称其为 (2D)²NMF。由

于在训练图像矩阵的行和列方向上进行分解,可以提取到图像的二维结构信息。这样的信息存储在行基矩阵 $L_{p \times d}$ 和列基矩阵 $R_{q \times r}$ 中。因此,对于任意图像 A_k ,它在行和列基矩阵上的投影系数 C_k 为:

$$C_k = L^T A_k R \quad (7)$$

利用行基 $L_{p \times d}$ 和列基 $R_{q \times r}$ 重构图像可以表示为:

$$A_k \approx LC_k R^T, k = 1, 2, \dots, m \quad (8)$$

则 (2D)²NMF 的二维基定义为:

$$E = L \cdot R^T \quad (9)$$

3 基矩阵正交化和图像变形策略

3.1 基矩阵正交化

NMF 与 PCA 等方法不同,PCA 方法中各个特征向量是正交归一化的。在正交基构成的空间中,适合利用欧氏距离度量样本间的相似性。而 NMF 得到的基向量显然没有正交特性。我们注意到前述 NMF 方法的评价函数(式(4))是用欧氏距离进行相似性度量,通常不会得到好的性能。为了解决这个问题,刘维湘^[1]提出了将 NMF 提取的基矩阵正交化策略:即先将非负矩阵分解得到的基正交化,然后用样本在这组基上的投影来作进一步分析。假设通过 NMF 学习得到的基矩阵为 $B_{d \times r}$,将其正交化得到 $O_{d \times r}$,然后利用数据在 $O_{d \times r}$ 上的投影再作分类。在利用最近邻法分类时,这样的处理能够提高分类精度。显然,这样的策略可以推广应用于二维 NMF 方法的图像识别。本文采取的正交化措施为:

$$L_o = \text{orth}(L), R_o = \text{orth}(R) \quad (10)$$

3.2 图像对角化和镜像

上述 (2D)²NMF 算法计算基矩阵具有并行特性,为进一步提高算法性能打下良好基础。可以对原图像作更进一步的变形以便提取更多的结构信息。其中,借鉴 Diagonal-PCA 算法^[5],可以先将每个样本图像 A 进行对角化变换得到对角矩阵 \bar{A} ,再对 \bar{A} 进行非负矩阵分解^[5];也可以变换得到 A 的镜像图像 \hat{A} (通常利用矩阵元素的左右互换得到),再提取其特征;或混合应用图像对角化和镜像等方法,得到变形的新图像。通过图像的这些变形,使得像素在不同方向上的结构信息得以利用,从而提高识别精度。

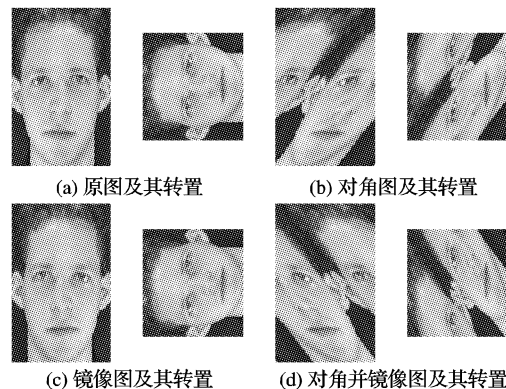


图2 图像对角化和镜像示例

4 利用加权的正交基矩阵进行图像分解

通过上述图像变形方法,我们可以得到原始图像的对角化图、镜像图,或者镜像图的对角化图像等。可以利用 (2D)²NMF 分别对这些图像进行学习得到行基矩阵 L_i 和列基矩阵 R_i ,然后分别将它们正交化为 L_{oi} 和 R_{oi} ,其中 i 表示从不同

的变形图像学习。由于这些行基和列基矩阵分别具有相同的阶次,可以利用加权方式得到混合的正交基矩阵。表示为:

$$L_{mix} = \sum_i w_i L_{oi} \quad (11)$$

$$R_{mix} = \sum_i w_i R_{oi} \quad (12)$$

其中, w_i 表示不同变形类别的图像权重。当某种变形类别的权重为 0 时,它的影响被消除。从这个角度而言,显然,已提出的一些二维人脸识别方法可以看作是本文方法中的特殊情况。

对于任意图像 A ,它在混合且正交化后的行和列基矩阵上的投影系数 C 为:

$$C = L_{mix}^T A R_{mix} \quad (13)$$

我们将利用这样的投影系数作为图像特征进行分类识别。

5 实验结果及分析

5.1 实验环境

为了测试上述改进的 NMF 算法的性能,本文利用 ORL 标准人脸数据库中 40×10 幅人脸图像进行了相关实验,实验涉及 $(2D)^2NMF$, $NMF^{[2]}$ 和 $2DNMF^{[4]}$ 。每幅图像大小为 112×92 。我们利用数据库中每个人的前 5 幅图像作训练样本,而后 5 幅图像作测试样本。分类器采用最近邻法,欧氏距离作为相似性度量。识别精度定义为测试集(200 个图像)中被正确分类的图像个数的百分比。

由于 NMF 是一种迭代分解法,其基矩阵和系数矩阵需要初始化,实际常采用随机初始化。而初始化可能影响分解结果,这可能导致分类结果的稳定性问题。为了更好地说明算法性能,我们对每次实验采用 3 次重复取平均的方法测量识别精度和速度等数据,来避免偏差。

文中采用的图像变形类别如图 2 示例,共有四种,分别为原图、对角图像、原图的镜像图像和镜像的对角图像,加权值分别为 0.3, 0.2, 0.2, 0.3。NMF 中子空间维数有两个,分别是行方向基向量个数 d 和列方向基向量个数 r ,为了方便实验,我们均设定 $r = d$, 以便比较。实验条件为 PC 机, CPU 1.7 GHz, 内存 512 M, Windows XP, Matlab 6.5。

5.2 实验比较及分析

1) 识别精度比较

图 3 比较了 $NMF^{[2]}$, $2DNMF^{[4]}$ 和 $(2D)^2NMF$ 算法的识别正确率。图中横坐标为子空间参考维数,而纵轴为识别正确率(三次平均)。

从图 3 数据可以看出, $1DNMF$ (正交基) 方法比 $1DNMF$ 的识别正确率高得多。说明对基矩阵进行正交化,能够消除基向量间的相关性,更适合利用欧氏距离作为相似性度量手段,因此识别精度的提高明显。

图中曲线显示,二维方法都比一维方法具有更高的识别精度。这是因为二维算法充分考虑了图像行方向和列方向的二维结构信息,因此提取的特征对目标能够更好地描述。

只利用原图像学习的情况下, $(2D)^2NMF$ 和 $2DNMF$ 的识别正确率指标几乎一致。说明本文提出的 $(2D)^2NMF$ 完全可以替代 $2DNMF$ 。如果利用变形图像得到混合正交基,在子空间维数合适的情况下 ($r = d = 20$), 识别精度会有少许提高,并且精度变化更小,会更稳定。这是因为采用图像变形策略后,实际增加了训练样本的数量。在这种情况下,分解并构建新的混合基矩阵,包含了更多的二维结构信息,因而能够提高识别正确率。另外,二维方法的子空间维数不会超过图像宽和

高的最小值,因此矩阵分解的内存耗费都远远小于一维方法。

然而值得注意的是,当子空间维数增加时, $(2D)^2NMF$ 识别率反而可能会下降。我们认为,随着维数的增加,一些与分类目的无关的特征会对分类产生影响,即对分类无用的特征增加了分类误差。为了解决这个问题,对 NMF 提取的特征加权将是下一步工作的方向。

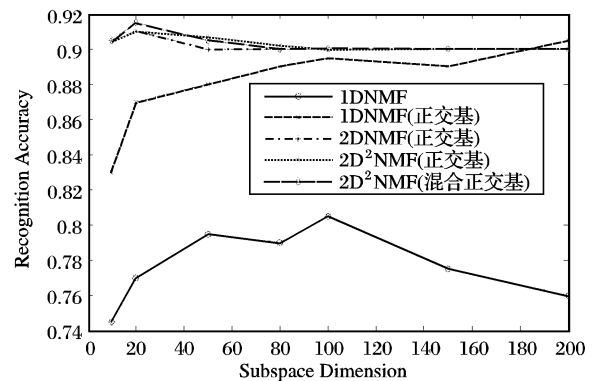


图 3 几种 NMF 人脸识别算法的比较

2) 矩阵分解的时间比较

表 1 显示了在图 3 最高识别正确率对应的子空间维数下,各方法分解矩阵的时间花费(三次平均)。

表 1 最佳子空间维数下算法训练时间比较

算法类别	训练时间/s
1DNMF	1040
2DNMF	48
$(2D)^2NMF$	50
$(2D)^2NMF$ (混合基)	200

从表 1 数据可以看出, $(2D)^2NMF$ 和 $2DNMF$ 都比 NMF 计算速度快 20 倍左右, $1DNMF$ 和 $2DNMF$ 在 PC 机上相比速度差别不大。 $(2D)^2NMF$ (混合基) 方法是 $(2D)^2NMF$ 的 4 倍时间花费,因为这里我们采用了 4 种变形图像得到 4 种基矩阵加权合成为一个混合基矩阵。限于条件,我们没有比较分布式计算时的速度。但是显而易见, $(2D)^2NMF$ 在并行计算方面具有优势,有更好的发展前景。

3) 图像重构

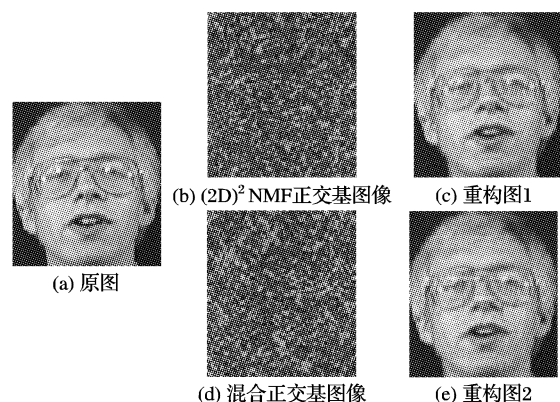


图 4 $(2D)^2NMF$ 得到的基图像及重构图像

$(2D)^2NMF$ 提取的基矩阵正交化后,由于是线性变换,仍然可以用来重构图像。如图 4 所示。原始图像经过在基矩阵上的映射投影后,能较准确地重构原图。其中,重构图 1 是正交基矩阵重构图,而重构图 2 是由混合正交矩阵得到的。这里的基图像类似于激光散斑效果,值得进一步研究。

(下转第 1666 页)

2.2 求解 Snake 能量函数最小值的贪婪算法

在 Kass 等提出了活动轮廓模型后,采用了变分法对能量函数进行求解,但存在许多问题,如由于轮廓点间的距离没有约束,各个点有可能沿着目标的轮廓线移动到使能量最小的区域。因此使 Snake 具有不断收缩最后成一点的趋势。在对能量函数进行计算时,需要进行大量的微分以保证收敛,因此要求能量函数必须可微,这样限制了其他有效强约束。另外,在使用变分法进行计算时,必须使用离散数据来估计高阶导数,而高阶导数的计算对噪声非常敏感,将导致计算不稳定。针对这些,有人提出了动态规划算法来求解能量函数。动态规划算法具有数值稳定,可加入强约束等优点,但是其计算量和存储量非常庞大。本文提出了求解能量函数最小值的贪婪算法,在克服了变分法的缺点基础上使计算量大大减少。

2.2.1 Snake 搜索区域的改进

在运动目标检测之后,通过以运动目标的心形为一端点引出 n 条射线,设定的初始轮廓点就在这些射线上。因此,可以让 Snake 轮廓点的搜索范围限制在这些射线上,即以初始轮廓点为中心,在两边位于射线上且和初始点距离在一定范围内的点进行搜索,这是一线性搜索区域,和原来 Snake 的以初始轮廓点为中心的矩形区域相比,大大减少了 Snake 的搜索范围,计算量相应减少。另一优点是把轮廓点限制在射线上,防止了轮廓点之间可能会产生的相互重叠现象,即在一定程度上阻止了 Snake 的轮廓点全部向能量最小的地方收缩的趋势。

2.2.2 求 Snake 能量函数最小值的贪婪算法

在定义了能量函数之后,应用贪婪算法可以求能量函数最小值,使初始轮廓逐渐趋向于实际的目标轮廓。算法具体步骤如下:

- 1) 根据式(15)计算得到梯度图像 $mag(x, y)$ 。
- 2) 从第一点 $i = 0$ 开始在射线 L_0 上以初始轮廓点为中心,在两边位于射线上且和初始点距离为 M ,根据式(9)进行能量计算,取能量最小的点为新的轮廓点。当计算射线上各点的能量时相邻点的位置不变,与其他位置均不相关。
- 3) 重复步骤2)直到 $i = n$ 为止,为一次迭代过程。迭代次数 $t = t + 1$ 。
- 4) 以上一次迭代过程中所得到的控制点的位置为初始位置重新对所有的控制点进行步骤2)的迭代计算。不停重复迭代计算,直至所有的控制点均不发生移动为止。

3 实验结果

实验用标准测试序列 Akiyo 的第二帧对本文所提出的算

法进行了测试。首先,应用本文的基于统计的运动检测和连通分量、填充技术处理后得到运动二值模板如图 2(c)。对运动模板进行初始轮廓点的提取结果如图 2(d)所示。经过改进的贪婪算法求得能量函数的最小值,得到最终的目标轮廓和运动目标如图 2(e)、(f)所示。从实验结果可以看到,用运动检测和活动轮廓相结合的运动目标分割算法可以得到准确的运动目标。



图 2 视频图像中运动目标分割图

参考文献:

- [1] CAVALLARO A, EBRAHIMI T. Video object extraction based on adaptive background and statistical change detection[C]// Proceedings of SPIE Electronic Imaging 2001 - Visual Communications and Image Processing. San Jose, California, USA: SPIE, 2001: 465 - 475.
- [2] SPAGNOLO P, D'ORAZIO T, LEO M, *et al.* Moving object segmentation by background subtraction and temporal analysis[J]. Image and Vision Computing, 2006, 24(5): 411 - 423.
- [3] ZENG W, DU J, GAO W. Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model[J]. Real-Time Imaging, 2005, 11(4): 290 - 299.
- [4] PARK H W, SCHOEPPFLIN T, KIM Y. Active contour model with gradient directional information: Directional snake[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(2): 252 - 256.
- [5] TSECHPENAKIS G, RAPANTZIKOS K, TSAPATSANLIS N, *et al.* A snake model for object tracking in natural sequences[J]. Signal Processing: Image Communication, 2004, 19(3): 219 - 238.
- [6] ZHANG J, GAO J B, LIU W S. Image sequence segmentation using 3-D structure tensor and curve evolution[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2001, 11(5): 629 - 641.
- [7] BILMES J A. A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models, TR-97-021[R]. U. C. Berkeley, Department of EE & CS, 1998.

(上接第 1662 页)

6 结语

本文对基于 NMF 的人脸识别方法进行了改进,提出可以并行处理的二维非负矩阵分解方法。在此基础上,提出利用适当的图像变形方法进一步增加图像结构信息,通过基向量正交化消除特征间的相关性等改进措施。通过人脸图像识别实验,集成了这些措施的算法能够加快特征提取速度、提高识别精度。实验结果证明了本文方法的有效性。

参考文献:

- [1] 刘维湘. 非负矩阵分解及其应用[D]. 西安: 西安交通大学, 2005.
- [2] LEE D D, SEUNG H S. Learning the parts of objects by non-nega-

tive matrix factorization[J]. Nature, 1999, 401: 788 - 791.

- [3] YANG J, YANG J Y. From image vector to matrix: a straightforward image projection technique - IMPCA vs. PCA[J]. Pattern Recognition, 2002, 35(4): 1997 - 1999.
- [4] ZHANG D Q, CHEN S C, ZHOU Z H. Two-dimensional non-negative matrix factorization for face representation and recognition[C]// Proceedings of the ICCV'05 Workshop on Analysis and Modeling of Faces and Gestures (AMFG'05), LNCS 3723. Beijing, China: [s. n.], 2005: 350 - 363.
- [5] ZHANG D Q, ZHOU Z H, CHEN S C. Diagonal principal component analysis for face recognition[J]. Pattern Recognition, 2006, 39(1): 140 - 142.