



# 基于鉴别字典学习的遮挡人脸姿态识别

邱益鸣<sup>1,3</sup> 廖海斌<sup>2</sup> 陈庆虎<sup>1</sup>

1 武汉大学电子信息学院,湖北 武汉, 430072

2 湖北科技学院计算机科学与技术学院,湖北 咸宁, 437100

3 江苏金陵科技集团有限公司,江苏 南京, 210008

**摘要:**利用字典学习与稀疏表示的信号重建与分类的性能,两步字典训练学习方法引入到鲁棒性人脸姿态识别中。首先,将人脸姿态离散化为不同的子空间,使用K-奇异值分解法(K-SVD)为每个子空间训练一个子字典使其对应一个类别;然后,将所有子字典组合成超完备字典;最后,采用基于Gabor特征与稀疏表示的方法进行姿态分类。为了提高字典的分类能力,本文采用两步字典训练学习方法,并在第二步学习中加入类别约束;为了提高算法的鲁棒性,本文重构一个遮挡人脸字典,解决人脸姿态识别中人脸遮挡问题。通过在公开的XJTU、PIE和CAS-PEAL-R1人脸库上的实验结果表明,本文方法在具有光照、噪声和遮挡变化的人脸库识别率均能达到95%左右,基本能达到实际应用的要求。

**关键词:**人脸姿态识别;人脸识别;遮挡人脸处理;字典学习;稀疏表示

中图法分类号:P237

文献标志码:A

人脸姿态识别就是根据图像确定人脸在三维空间中姿态参数的过程。人脸姿态识别在智能视频监控、人脸识别、人机交互和虚拟现实领域具有巨大的应用前景。例如,在智能视频监控方面,人脸姿态识别可以应用于驾驶监控系统,通过监控司机的人脸姿态变化来识别司机是否集中注意力开车,避免撞车情况的发生。此外,人脸姿态估计对人脸识别的准确度有很大的影响,许多人脸识别算法对正面人脸图像能够达到很好的识别率,但对于多姿态的非正面人脸图像,识别准确率会严重下降,而人脸姿态预估计是解决多姿态人脸识别的一种重要途径。

现有的人脸姿态检测方法大体上可以分为纹理子空间方法、3D方法和其他类方法3类。第1类方法通过基于2D人脸外观的学习方法实现姿态的检测与估计。其中比较典型的有主成分分析(principal component analysis, PCA)<sup>[1]</sup>、线性判别分析(linear discriminant analysis, LDA)<sup>[2]</sup>和独立子空间分析<sup>[3]</sup>等。由于PCA是一种线性降维方法,而人脸姿态3D旋转变化很大程度上是一种非线性变化。因此学者们使用核主成分分析(Kernel-PCA, KPCA)<sup>[4]</sup>方法、流型学习方法<sup>[5]</sup>

解决这种非线性变化问题。但是,核方法和流型学习方法有一个缺陷:随着人脸训练样本增加,它很难分离出身份和姿态。这就意味着,当人脸训练库足够大时,姿态估计的准确率会根据人的不同而变化。此类方法最大的特点是处理速度快,容易实现,但是需要通过大量样本的训练,对人脸的光照、表情等变化较为敏感,特别是对光照极差的视频人脸图像其准确率下降明显。

第2类方法认为人脸姿态检测本身就是一个3D问题,只有通过3D信息才能表征人脸姿态的本质特征。因此,这类方法往往通过抽取3D特征来表征不同姿态,或者利用不同视角下的多幅图像在三维空间中重建人脸的3D模型,实现姿态的检测<sup>[6-8]</sup>。这类方法往往对图像的大小和质量要求很高,并且会花费大量的运算时间。此类方法能够得到很高准确率,但是实时性不高,同时对视频监控中的超低分辨率和遮挡人脸图像效果不是很好。

第3类方法是一些非主流方法<sup>[9-13]</sup>,只能解决人脸姿态估计中部分问题或只能应用于某些特定场合。例如,文献[9]提出多相机的人脸姿态估计方法,为了正确估计人脸姿态,需要利用前后左

右6个相机拍照的6幅图像进行融合判别。文献[10]提出块聚类的方法进行人脸姿态估计,取得了不错的效果,但是他们的方法估计的姿态范围有限(只能识别45°范围的姿态变化)。文献[12]利用人脸对称性,文献[13]采用椭圆的方法,文献[11]采用三角形的方法进行人脸姿态识别,都得到了不错准确率,但是他们的方法只能对人脸左右偏转估计有效,而对人脸上下旋转情况失效。

目前,光照、噪声、遮挡、分辨率、身份、表情等因素的变化都会对姿态估计的准确性产生巨大的影响,如何消除这些因素的影响是目前亟需解决的问题。最近法国里昂大学和法国国立科学研究院的Zhang和Idrissi等人提出基于字典学习的人脸姿态识别框架<sup>[14]</sup>,他们的方法对人脸光照、噪声和分辨率变化具有鲁棒性。但可惜的是他们只研究了人脸左右偏转的情况,并没有对人脸上下偏转进行研究。另外,他们并没有充分挖掘字典学习与稀疏表示的特点与优势,导致没有对人脸遮挡问题做出相应的解决方案。

因此,本文进一步研究基于字典学习和稀疏表示的人脸姿态识别框架,为了使此框架能进行人脸上下偏转姿态识别,重新构建了人脸姿态超完备字典,使其囊括上下偏转姿态。同时,为了使算法对人脸遮挡变化具有鲁棒性,还增建了一个遮挡人脸字典。另外,为了提高识别率,在字典训练学习过程中加入类别约束信息,使字典不但具有高可重构性,还具有高分类性。

## 1 基于字典学习与稀疏表示的人脸姿态识别框架

### 1.1 人脸姿态Gabor特征稀疏表示

将人脸姿态以左右20°和上下30°偏转为间隔进行离散化,把人脸姿态化分为L种不同类别。假设有姿态训练样本集 $A=\{A_1 A_2 \dots A_L\}$ ,其中,第*i*( $i=1,2,\dots,L$ )类训练样本用特征向量矩阵表示为: $A_i=[S_{i,1} S_{i,2} \dots S_{i,n_i}] \in \mathbb{R}^{m \times n_i}$ 。其中, $n_i$ 表示第*i*类姿态样本数目, $m$ 表示样本特征维数, $S_{i,1}$ 是第*i*类姿态中第1个人脸的特征向量。研究表明,Gabor滤波器能高效地以多尺度形式获取图像的局部方向信息,对人脸光照、表情具有不变性。文献[15]指出采用基于Gabor特征的稀疏表示人脸识别方法比采用原始随机采样和PCA特征的稀疏表示方法要好。因此,本文采用Gabor特征向量作为字典的原子。人脸姿态图像的Gabor特征为:

$$S = (a_{0,0}^{(\rho)}; a_{1,0}^{(\rho)}; \dots; a_{7,4}^{(\rho)}) \quad (1)$$

式中, $a_{\mu,v}^{(\rho)}$ 是通过对Gabor滤波系数的模 $M_{\mu,v}^{(\rho)}$ 进行 $\rho$ 下采样而得到的列向量; $\mu, v$ 为Gabor滤波器的方向与尺度。 $M_{\mu,v}^{(\rho)}$ 为图像与Gabor核 $\psi_{\mu,v}$ 的卷积。Gabor核定义为:

$$\psi_{\mu,v}(z) = \frac{\|k_{\mu,v}\|^2}{\sigma^2} e^{-\|k_{\mu,v}\|^2 \|z\|^2/2\sigma^2} (e^{ik_{\mu,v}z} - e^{-\sigma^2/2}) \quad (2)$$

式中, $z(x,y)$ 表示像素; $k_{\mu,v}$ 为小波项; $\sigma$ 控制着高斯窗口宽度与波长的比例。

根据信号的稀疏表示原理,测试样本 $w_0$ 可以由训练样本中少数样本线性组合表示:

$$\hat{x} = \operatorname{argmin} \|x\|_0 \quad \text{s.t. } Ax = y \quad (3)$$

其中 $x$ 为线性组合系数。原子追踪、压缩感知与稀疏表示方法的研究成果<sup>[16]</sup>表明,如果上述方程组的解足够稀疏,就可以由下面的 $l_1$ 范数正则化的极小化问题求出:

$$(l^1): \hat{x}_1 = \operatorname{argmin} \|x\|_1 \quad \text{s.t. } Ax = y \quad (4)$$

该问题可以通过标准的线性规范方法进行求解。

理论上 $\hat{x}_1$ 应该只与训练样本中的某一类姿态样本的关系密切,其对应的表征系数非零。因此,可以清楚地对该待测姿态进行分类,如图1所示,输入的侧脸属于第4类(图中第2列)。因此,其线性组合表示的非零项主要集中在第2列(图1中红色方框)。

一般的,可以直接使用上述训练集(Gabor特征向量)作为人脸姿态字典。但这种方式构建的字典并非最优,同时可能导致字典矩阵过大。为了对字典进行优化,本文采用文献[17]中带类别信息的字典学习方法进行两步字典学习优化。

第1步:对每类样本采用经典K-means聚类和奇异值分解的字典学习方法(K-SVD)进行单独训练优化得到字典 $D_i$ ,字典训练的目标函数为:

$$\begin{aligned} & \langle D_i, \alpha_i \rangle = \\ & \operatorname{argmin}_{D_i, \alpha_i} \|S_i - D_i \alpha_i\|_2 \quad \text{s.t. } \|\alpha_i\|_0 \leq T \end{aligned} \quad (5)$$

式中, $S_i$ 为第*i*类训练样本矩阵; $\alpha_i$ 为线性组合系数; $T$ 为稀疏约束项。然后将各类子字典 $D_i$ 融合组成超完备字典 $D=[D_1, D_2 \dots D_L]$ 。

K-SVD方法在图像去噪和压缩等方面取得了非常好的效果。但由于其目标函数仅考虑了信号的重建误差和系数的稀疏性。所以通过上式训练得出的字典对于信号分类来说并一定是最优的。为了使字典具有很强的鉴别能力,文献[17]中的方法以第1步训练出字典作为初始化字典进行第2步训练学习:

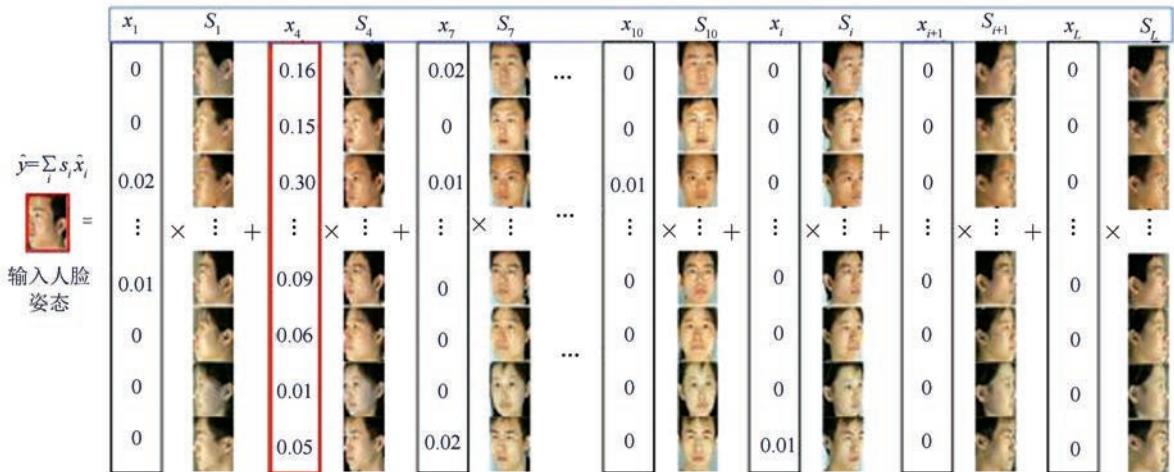


图 1 基于 SRC 人脸姿态分类识别

Fig.1 Face Poses Recognition Based on SRC

$$<\mathbf{D}', \mathbf{W}, \boldsymbol{\alpha}> =$$

$$\begin{aligned} \operatorname{argmin}_{\mathbf{D}', \mathbf{W}, \boldsymbol{\alpha}} & \|(\mathbf{S} \sqrt{\gamma} \times \mathbf{H}) - (\mathbf{D}' \sqrt{\gamma} \times \mathbf{W}) \times \boldsymbol{\alpha}\|_2 \\ \text{s.t. } & \|\boldsymbol{\alpha}\|_0 \leqslant T \end{aligned} \quad (6)$$

式中,  $\mathbf{H}$  为训练样本的类别;  $\mathbf{W}$  为引入的线性分类器;  $\boldsymbol{\alpha}$  为稀疏表示系数;  $\mathbf{S}$  为全体训练样本矩阵;  $\gamma$  为平衡控制因子。

## 1.2 人脸姿态分类

有了稀疏表示系数  $\boldsymbol{\alpha}$  和分类器  $\mathbf{W}$  后, 可以得出输入姿态的类别  $\mathbf{l}$ :

$$\mathbf{l} = \mathbf{W} \times \boldsymbol{\alpha} \quad (7)$$

式中,  $\boldsymbol{\alpha}$  可看成  $\mathbf{D}'$  中每个原子项  $d'_{ik}$  对重构输入图像的权值;  $\mathbf{W}$  中每列  $w_k$  可看成用来计算  $d'_{ik}$  与每类的相似度。因此,  $\mathbf{l}$  可以认为是输入姿态  $y$  与各类的加权相似度。理想情况下,  $\mathbf{l}$  只有一个非零项, 其他项都为零, 即  $\mathbf{l} = [0 \ 0 \cdots 1 \cdots 0 \ 0]$ 。因此, 可以根据其非零项确定其类别。但在实际应用中, 由于建模误差和复杂环境变化等因素影响导致  $\mathbf{l}$  中可能有多个非零项, 本文取最大项所对应的类别。

## 2 遮挡人脸姿态识别

根据稀疏表示的人脸识别方法<sup>[18]</sup>可知, 当测试人脸具有遮挡或腐蚀变化时, 式(4)可重写为:

$$\mathbf{y} = \mathbf{y}_0 + \mathbf{e}_0 = \mathbf{D}' \mathbf{x} + \mathbf{e}_0 = [\mathbf{D}' \quad \mathbf{D}_e] \begin{bmatrix} \mathbf{x} \\ \mathbf{x}_e \end{bmatrix} = \mathbf{B} \boldsymbol{\omega} \quad (8)$$

式中,  $\mathbf{B} = [\mathbf{D}' \quad \mathbf{D}_e] \in \mathbf{R}^{m \times (n+e_0)}$ ; 无遮挡图像  $\mathbf{y}_0$  与遮挡误差图像  $\mathbf{e}_0$  分别可由字典  $\mathbf{D}'$  和遮挡字典  $\mathbf{D}_e \in \mathbf{R}^{m \times e_0}$  稀疏表示。在稀疏表示人脸识别中,  $\mathbf{D}_e$  通常为正交单位矩阵。单位矩阵对图像误差

和噪声描述不够准确与直观, 且维数过高。首先, 收集不同姿态变化下的遮挡人脸图像, 每个人每种姿态下收集 4 种不同遮挡(可以用黑方块模拟)的人脸样本, 共收集 100 个人; 其次, 用遮挡人脸样本减去其对应类的均值人脸得到误差图像, 所有的误差图像构成了遮挡字典  $\mathbf{D}_e$  的训练集; 然后, 将全部训练样本与误差图像组成最终的字典训练集  $\mathbf{S}'$ 。因此, 遮挡人脸姿态识别中鉴别字典  $\mathbf{B}$  的学习问题转化成如下的优化问题:

$$\begin{aligned} & <\mathbf{B}, \mathbf{W}, \boldsymbol{\alpha}> = \\ & \operatorname{argmin}_{\mathbf{B}, \mathbf{W}, \boldsymbol{\alpha}} \| \begin{pmatrix} \mathbf{S}' \\ \sqrt{\gamma} \times \mathbf{H} \end{pmatrix} - \begin{pmatrix} \mathbf{B} \\ \sqrt{\gamma} \times \mathbf{W} \end{pmatrix} \times \boldsymbol{\alpha} \|_2 \\ & \text{s.t. } \|\boldsymbol{\alpha}\|_0 \leqslant T \end{aligned} \quad (9)$$

通过式(9)便可求出带遮挡部分的稀疏表示系数  $\boldsymbol{\omega}$ , 并从中分解出稀疏表示系数  $\mathbf{x}$  和  $\mathbf{x}_e$ 。进行遮挡人脸姿态识别。可见, 本文通过增加遮挡字典  $\mathbf{D}_e$  的方法可以成功地解决人脸遮挡问题。

实际应用中, 人脸遮挡主要包括墨镜与围巾遮挡。目前还没有出现不同姿态下包括墨镜与围巾遮挡的人脸库。因此, 本文用黑方块遮挡眼睛和下巴区域来模拟人脸遮挡情况来补充遮挡训练样本不足的缺陷, 并以此构建遮挡人脸字典和测试人脸库。

## 3 实验与分析

本节将利用 XJTU(Xi'an Jiaotong University Database)<sup>[19]</sup>、PIE (Pose, Illumination and Expression Database)<sup>[20]</sup> 和 CAS-PEAL-R1 Database<sup>[21]</sup> 验证本文提出的人脸姿态估计算法的有效性。并与目前主流的 PCA<sup>[1]</sup>、ICA<sup>[3]</sup> 和先进的

基于字典学习和稀疏表示的方法<sup>[14]</sup>( dictionary-learning sparse representation, DLSR)相比较。

XJTU 人脸库是由西安交通大学建立的,共采集了 1247 位志愿者的 33 669 幅人脸图像,包括视点(左右偏转)和光照两个子库,每位志愿者拍摄 19 张视点图像和 8 张光照图像。本文采用了相同光照条件下 130 人的视点图像进行实验,



图 2 XJTU 人脸姿态库像示例

Fig.2 The Sample Images of XJTU Pose Database

PIE 人脸数据库由美国卡耐基梅隆大学创建,包括不同光照下 68 位志愿者多种姿态和表情的共 41 368 幅人脸图像。其中,光照和姿态变化的图像都是在严格的控制条件下采集的,目前该数据库已成为人脸算法领域的一个重要的测

试集合之一。该数据库在平面外左右偏转方向上人脸姿态角度的变化范围为  $-90^\circ \sim 90^\circ$ , 每人有 9 个不同姿态,如图 3 所示。由于该数据库每种姿态都有光照变化,因此,本文使用该人脸库验证算法对光照的鲁棒性。



图 3 PIE 人脸姿态库示例

Fig.3 The Sample Images of PIE Pose Database

CAS-PEAL 人脸库是中国科学院建立的具有不同表情、姿态、装饰物、年龄和光照等变化的大规模人脸库。当前公开的人脸库版本(CAS-PEAL-R1)含有 1 040 人的 30 864 个不同条件下的样本。本文使用其中 pose 库(由 167 个人,每个人包含 22 种不同上下、左右偏转姿态组成)作姿态识别实验。本文将其人脸姿态划分为 9 类

(左右偏转 7 类,上下偏转 2 类),取其中 100 人作为训练字典库,剩下的 67 人作为测试库,人脸姿态划分示例如图 4 所示。为了验证算法对遮挡的处理能力,对测试图像进行遮挡处理。图 5 为人脸图像加噪和遮挡的示例样本,测试时加噪图像样本数量为  $30 \times 9 = 270$ , 遮挡图像样本数量为  $67 \times 9 = 603$ 。



图 4 CAS-PEAL-R1 人脸姿态库示例

Fig.4 The Sample Images of CAS-PEAL-R1 Pose Database



图 5 人脸图像噪声示例样本

Fig.5 The Sample Face Images with Noise and Occlusion

本实验采用的图像均为手动对齐后的人脸图像,为了验证各方法对人脸图像光照、噪声和遮挡的鲁棒性,本文在手动对齐的人脸图像数据上进行训练,然后在有光照、噪声和遮挡的待测图像上进行人脸姿态判别,分别统计不同姿态的识别准确率。由于 DLSR 方法只进行无遮挡人脸左右姿态识别,为了公平比较,本文将在 XJTU 和 PIE 两个人脸库进行算法性能比较。在 CAS-PEAL 人脸库上验证本文算法对上下、左右偏转的识别能力和对遮挡的处理能力。实验时,每类姿态进行 10 次实验,所有实验均重复 10 次,统计其平均识别率。

4 种不同方法在 XJTU 人脸姿态库上的识别如图 6 所示,从图中可以看出:图像无光照、噪声和遮挡变化时,3 种方法都能得到很好的效果,本文方法和 DLSR 方法识别率都达到 99% 以上。

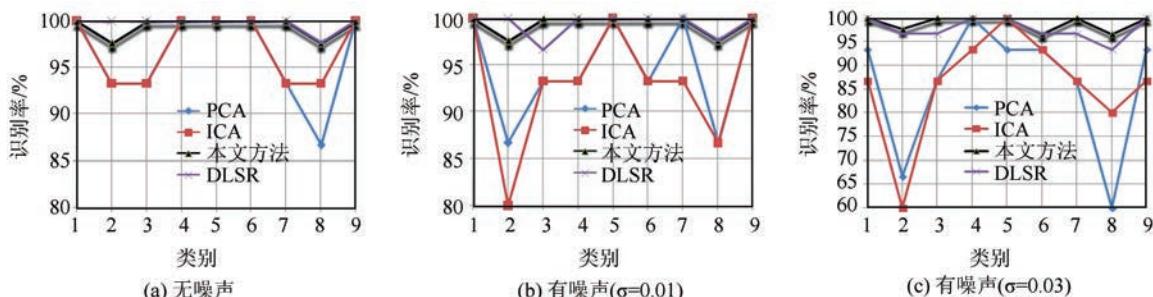


图 6 基于 XJTU 的不同姿态分类方法比较

Fig.6 The Compare Results with XJTU

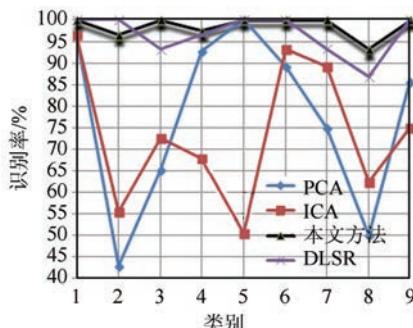


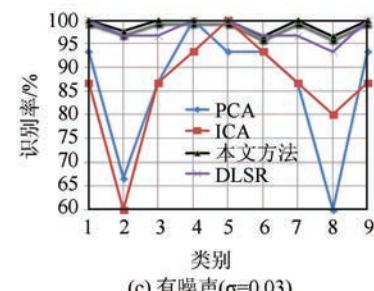
图 7 基于 PIE 的不同姿态分类方法比较(有光照变化)

Fig.7 The Comparing Results with PIE (with Illumination)

为了验证本文算法对遮挡的鲁棒性,本文对测试人脸加长方形方块进行遮挡模拟实验。采用本文方法得到的 9 类不同姿态的混淆矩阵识别结果如表 1 所示,表中对角线加粗部分的识别率为正确分类的概率,其他部分表示错分的概率(本来属于对角线类别,但分到其它类别的概率),每一行相加为 1。为了突出本文方法的优势,表 2 给出了采用 DLSR 方法进行遮挡模拟实验的结果。而 PCA 和 ICA 方法在遮挡模拟实验中其识别效

但是当图像有噪声情况后,基于 PCA 和 ICA 的姿态判别方法性能下降很快(见图 6(b)、6(c)),而本文方法基本保持不变,DLSR 方法受噪声干扰比 PCA 和 ICA 方法小,比本文方法大。分析其原因,主要是 DLSR 采用图像分块的方法,此方案虽然在无光照、噪声和遮挡情况下获得了比本文好的结果。但是,图像块对噪声的敏感性要比整体图像大,导致其抗噪性不强。

为了比较算法对光照的鲁棒性,采用 PIE 人脸姿态库作进一步实验,实验结果如图 7 所示。从图中可以看出,当人脸出现光照变化时,PCA 和 ICA 方法识别率急剧下降,而 DLSR 和本文方法基本比较稳定。由于本文采用的 Gabor 特征具有光照和表情不变性,因此本文方法比 DLSR 方法更具稳定性。



果较差平均识别率低于 70%,基于篇幅限制考虑并没有给出其混淆矩阵结果。对比表 1 和表 2 可以看出,本文方法在处理遮挡人脸图像时有优势,这是由于本文方法加了遮挡字典的缘故而导致。

为了进一步说明本文算法的性能,表 3 给出了几种算法的平均运行时间(单位:s)比较结果(配置为:hp Core i3 M330 2.13GHz 2GB, Matlab 2012R)。从表中可以看出:本文算法的运行时间多于线性子空间方法,低于 DLSR 方法。但是,本文算法对人脸噪声和光照变化的鲁棒性是线性子空间方法所不能比拟的。

## 4 结语

人类可以轻易地进行人脸姿态识别是因为人们通过大量训练学习的结果。本文仿照人类训练学习机制进行人脸姿态超完备字典的构建与学习,独特性地提出基于字典学习与稀疏表示的人脸姿态识别方法。为了使算法更具鲁棒性和可行

表1 本文方法遮挡人脸姿态分类混淆矩阵结果

Tab.1 Confusion Matrix of Face Pose Classification on the Occlusion Database of the Proposed Method

类别	1	2	3	4	5	6	7	8	9
1	1								
2	0.0207	<b>0.9643</b>	0.015						
3		0.0265	<b>0.933</b>	0.0402					
4				1					
5					<b>0.975</b>	0.025			
6					0.1	<b>0.867</b>	0.033		
7							1		
8			0.013	0.01	0.002			<b>0.975</b>	
9			0.003	0.015	0.002				<b>0.98</b>

表2 DLSR方法遮挡人脸姿态分类混淆矩阵结果

Tab.2 Confusion Matrix of Face Pose Classification on the Occlusion Database of DLSR

类别	1	2	3	4	5	6	7	8	9
1	0.955	0.02	0.025						
2	0.0307	<b>0.9200</b>	0.035	0.010		0.0043			
3		0.046	<b>0.900</b>	0.050		0.004			
4			0.10	<b>0.875</b>	0.025				
5				0.014	<b>0.961</b>	0.025			
6					0.103	<b>0.867</b>	0.030		
7							<b>0.98</b>	0.02	
8			0.013	0.01	0.002		0.05	<b>0.920</b>	0.005
9				0.015	0.002	0.003		0.03	<b>0.95</b>

表3 运行时间比较/s

Tab.3 Running Times/s

人脸库	算法			
	PCA	ICA	DLSR	本文算法
XJTU	0.153	0.194	0.643	<b>0.360</b>
PIE	0.152	0.195	0.571	<b>0.360</b>
CAS-PEAL	0.163	0.214	0.600	<b>0.400</b>

性,本文采用Gabor特征作为人脸姿态图像的特征,并在字典训练学习中加入类别约束信息使字典不但具有重建性能还具有分类性能。同时通过构建人脸姿态图像遮挡字典的方法解决人脸遮挡问题。实验结果表明本方法不但具有稀疏表示人脸识别方法(SRC)中对光照和遮挡的鲁棒性,同时还能克服SRC中的“小样本”问题。SRC要求训练样本“足够”多,这在实际应用中有时是相当苛刻的。而基于SRC的姿态分类方法却能克服“小样本”问题。因为人脸姿态离散化为有限类(假定9类),每类包含100个姿态样本,总共也只需900个样本,这在现实中很容易满足。因此,相比于人脸识别问题,基于稀疏表示的分类方法更适合于姿态识别问题。

## 参考文献

[1] Srinivasan S, Boyer K L. Head Pose Estimation Using View Based Eigenspaces [C]. The 16th Interna-

national Conference on Pattern Recognition, Quebec City, Canada, 2002

- [2] Wang Huaqing. Face Pose Estimation with Local Geometry Preserving [D]. Xi'an: Xidian University, 2013 (王华青.基于局部几何结构的人脸图像姿态估计[D].西安:西安电子科技大学,2013)
- [3] Lis Z. Learning Multi-view Face Subspaces and Facial Pose Estimation Using Independent Component Analysis[J]. *IEEE Trans. Image Process.*, 2005, 14(6): 705-712
- [4] Wu J, Trivedi M M. A Two-Stage Head Pose Estimation Framework and Valuation[J]. *Pattern Recognition*, 2008, 41(5): 1138-1158
- [5] Wang C, Song X B. Robust Head Pose Estimation Using Supervised Manifold Projection[C]. International Conference on Image Processing, USA, 2012: 161-164
- [6] Dou P F, Wu Y H, Shah S K, et al. Benchmarking 3D Pose Estimation for Face Recognition[C]. International Conference on Pattern Recognition, Stockholm, Sweden, 2014
- [7] Akos P, Levente H. Precise 3D Pose Estimation of Human Faces[C]. The 9th International Conference on Computer Vision Theory and Applications, Lisbon, Portugal, 2014
- [8] Li D Q, Pedrycz W. A Central Profile-Based 3D Face Pose Estimation [J]. *Pattern Recognition*,

- 2014, 47(2): 525-534
- [9] Muñoz-Salinas R, Yeguas-Bolívar E, Saffiotti A, et al. Multi-camera Head Pose Estimation [J]. *Machine Vision and Applications*, 2012, 23(3): 479-490
- [10] Nuevo J, Bergasa L M, Jiménez P. RSMAT: Robust Simultaneous Modeling and Tracking [J]. *Pattern Recognition Letters*, 2010, 31(16): 2455-2463
- [11] Chen Zhenxue, Chang Faliang, Liu Chunsheng, et al. Pose Parameters Estimate Based on AdaBoost Algorithm and Facial Feature Triangle [J]. *Geomatics and Information Science of Wuhan University*, 2011, 36(10): 1164-1167(陈振学, 常发亮, 刘春生, 等. 基于 Adaboost 算法和人脸特征三角形的姿态参数估计[J]. 武汉大学学报·信息科学版, 2011, 36(10): 1164-1167)
- [12] Ma B P, Li A A, Chai X J, et al. CovGa: A novel Descriptor Based on Symmetry of Regions for Head Pose Estimation [J]. *Neurocomputing*, 2014, 143: 97-108
- [13] Chen X P, Yang Q Q, Liao H H, et al. Real-Time Face Pose Estimation in Video Sequence[C]. The Second International Workshop on Education Technology and Computer Science, Wuhan, 2010
- [14] Zhang Y Y, Idrissi K, Garcia C. A Dictionary-learning Sparse Representation Framework for Pose Classification[C]. IEEE International Workshop on Machine Learning for Signal Processing, UK, 2013
- [15] Yang M, Zhang L, Shiu S C K, et al. Gabor Feature Based Robust Representation and Classification for Face Recognition with Gabor Occlusion Dictionary [J]. *Pattern Recognition*, 2013, 46(2): 1865-1878
- [16] Xiang Jinhai, Fan Heng, Xu Jun, et al. Object Tracking Based on Local Sparse Representation [J]. *J. Huazhong Univ. of Sci. & Tech. (Natural Science Edition)*, 2014, 42(7): 187-193(向金海, 樊恒, 徐俊, 等. 基于局部稀疏表示的目标跟踪[J]. 华中科技大学学报(自然科学版), 2014, 42(7): 187-193)
- [17] Zhang Q, Li B. Discriminative K-SVD for Dictionary Learning in Face Recognition[C]. IEEE Conference on Computer Vision & Pattern Recognition, USA, 2010
- [18] Meng Yang, Luc Van Gool, Lei Zhang. Sparse Variation Dictionary Learning for Face Recognition with a Single Sample Per Person[C]. IEEE International Conference on Computer Vision, San Francisco, USA, 2013
- [19] AI & R Institute of Artificial Intelligence and Robotics Xi'an Jiaotong University. Oriental Face Database [EB/OL]. <http://www.aiar.xjtu.edu.cn/groups/face/Chinese/Homepage.htm>, 2013
- [20] The CMU PIE database [EB/OL]. [http://www.ri.cmu.edu/projects/project\\_418.html](http://www.ri.cmu.edu/projects/project_418.html), 2013
- [21] Gao W, Cao B, Shan S G, et al. The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations [J]. *IEEE Transaction on System Man, and Cybernetics (Part A)*, 2008, 38: 149-161

## Occluded Face Pose Recognition Based on Dictionary Learning with Discrimination Performance

QIU Yiming<sup>1,3</sup> LIAO Haibin<sup>2</sup> CHEN Qinghu<sup>1</sup>

<sup>1</sup> School of Electronic Information, Wuhan University, Wuhan 430072, China

<sup>2</sup> School of Computer Science and Technology, Hubei University of Science and Technology, Xianning 437100, China

<sup>3</sup> Jiangsu Jinling Sci&Tech Group Co., Ltd, Nanjing 210008, China

**Abstract:** This paper make full use of dictionary learning and sparse representation for signal reconstruction and classification, and present a two-step dictionary learning method and apply it to robust face pose recognition. The proposed approach models the appearance of face images from the subspace via K-SVD that learns the sub-dictionary from a set of images. A combination of the trained sub-dictionaries of all pose classes are used as an over-complete dictionary. Finally, the Gabor features are extracted for sparse representation and classification. In order to improve the classification ability, we put forward a two-step dictionary learning method, and carry out dictionary learning with label con-

# A Novel Approach Combining KI Criterion and Inverse Gaussian Model to Unsupervised Change Detection in SAR Images

ZHUANG Huifu<sup>1</sup> DENG Kazhong<sup>1</sup> YU Mei<sup>1</sup> FAN Hongdong<sup>1,2</sup>

1 Jiangsu Key Laboratory of Resources and Environmental Information Engineering, China University of Mining and Technology, Xuzhou 221116, China

2 State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu 610059, China

**Abstract:** In this context, a novel approach combining inverse Gaussian model (IGM) and the Kittler-Illingworth (KI) criterion has been proposed to carry out unsupervised change detection in synthetic aperture radar (SAR) images. The minimum error threshold could be computed by exploiting the Bayes decision theory under the assumption that hybrid IGM could describe the distribution of the changed and unchanged class in difference image. Experiments carried out on two sets of multi-temporal SAR images indicate that the proposed approach can effectively estimate the probability density function of the unchanged and changed classes in the difference image and acquire a reasonable threshold for yielding a better change map from the difference image.

**Key words:** synthetic aperture radar (SAR); Kittler-Illingworth criterion; inverse Gaussian model (IGM); Bayes decision theorem; threshold selection; change detection

**First author:** ZHUANG Huifu, PhD candidate, specializes in change detection of remote sensing images. E-mail: huifuzhuang@163.com

**Corresponding author:** DENG Kazhong, PhD, professor. E-mail: kzdeng@cumt.edu.cn

**Foundation support:** The Natural Science Foundation of China, No. 51774270; the Opening fund of State Key Laboratory of Geohazard Prevention and Geoenvironment Protection (Chengdu University of Technology), No. SKLGP2016K008.

-----  
(上接第 281 页)

straints in the second step. Additionally, in order to improve robustness against face occlusion, we introduce a pose occlusion dictionary to code the occluded portions of face images. Several experiments were performed on XJTU, PIE, and CAS-PEAL-R1 databases. Recognition results show that the proposed method can achieve a recognition rate of about 95% under illumination, noise, and occlusion variations. It can satisfy the requirements of practical applications.

**Key words:** face pose recognition; face recognition; occlusion face processing; dictionary-learning; sparse representation

**First author:** QIU Yiming, PhD, researcher, specializes in image processing and pattern recognition. E-mail: fhqim@sina.com

**Corresponding author:** LIAO Haibin, PhD, associate professor. E-mail: liao\_haibing@163.com

**Foundation support:** The Program of the Natural Science Foundation of Hubei Province, No. 2017CFB300; the Hubei Provincial Education Department Science and Technology Research Projects, No. Q20172805; the Hubei Provincial Education Science Plan Project, No. 2016GB086.