

DOI: 10.11992/tis.201609002

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20170317.1937.002.html>

## 鲁棒的正则化编码随机遮挡表情识别

刘帅师, 郭文燕, 张言, 程曦

( 长春工业大学 电气与电子工程学院, 吉林 长春 130000 )

**摘 要:** 为了提高随机遮挡下人脸表情的识别率, 提出一种新的人脸表示模型, 即鲁棒的正则化编码, 通过正则回归系数对给定信号进行鲁棒回归。首先, 为了减少遮挡对人脸表情识别系统的影响, 待识别表情图像的每个像素点将被分配不同的权重; 然后, 由于被遮挡部分像素点应分配较小的值, 通过连续迭代直到权重收敛于设定的权重阈值; 最后, 待测图像的稀疏表示将通过最优权重矩阵计算, 且待测表情图像分类结果由训练样本逼近待测图像的最小残差决定。应用该方法在日本的 JAFFE 表情数据库和 Cohn-Kanade 数据库上取得较理想的结果, 且实验结果表明该方法对随机遮挡表情识别具有鲁棒性。

**关键词:** 随机遮挡; 正则化编码; 自动更新权重; 表情识别

**中图分类号:** TP391.4    **文献标志码:** A    **文章编号:** 1673-4785(2018)02-0261-08

**中文引用格式:** 刘帅师, 郭文燕, 张言, 等. 鲁棒的正则化编码随机遮挡表情识别[J]. 智能系统学报, 2018, 13(2): 261-268.

**英文引用格式:** LIU Shuaishi, GUO Wenyan, ZHANG Yan, et al. Recognition of facial expression in case of random shielding based on robust regularized coding[J]. CAAI transactions on intelligent systems, 2018, 13(2): 261-268.

## Recognition of facial expression in case of random shielding based on robust regularized coding

LIU Shuaishi, GUO Wenyan, ZHANG Yan, CHENG Xi

(College of Electrical and Electronic Engineering, Changchun University of Technology, Changchun 130000, China)

**Abstract:** In order to improve facial expression recognition rate under the random shielding, a new face representation model was proposed: robust regularized coding. Regularized regression coefficients are used for carrying out robust regression for the given signals. Firstly, in order to reduce the influence of shielding on facial expression identification system, all pixels of the expression image to be identified will be assigned with different weights; then, because the occluded pixels should have lower weight values, hence, successive iteration is applied until the weight converges to the set weight threshold; finally, the sparse representation of image to be tested can be calculated by using the optimal weight matrix, in addition, the classified results of the expression image to be tested are determined by the minimal residual that the training samples approximate to the test image. The proposed method achieved an ideal performance in Japanese JAFFE expression database and Cohn-Kanade database, in addition, the experimental results show that the method is robust for the recognition of the facial expression randomly shielded.

**Keywords:** random shielding; regularized coding; automatic update of weight; recognition of facial expression

人脸表情识别技术是生理学、心理学、图像处理、模式识别和计算机视觉等领域的一个具有挑战性的交叉学科<sup>[1]</sup>。为了确保信息的完整性, 研究人员们使用不存在遮挡人脸表情图像在受控的实验室

条件下进行实验和研究<sup>[2]</sup>。然而, 人脸遮挡在现实生活中很常见, 例如, 太阳镜可以遮挡眼睛区域、一条围巾或外科口罩遮挡嘴部<sup>[3]</sup>。因此, 在人脸存在遮挡的表情识别仍然是人脸表情识别系统在实际应用中最重要的瓶颈之一。

近年来, 针对部分遮挡人脸表情识别, 研究人员已经提出了许多方法来减少遮挡对表情识别的影

收稿日期: 2016-09-06. 网络出版日期: 2017-03-17.

基金项目: 吉林省教育厅“十三五”科学技术项目 (JJKH20170571KJ).

通信作者: 刘帅师. E-mail: [liu-shuaishi@126.com](mailto:liu-shuaishi@126.com).

响。Kotsia 等<sup>[4]</sup>受 Fisher 的线性判别分析和支持向量机 (support vector machine, SVM) 的启发, 提出了一种新颖的最小类内方差的多类分类器来研究在不同人脸器官遮挡的情况下对人脸表情识别的影响。Tarrés 等<sup>[5]</sup>提出了基于 PCA(principal component analysis) 和 LDA(linear discriminant analysis), 并结合直方图均衡化和均值、方差归一化预处理的方法, 减少了遮挡部分对人脸识别过程的影响。Kotsia 等<sup>[6]</sup>对遮挡的人脸图像进行 Gabor 小波滤波提取纹理特征, 利用监督的判别非负矩阵因子分解 (discriminant non-negative matrix factorization, DNMF) 进行图像分解, 再采用基于模型的方法来描述特定面部特征的几何位移来完成遮挡图像的特征表征。Zhang 等<sup>[7]</sup>利用蒙特卡罗算法对表情图像提取 Gabor 特征, 并遍历表情图像的每个区域进行模板匹配以产生对遮挡具有鲁棒性的特征向量。Wang 等<sup>[8]</sup>利用改进的中心对称局部二值模式和梯度中心对称局部方向模式 GCS-LDP, 利用卡方距离求取测试集图像与训练集图像特征直方图之间的距离。但是, 上述方法主要研究眼部遮挡和嘴部遮挡对人脸表情识别效果的影响, 没有充分考虑遮挡在现实生活中出现的特点, 对随机遮挡情况的适应性较弱。人脸遮挡的特点是遮挡可以在人脸的任何地方发生, 并且遮挡范围的大小和遮挡的形状都是未知的, 没有任何关于它的先验知识<sup>[9]</sup>。因此, 不能只考虑脸部某个区域对人脸表情识别的影响情况, 应该根据遮挡的特点来展开研究, 并提出一些可以克服这个问题的方法。Wright 等<sup>[10]</sup>采用稀疏编码方法完成人脸识别任务, 并提出使用已知类别的训练人脸图像对测试人脸图像进行稀疏表示的方法, 在随机遮挡人脸识别系统中取得了较为理想的识别效果。M. Zhu 等<sup>[11]</sup>利用稀疏分解求出待测图像的稀疏表示系数, 并在待测图像所在的子空间内实现表情类别判断, 该方法使得待测图像的分解系数变得更稀疏, 同时避免身份特征对表情分类的干扰。

为了提高稀疏表示的鲁棒性和有效性, 本文提出了基于鲁棒的正则化编码和自动更新权重的随机遮挡表情识别方法。受鲁棒回归理论的启发<sup>[10]</sup>, 即通过自适应地不断迭代的方法来给残差分配不同的权重, 直到估计过程收敛。本文通过假设编码残差和编码系数分别是独立同分布的, 并基于最大后验估计的原则来对给定的信号进行鲁棒回归, 为了方便实现, 正则化编码的最小化问题将转换成一个自动更新权重的问题, 通过设计合理的权重函数可以鲁棒地识别出遮挡部分从而减小它们对编码过程的影响。在 JAFFE 和 Cohn-Kanade 数据库上, 本文方

法对随机遮挡表情识别的鲁棒性进行了验证, 取得了较理想的识别效果。

## 1 鲁棒的正则化编码

通常情况下, 稀疏编码问题可以定义为

$$\hat{\alpha} = \arg \min \|\alpha\|_1 \text{ s.t. } \|\mathbf{y} - \mathbf{T}\alpha\|_2^2 \leq \varepsilon \quad (1)$$

式中:  $\mathbf{y}$  是待测的表情图像,  $\mathbf{T}$  是训练表情图像字典,  $\alpha$  是待测的表情图像  $\mathbf{y}$  在训练表情图像字典  $\mathbf{T}$  上的编码向量, 并且  $\varepsilon > 0$ 。保真项定义为  $\|\mathbf{y} - \mathbf{T}\alpha\|_2^2$ 。

如果事先假设编码残差  $\mathbf{e} = \mathbf{y} - \mathbf{T}\alpha$  服从高斯分布, 式 (1) 的解将变成最大似然估计的解, 如果  $\mathbf{e}$  服从拉普拉斯分布,  $l_1$  稀疏限制的最大似然估计解变为  $\hat{\alpha} = \arg \min \|\alpha\|_1 \text{ s.t. } \|\mathbf{y} - \mathbf{T}\alpha\|_1 \leq \varepsilon$ 。

其实,  $\hat{\alpha} = \arg \min \|\alpha\|_1 \text{ s.t. } \|\mathbf{y} - \mathbf{T}\alpha\|_1 \leq \varepsilon$  是式 (1) 的另一种表现形式, 因为它们都具有相同的拉格朗日公式:  $\arg \min \{\|\alpha\|_1 + \lambda \|\mathbf{y} - \mathbf{T}\alpha\|_1\}$ 。

在实际应用中这种假设是不成立的, 尤其是当人脸表情图像  $\mathbf{y}$  被遮挡时。为了构建一个鲁棒性更强的人脸表情图像的稀疏编码模型, 本文提出了一个更通用的而且效率更高的鲁棒正则化编码模型。

### 1.1 鲁棒的正则化编码模型

贝叶斯估计的观点确切地说是从最大后验概率估计观点考虑人脸表示的问题。通过字典  $\mathbf{T}$  对待测表情图像  $\mathbf{y}$  进行编码, 编码向量  $\alpha$  的最大后验概率估计变成  $\hat{\alpha} = \arg \max \ln P(\alpha | \mathbf{y})$ 。利用贝叶斯公式得

$$\hat{\alpha} = \arg \max \{\ln P(\mathbf{y} | \alpha) + \ln P(\alpha)\} \quad (2)$$

将式 (1) 中的字典  $\mathbf{T}$  改写成  $\mathbf{T} = [\mathbf{r}_1; \mathbf{r}_2; \dots; \mathbf{r}_n]$ , 其中,  $\mathbf{r}_i$  表示  $\mathbf{T}$  的第  $i$  列, 而且  $\mathbf{e} = \mathbf{y} - \mathbf{T}\alpha = [\mathbf{e}_1; \mathbf{e}_2; \dots; \mathbf{e}_n]$ , 其中,  $\mathbf{e}_i = \mathbf{y}_i - \mathbf{r}_i\alpha, i = 1, 2, \dots, n$ 。假设元素  $\mathbf{e}_i$  是独立同分布的, 并且概率密度函数为  $f_\theta(\mathbf{e}_i)$ , 而且  $P(\mathbf{y} | \alpha) = \prod_{i=1}^n f_\theta(\mathbf{y}_i - \mathbf{r}_i\alpha)$ 。

与此同时, 假设编码向量  $\alpha = [\alpha_1; \alpha_2; \dots; \alpha_m]$  中的元素  $\alpha_j, j = 1, 2, \dots, m$  是独立同分布的并且概率密度函数为  $f_0(\alpha_j)$ , 而且  $P(\alpha) = \prod_{j=1}^m f_0(\alpha_j)$ 。从而式 (2) 中  $\alpha$  的最大后验概率估计为

$$\hat{\alpha} = \arg \max \left\{ \prod_{i=1}^n f_\theta(\mathbf{y}_i - \mathbf{r}_i\alpha) + \prod_{j=1}^m f_0(\alpha_j) \right\} \quad (3)$$

令  $\rho_\theta(\mathbf{e}) = -\ln f_\theta(\mathbf{e})$  和  $\rho_0(\alpha) = -\ln f_0(\alpha)$  式 (3) 转成:

$$\hat{\alpha} = \arg \min \left\{ \sum_{i=1}^n \rho_\theta(\mathbf{y}_i - \mathbf{r}_i\alpha) + \sum_{j=1}^m \rho_0(\alpha_j) \right\} \quad (4)$$

本文把式 (4) 的模型称为鲁棒的正则编码, 由于保真项  $\rho_\theta(\mathbf{y}_i - \mathbf{r}_i\alpha)$  对遮挡造成的异常值具有鲁棒性, 且根据先验概率  $P(\alpha), \sum_{j=1}^m \rho_0(\alpha_j)$  是正则化项。

因此当 $\alpha_j$ 服从拉普拉斯分布的时候有: $P(\alpha) = \prod_{j=1}^m \exp(-\|\alpha_j\|_1/\sigma_\alpha)/2\sigma_\alpha; \sum_{j=1}^m \rho_0(\alpha_j)$ 将变成稀疏限制 $l_1$ 范数。对于分类问题,理想的结果是只有待测表情图像与训练样本组成的完备字典中对应的目标类别的表示系数有很大的绝对值。由于事先不知道待测图像属于哪个类别,可以进行一个合理的推理,是只有很少一部分的稀疏表示系数具有显著值。因此,假设稀疏表示系数 $\alpha_j$ 服从高斯分布,则有:

$$f_0(\alpha_j) = \beta \exp\left\{-\left(\|\alpha_j\|_1/\sigma_\alpha\right)^\beta\right\}/(2\sigma_\alpha \Gamma(1/\beta)) \quad (5)$$

式中 $\Gamma$ 表示伽马函数。

由于表情图像的变化多样性,很难预先确定稀疏表示残差的分布。通常,假设概率密度函数 $f_\theta(e)$ 是对称的、单调并且可微的。因此, $\rho_\theta(e)$ 具有以下性质:

- 1)  $\rho_\theta(0)$ 是 $\rho_\theta(e)$ 的局部最小值;
- 2) 对称性:  $\rho_\theta(e_i) = \rho_\theta(-e_i)$ ;
- 3) 单调性: 当 $|e_1| > |e_2|$ 时,  $\rho_\theta(e_1) > \rho_\theta(e_2)$ 。不失一般性,令 $\rho_\theta(0) = 0$ 。

鲁棒的正则化编码模型需要解决的两个关键问题:如何确定 $\rho_\theta$ (或者 $f_\theta$ )的分布和最小化能量函数。如果只是简单的使 $f_\theta$ 服从高斯分布或者拉普拉斯分布并且 $f_\theta$ 服从拉普拉斯分布,本文提出的模型将退化成为式(1)所示的传统的稀疏表示问题。为了解决这两个问题并更有效地获得鲁棒的正则化编码模型的最大后验概率,本文将式(4)的最小化问题转换成迭代权重正则编码问题。

## 1.2 迭代权重优化鲁棒的正则化编码模型

定义 $F_\theta(e) = \sum_{i=1}^n \rho_\theta(e_i)$ 。 $F_\theta(e)$ 在定义域内某点 $e_0$ 处的一阶泰勒展开公式:

$$\tilde{F}_\theta(e) = F_\theta(e_0) + (e - e_0)^T F'_\theta(e_0) + R_1(e) \quad (6)$$

式中: $F'_\theta(e)$ 是 $F_\theta(e)$ 的一阶导数, $R_1(e)$ 是 $F_\theta(e)$ 的高阶导数的余项式。定义 $\rho'_\theta$ 是 $\rho_\theta$ 的一阶导数,并且有: $F'_\theta(e_0) = [\rho'_\theta(e_{0,1}); \rho'_\theta(e_{0,2}); \dots; \rho'_\theta(e_{0,n})]$ ;  $e_{0,i}$ 是 $e_0$ 的第 $i$ 个元素。使 $F'_\theta(e)$ 严格显凸性便于最小化,近似余项为 $R_1(e) \approx \frac{1}{2}(e - e_0)^T W(e - e_0)$ 其中, $W$ 是对角矩阵,使 $e$ 中的元素独立且在 $F_\theta(e)$ 中 $e_i$ 和 $e_j$ ( $i \neq j$ )没有交叉项。

$F_\theta(e)$ 在 $e = 0$ 取得最小值的同时,它的近似值 $\tilde{F}_\theta(e)$ 在 $e = 0$ 也应取得最小值。令 $F'_\theta(0) = 0$ ,可以得到 $W$ 的对角元素如式(7):

$$W_{i,i} = \rho'_\theta(e_{0,i})/e_{0,i} \quad (7)$$

根据 $\rho_\theta$ 的性质, $\rho'_\theta(e_i)$ 和 $e_i$ 符号相同,所以 $W_{i,i}$ 是非负的标量。因此进一步,从而 $\tilde{F}_\theta(e)$ 可以写成 $\tilde{F}_\theta(e) = \frac{1}{2} \|W^{1/2}e\|_2^2 + b_{e_0}$ 。

式中: $b_{e_0} = \sum_{i=1}^n (\rho_\theta(e_{0,i}) - \rho'_\theta(e_{0,i})e_{0,i}/2)$ 是由 $e_0$ 决定的标

量常量。不考虑 $b_{e_0}$ ,式(4)模型可近似为式(8):

$$\hat{\alpha} = \arg \min \left\{ \frac{1}{2} \|W^{1/2}(y - T\alpha)\|_2^2 + \sum_{j=1}^m \rho_0(\alpha_j) \right\} \quad (8)$$

虽然,式(8)是式(4)的局部近似值,但是这样做可以将鲁棒的正则化编码模型的最小化问题通过迭代再加权重 $l_2$ 正则编码来解决,也就是通过式(7)不断更新权重 $W$ 。这样最小化问题转变成了如何计算对角权重矩阵 $W$ 。

## 1.3 权重 $W$

$W_{i,i}$ 表示分配给待测表情图像 $y$ 每个像素点 $i$ 的权值。从人的感官认识出发,被遮挡部分的像素点应该具有较低的权重,这样可以减少它们对编码过程的影响。由于完备字典是由非遮挡的人脸表情图像构成的,可以很好地表征人脸表情,然而遮挡部分像素造成的异常值将具有较大的编码残差,因此,这些具有较大的编码残差像素点应具有较小的权重。通过观察式(7)可以得到 $W_{i,i}$ 与 $e_i$ 成反比,与 $\rho'_\theta(e_i)$ 成正比。由于 $\rho_\theta$ 可微、对称、单调并且在原点取得最小值,可以假设 $W_{i,i}$ 是连续且对称的,与 $e_i$ 成反比而且有界。不失一般性,令 $W_{i,i} \in [0, 1]$ ,综合多方面的考虑,逻辑函数是权重函数的最好选择。本文选用与逻辑函数具有相似性质的SVM hinge loss函数<sup>[12]</sup>作为权重函数。

初值对人脸表情识别取得较理想的识别效果至关重要。为了对待测表情图像 $y$ 设置初值,首先应该初始化 $y$ 的编码残差 $e$ 。本文对 $e$ 初始化为 $e = y - T\alpha^{(0)}$ ;  $\alpha^{(0)}$ 是初始编码向量。由于待测表情图像 $y$ 所属类别事先未知,因此 $\alpha^{(0)}$ 的合理初始编码向量可以设置为 $\alpha^{(0)} = \left[ \frac{1}{m}; \frac{1}{m}; \dots; \frac{1}{m} \right]$ 。这样 $T\alpha^{(0)}$ 表示的就是所有训练表情图像的平均表情图像。

通过不断迭代优化更新权重 $W$ ,直到权重收敛为止,即相邻迭代权重之间的差异足够小。具体来说,应当式(9)成立时停止迭代:

$$\|W^{(t)} - W^{(t-1)}\|_2 / \|W^{(t-1)}\|_2 < \gamma \quad (9)$$

式中: $\gamma$ 是较小的正数。本文迭代15次权重就趋于收敛。得到收敛后的权重矩阵 $W$ 后,最优的稀疏表示 $\hat{\alpha}$ 可以通过式(8)计算得到。最后,通过式(10)计算每类训练表情图像逼近待测表情图像 $y$ 的逼近残差。

$$r_i(y) = \|W_{\text{final}}^{1/2}(y - T\delta_i(\hat{\alpha}))\|_2, i = 1, 2, \dots, k \quad (10)$$

式中: $\delta_i(\hat{\alpha})$ 是 $\hat{\alpha}$ 第 $i$ 类训练样本空间最终编码向量, $W_{\text{final}}$ 是最终的权重矩阵, $k$ 表示表情类别数。

根据最小逼近残差的准则公式(11),待测表情图像 $y$ 最终将被分类到训练表情图像逼近待测表情图像残差最小的类别。可由式(11)进行判断:



$$\text{identity}(\mathbf{y}) = \arg \min_{i \in \{1, 2, \dots, k\}} (r_i(\mathbf{y})) \quad (11)$$

因此,本文方法的流程如图1所示。首先,待测的人脸表情图像的每个像素点赋予不同的权重。其次,通过连续迭代得到收敛的权重矩阵。权重随着每次迭代的收敛曲线如图2所示。然后,得到收敛的权重矩阵 $\mathbf{W}$ 后,待测表情图像的最优稀疏表示也可以通过计算得到。最后,计算每类训练表情图像逼近待测表情图像 $\mathbf{y}$ 的编码残差,并根据最小逼近残差的准则将待测表情图像 $\mathbf{y}$ 分类到训练表情图像

逼近待测表情图像最小逼近残差所对应的类别。每类训练表情图像逼近待测表情图像的逼近残差如图3所示。为了减少原始表情特性的特征维数,本文应用PCA的方法对特征进行降维处理,Eigen-face特征应用到与本文方法进行对比的其他算法中。定义 $\mathbf{P}$ 为PCA的投影矩阵,那么,式(8)将变为

$$\hat{\alpha} = \arg \min \left\{ \frac{1}{2} \|\mathbf{P}\mathbf{W}^{1/2}(\mathbf{y} - \mathbf{T}\alpha)\|_2^2 + \sum_{j=1}^m \rho_0(\alpha_j) \right\} \quad (12)$$

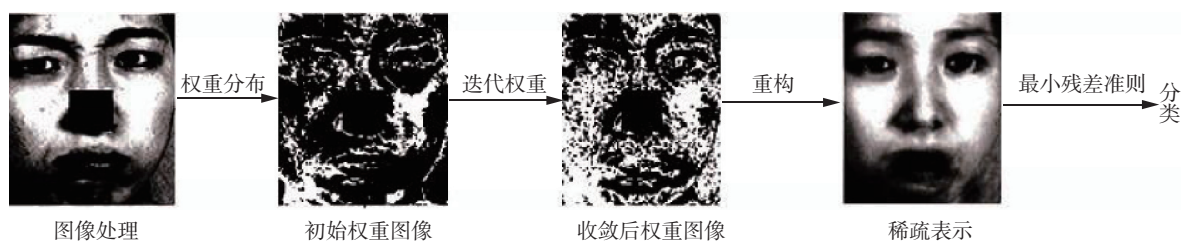


图1 本文方法的流程

Fig. 1 The structure of our method

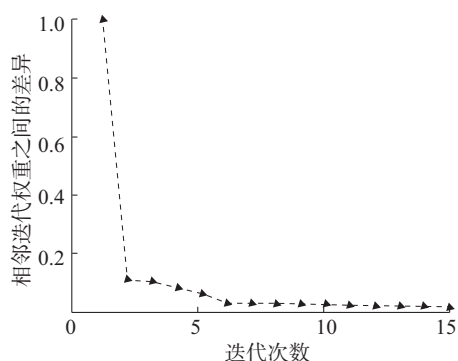


图2 权重收敛曲线

Fig. 2 The convergence curve of the weigh

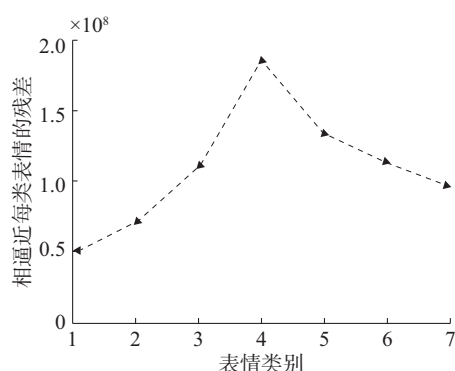


图3 不同类别训练图像逼近待测图像的残差图

Fig. 3 The residual of each training class approximates the test image

## 2 实验描述与结果分析

实验采用日本女性表情图像 JAFFE 数据库和 Cohn-Kanade 数据库来验证本文所提方法的可行性

和有效性。JAFFE 数据库包含 10 个女性共 213 张人脸表情图像,并且每个人都有 7 种表情,每种表情有 3 或 4 张表情图像样本。实验时选用 10 个人共 137 表情图像作为训练样本,其中(高兴-19,惊讶-20,悲伤-20,恐惧-20,厌恶-18,愤怒-20 和中性-20)。其余的 76 张人脸表情图像作为测试样本。基于 JAFFE 数据库的表情图像数量少,实验将遍历 3 种情况来取得平均识别率。而 Cohn-Kanade 人脸表情数据库是由 100 名大学生按照指定的方式来从中性表情呈现 23 幅表情序列,这些人都来自 18~30 岁的大学心理系的学生。其中 15% 为欧洲人(包括黑种人和白种人),3% 为亚洲人或拉丁人种,65% 为女性。该数据库也包含与 JAFFE 数据库一样的 7 种表情类别。对于 Cohn-Kanade 数据库,选用 10 个人 7 种表情(高兴-6,惊讶-6,悲伤-6,恐惧-6,厌恶-6,愤怒-6 和中性-6)共 420 张表情序列进行实验。其中,10 个人 7 种表情共 210 张表情图像作为训练样本,其余的作为测试样本。为了验证算法在 Cohn-Kanade 数据库的泛化性能实验遍历 6 种情况来取得平均识别率。

### 2.1 实验描述

由于 JAFFE 数据库和 Cohn-Kanade 数据库中的表情图像稍有头部倾斜和尺寸大小不一,需要经过预处理来消除这些差异。本文采用类似文献[12]的预处理方法:通过旋转使眼睛水平面对准,并根据两眼间的距离来从原始的表情图像裁剪出实验用的只含正面人脸表情的矩形区域。JAFFE 数据库中的原始人脸表情图像的尺寸为 256×256, Cohn-Kanade

数据库中的原始人脸表情图像尺寸为  $640 \times 490$ 。两数据库中的表情图像进行尺寸归一化  $128 \times 104$ , 利用直方图均衡化来增强表情图像某些区域的局部对比度, 如图4所示。



(a) JAFFE



(b) Cohn-Kanade

图4 实验用的两数据库中的部分随机遮挡表情图像

Fig. 4 Some samples of occluded facial images in two databases

与其他的影响因素不同, 如姿势的变化, 它的变化特点是可以事先预测判别出来的。然而, 面部遮挡是特别难以处理的, 因为它具有随机性的特点, 也就是说, 遮挡可以发生在人脸表情图像的任意位置并且大小也是任意的。我们对于遮挡发生的位置和遮挡面积的大小没有任何明确的先验知识。关于遮挡唯一有的先验信息就是遮挡毁坏的像素点可能是彼此相邻的, 就是说某个区域的像素点可能是连续毁坏。图4表示的两个表情数据中一些在不同遮挡程度块遮挡级别下的人脸表情图像。遮挡级别表示的是遮挡的部分占整个人脸表情图像的百分比是多少。所以遮挡级别是正数并且处于 $[0, 1]$ , 遮挡级别为0表示图像没有被遮挡, 1表示图像全部被遮挡, 例如: 遮挡级别0.1表示图像的10%被遮挡。基于稀疏表示的人脸表情识别方法最重要特点对于人脸遮挡具有鲁棒性。为更好地验证本文提出的方法对于随机遮挡的鲁棒性, 采用表情识别方法:

KNN<sup>[14]</sup>(K-nearest neighbor)、SVM<sup>[15]</sup>、SRC<sup>[10]</sup>(sparse representation-based classifier)、GSRC<sup>[16]</sup>(gabor feature based sparse representation) 与本文方法在以下两种数据库进行对比。

## 2.2 结果分析

表1和表2分别表示的是本文提出的方法与其他方法分别在 JAFFE 数据库上和 Cohn-Kanade 数据库上对应不同遮挡级别的平均识别率。

表1 不同方法在 JAFFE 数据库上的识别率

Table 1 The accuracies of different methods on JAFFE %

遮挡级别	本文方法	KNN	SVM	SRC	GSRC
0	96.05	85.57	89.47	90.79	92.11
0.1	96.05	84.26	88.16	90.79	91.52
0.2	94.47	81.58	86.84	88.16	89.47
0.3	93.25	80.55	83.67	86.84	88.16
0.4	90.06	84.26	88.16	90.79	91.52
0.5	89.57	80.55	83.67	86.84	88.16
0.6	85.53	75.00	77.32	79.78	81.58
0.7	61.84	57.43	58.82	59.23	60.32
0.8	53.33	47.23	48.64	49.42	51.67
0.9	22.81	18.14	19.32	20.43	21.76

表2 不同方法在 Cohn-Kanade 数据库上的识别率

Table 2 The accuracies of different methods on Cohn-Kanade %

遮挡级别	本文方法	KNN	SVM	SRC	GSRC
0	98.10	86.84	91.23	92.79	93.21
0.1	98.10	84.35	89.32	90.23	92.79
0.2	97.62	82.12	85.93	88.69	89.23
0.3	96.19	79.98	84.13	85.97	88.69
0.4	94.29	86.84	91.23	92.79	93.21
0.5	93.33	75.21	79.87	82.46	84.27
0.6	92.86	73.87	74.97	76.78	78.87
0.7	70.48	68.12	68.94	69.23	69.98
0.8	32.86	27.64	28.32	29.43	30.76
0.9	18.10	15.33	16.92	17.23	17.89

从表1和表2可看出, 随遮挡级别增大人脸表情的识别率逐渐减小, 符合人们的感性认识。表情识别方法在遮挡级别为0.1~0.5会取得较理想的识别效果。由于训练样本和测试样本用自身像素值不需特征提取过程, KNN<sup>[16]</sup>和 SVM<sup>[17]</sup>方法在遮挡级别很大时没有很好的识别效果。且这两种方法要结合提取较好区分性特征的提取方法才可发挥较好的分



类效果,可看出本文的方法比 SRC<sup>[10]</sup>和 GSRC<sup>[12]</sup>表情识别率略高。

本方法在随机遮挡的情况下取得比其他方法较为理想的识别效果,接下来分析本文在不同遮挡级别对于每种表情识别的影响大小。两种数据库不同遮挡级别对每种表情的遮挡情况如表3和表4。

表3 每种表情在 JAFFE 数据库不同遮挡级别的识别率  
Table 3 The accuracies of each expression on JAFFE in different levels of block occlusions %

遮挡级别	生气	厌恶	恐惧	高兴	中性	悲伤	惊讶
0	90	81.82	100	100	100	100	100
0.1	90	81.82	100	100	100	100	100
0.2	90	81.82	100	100	100	100	90
0.3	90	81.82	100	100	100	90.91	90
0.4	90	81.82	91.67	100	100	90.91	90
0.5	90	81.82	83.33	100	100	81.82	90
0.6	90	81.82	75.00	91.67	100	81.82	90
0.7	60	63.64	58.33	41.67	90	72.73	50
0.8	60	45.45	41.67	41.67	80	54.55	50
0.9	20	18.18	16.67	16.67	60	18.18	10

表4 每种表情在 Cohn-Kanade 数据库不同遮挡级别的识别率  
Table 4 The accuracies of each expression on Cohn-Kanade in different levels of block occlusions %

遮挡级别	生气	厌恶	恐惧	高兴	中性	悲伤	惊讶
0	96.67	90	100	100	100	100	100
0.1	96.67	90	100	100	100	100	100
0.2	96.67	90	100	100	100	100	96.67
0.3	93.33	90	100	100	100	93.33	96.67
0.4	93.33	90	100	96.67	100	90	90
0.5	86.67	90	100	96.67	100	90	90
0.6	86.67	90	100	93.33	100	90	90
0.7	63.33	50	70	76.67	86.67	76.67	70
0.8	33.33	23.33	23.33	36.67	60	20	33.33
0.9	10	16.67	3.33	23.33	46.67	6.67	20

从表3和表4可以看出,随着表中给出的人脸表情图像的随机遮挡级别的增大,两种数据库上不同的人脸表情识别率随之降低。在两种数据库中,生气、高兴、中性、悲伤和惊讶表情在图像遮挡级别为0~0.1取得了较为理想的识别效果。这是由于遮

挡级别较小,这些表情图像的决策信息缺失的少。在这两种数据库上所有表情只有惊讶表情的识别效果在遮挡级别为0.2时受到了影响,其他表情都没有受到影响。当图像遮挡级别为0.3时,在这两种数据库上所有表情只有生气、悲伤和惊讶表情识别效果受到了影响。当图像遮挡级别为0.4~0.5,在 JAFFE 数据库上高兴和中性表情取得了较好的识别效果,在图像遮挡级别为0.6时中性表情取得了较好的识别效果。然而在 Cohn-Kanade 数据库上,当图像遮挡级别为0.4~0.6时,恐惧和中性表情取得了较好的识别效果。当图像遮挡级别为0.7~0.9时,所有的表情的识别率(除了中性表情外)都受到了较为严重的影响。

从两表中可以看出中性表情识别率在不同的图像遮挡级别下都可以保持了较高的识别率。即使在遮挡级别为0.9的 JAFFE 数据库上,中性表情的识别率仍为60%。这是由于本文在赋予编码残差 $e$ 的初值时,选用的是所有训练表情图像的平均表情作为 $e$ 的初值,中性表情和平均人脸表情很相似。因此,即使在遮挡级别很大是,中性表情也较其他表情更容易更有效地识别。

图5表示的是 JAFFE 数据库上所有测试的中性表情图像和所有训练表情图像的平均表情图像。虽然在 JAFFE 数据库上中性表情的识别效果在遮挡级别很高的时候也能取得较好识别效果,但是在 Cohn-Kanade 数据库上这种现象表现的并不是十分明显。在遮挡级别为0.9时的中性表情识别率为46.67%。尽管中性表情的识别率较其他的表情识别率高,但是与在 JAFFE 数据上的识别率相比还是相差较大。这是由于 JAFFE 数据库上的人脸表情图像都是女性,并且都属于同一国家的。

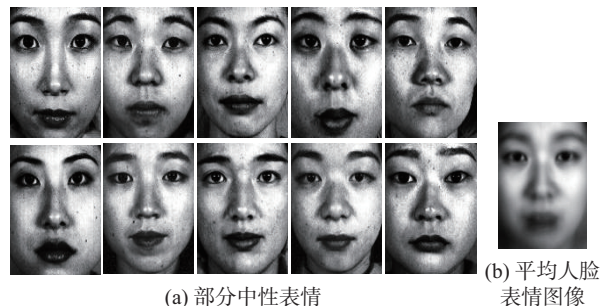
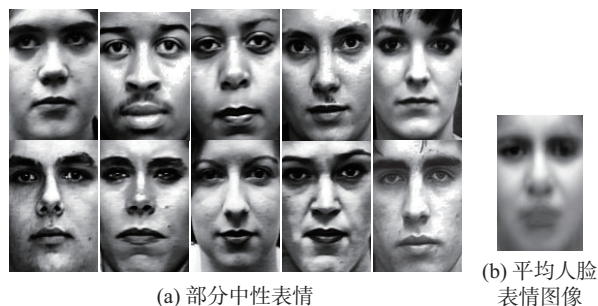


图5 JAFFE 数据库

Fig. 5 JAFFE

而 Cohn-Kanade 数据库中的人脸表情图像是来自不同的国籍和不同性别的。这将造成如图6所示的中性表情和平均人脸表情之间的相似性较小。因此,当遮挡级别为0.7~0.9时,在 Cohn-Kanade 数据库上的中性表情较 JAFFE 上的中性表情的识别率

低。尽管, 中性表情的识别率在两种表情数据库上相差较大, 但是 Cohn-Kanade 数据库中的表情图像来自不同的国籍和性别更符合实际情况, 在该数据库上进行实验更有利于算法的推广和实际应用。



(a) 部分中性表情

(b) 平均人脸表情图像

图6 Cohn-Kanade 数据库

Fig. 6 Cohn-Kanade

### 3 结束语

本文提出了基于鲁棒的正则化编码模型和自动更新权重的随机遮挡表情识别方法。根据人脸表情遮挡随机性的特点, 提高了稀疏表示的鲁棒性和有效性并且减少随机遮挡部分对人脸表情识别的影响。本文方法使用原始图像数据 (像素点) 即可不需要采用特征降维、特征提取、综合训练样本和特定领域信息等, 通过求取编码问题的最大后验概率, 从而实现对遮挡的鲁棒性。根据编码残差来对待测图像的所有像素点自适应的分配和反复迭代权重, 这样可以鲁棒地辨别出遮挡造成的奇异值并减少它们对编码过程的影响。在 JAFFE 数据库和 Cohn-Kanade 数据库上与其他几种方法进行了不同遮挡级别情况下识别率的对比实验, 由结果可以看出本文提出的方法取得了较好的识别效果, 较其他几种方法有效并对随机遮挡具有较强的鲁棒性。

### 参考文献:

- [1] MERY D, BOWYER K. Face recognition via adaptive sparse representations of random patches[C]//IEEE International Workshop on Information Forensics and Security. London, UK, 2015: 13–18.
- [2] WANG J, LU C, WANG M, et al. Robust face recognition via adaptive sparse representation[J]. IEEE transactions on cybernetics, 2014, 44(12): 2368.
- [3] 赵军, 赵艳, 杨勇, 等. 基于降维的堆积降噪自动编码机的表情识别方法[J]. 重庆邮电大学学报: 自然科学版, 2016, 28(6): 844–848.  
ZHAO Jun, ZHAO Yan, YANG Yong, et al. Facial expression recognition method based on stacked denoising auto-encoders and feature reduction[J]. Journal of Chongqing university of posts and telecommunications: natural science edition, 2016, 28(6): 844–848.
- [4] KOTSIA I, PITAS I, ZAFEIRIOU S, et al. Novel multi-class classifiers based on the minimization of the within-class variance[J]. IEEE transactions on neural networks, 2009, 20(1): 14–34.
- [5] TARRÉS F, RAMA A, TORRES L. A novel method for face recognition under partial occlusion or facial expression variations[C]//Proceedings of the 47th International Symposium ELMAR. Zadar, Croatia, 2005: 163–166.
- [6] KOTSIA I, BUCIU I, PITAS I. An analysis of facial expression recognition under partial facial image occlusion[J]. Image and vision computing, 2008, 26(7): 1052–1067.
- [7] ZHANG Ligang, TJONDRONEGORO D, CHANDRAN V. Toward a more robust facial expression recognition in occluded images using randomly sampled Gabor based templates[C]//Proceedings of 2011 IEEE International Conference on Multimedia and Expo. Barcelona, Spain, 2011: 1–6.
- [8] 王晓华, 李瑞静, 胡敏, 等. 融合局部特征的面部遮挡表情识别[J]. 中国图象图形学报, 2016, 21(11): 1473–1482.  
WANG Xiaohua, LI Ruijing, HU Min, et al. Occluded facial expression recognition based on the fusion of local features[J]. Journal of image and graphics, 2016, 21(11): 1473–1482.
- [9] 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.
- [10] 朱明早, 李树涛, 叶华. 基于稀疏表示的遮挡人脸表情识别方法[J]. 模式识别与人工智能, 2014, 27(8): 708–712.  
ZHU Minghui, LI Shutao, YE hua. An occluded facial expression recognition method based on sparse representation[J]. Pattern recognition and artificial intelligence, 2014, 27(8): 708–712.
- [11] 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.
- [12] CAO J, ZHANG K, LUO M, et al. Extreme learning machine and adaptive sparse representation for image classification[J]. Neural networks the official journal of the international neural network society, 2016, 81(c): 91.
- [13] ZHANG Jian, JIN Rong, YANG Yiming. Modified logistic regression: an approximation to SVM and its applications in large-scale text categorization[C]//Proceedings of the Twentieth International Conference on Machine Learning. Washington, DC, USA, 2003: 888–895.
- [14] LIU Shuaishi, ZHANG Yan, LIU Keping, et al. Facial expression recognition under partial occlusion based on

Gabor multi-orientation features fusion and local Gabor binary pattern histogram sequence[C]//Proceedings of the 9th International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Beijing, China, 2013: 218–222.

[15] LIU Licheng, CHEN Long, CHEN C L. Weighted joint sparse representation for removing mixed noise in image [J]. IEEE transactions on cybernetics includes computational approaches to the field of cybernetics, 2016: 1–12.

[16] YANG M, SONG T, LIU F, et al. Structured regularized robust coding for face recognition[J]. IEEE transactions on image processing a publication of the IEEE signal processing society, 2013, 22(5): 1753–1766.

[17] 罗元, 吴彩明, 张毅. 基于 PCA 与 SVM 结合的面部表情识别的智能轮椅控制[J]. 计算机应用研究, 2012, 29(8): 3166–3168.

LUO Yuan, WU Caiming, ZHANG Yi, et al. Facial expression recognition based on principal component analysis and support vector machine applied in intelligent wheelchair[J]. The research and application of computer, 2012, 29(8): 3166–3168.

#### 作者简介:



刘帅师, 女, 1981 年生, 副教授, 博士, 主要研究方向为模式识别、计算机视觉。



郭文燕, 女, 1991 年生, 硕士研究生, 主要研究方向为模式识别、机器学习。



张言, 男, 1989 年生, 硕士研究生, 主要研究方向为模式识别、机器学习。

## 2018 年第三届智能机器人系统亚太会议 (ACIRS 2018) 2018 3rd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS 2018)

2018 3rd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS 2018) will be held during July 21-23, 2018 in Singapore.

ACIRS 2018 provides a forum for scientific advances in the theory and practice of Intelligent Robot Systems. It is a highly selective, single-track meeting that will be soliciting submissions presenting significant, original, and previously unpublished research. ACIRS 2018 aims to be one of the leading international conferences in the Asia Pacific region, and will provide an exciting environment for researchers to present and discuss the latest technologies, algorithms, system architectures, and applications.

All full paper submissions will also be peer reviewed and evaluated based on originality, technical and/or research content/depth, correctness, relevance to conference, contributions, and readability. The full paper submissions will be chosen based on technical merit, interest, applicability, and how well they fit a coherent and balanced technical program.

Accepted papers of ACIRS 2018 will be published in Conference Proceedings and will be submitted to EI Compendex and Scopus.