

1-2-2019

Facial Expression Recognition Based on Improved Dictionary Learning and Sparse Representation

Li Ming

1.School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China; ;2.College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

Xiujiao Peng

1.School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China; ;

Wang Yan

2.College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Facial Expression Recognition Based on Improved Dictionary Learning and Sparse Representation

Abstract

Abstract: In order to overcome the problems induced by illumination and occlude in facial expression recognition and reduce the time required by sparse representation classification, *the facial expression recognition algorithm with fusion of HOG feature and improved KC-FDDL dictionary learning sparse representation is put forward. Improved K-means cluster and Fisher discrimination dictionary learning is implemented on the training set generated by extracting HOG features of normalized expression image. Facial expression classification is conducted by the sparse representation with weighted residuals.* Experimental results on the Cohn-Kanade database show that this method is lower time-consuming and more accurate for similar facial expression classification than other facial expression classification methods.

Keywords

occlude, HOG features, KC-FDDL dictionary learning, sparse representation, weighted residuals

Recommended Citation

Li Ming, Peng Xiujiao, Wang Yan. Facial Expression Recognition Based on Improved Dictionary Learning and Sparse Representation[J]. Journal of System Simulation, 2018, 30(1): 28-36.

基于改进的字典学习与稀疏表示的人脸表情识别

黎明^{1,2}, 彭秀姣¹, 王艳²

(1.南昌航空大学信息工程学院, 江西 南昌 330063; 2.南京航空航天大学自动化学院, 江苏 南京 210016)

摘要: 为克服人脸表情图像识别过程中光照、遮挡等带来的影响, 减少稀疏表示分类的时间, 提出一种融合HOG特征和改进KC-FDDL(K-means Cluster and Fisher Discrimination Dictionary Learning)字典学习稀疏表示的人脸表情识别算法。对归一化后的表情图像提取HOG特征构成训练集, 对训练集进行改进的K-均值聚类的Fisher判别字典学习, 利用残差加权的稀疏表示进行表情分类。Cohn-Kanade数据库上的实验结果表明, 该算法相比其他的人脸表情分类方法具有耗时低、相似表情分类更准确的优势。

关键词: 遮挡; HOG特征; KC-FDDL字典学习; 稀疏表示; 残差加权

中图分类号: TP391 文献标识码: A 文章编号: 1004-731X(2018)01-0028-09

DOI: 10.16182/j.issn1004731x.joss.201801004

Facial Expression Recognition Based on Improved Dictionary Learning and Sparse Representation

Li Ming^{1,2}, Peng Xiujiao¹, Wang Yan²

(1. School of Information Engineering, Nanchang Hangkong University, Nanchang 330063, China;

2. College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: In order to overcome the problems induced by illumination and occlude in facial expression recognition and reduce the time required by sparse representation classification, the facial expression recognition algorithm with fusion of HOG feature and improved KC-FDDL dictionary learning sparse representation is put forward. Improved K-means cluster and Fisher discrimination dictionary learning is implemented on the training set generated by extracting HOG features of normalized expression image. Facial expression classification is conducted by the sparse representation with weighted residuals. Experimental results on the Cohn-Kanade databases show that this method is lower time-consuming and more accurate for similar facial expression classification than other facial expression classification methods.

Keywords: occlude; HOG features; KC-FDDL dictionary learning; sparse representation; weighted residuals

引言

人脸表情是人类用来表达内心情感的一种最

普遍的非语言方式之一。人脸表情识别技术的目的主要是让计算机能够做到看人的情绪表情状态做出相应的决策, 从而营造真正和谐的人机环境。因此, 为了促进人与计算机之间可以更智能、和谐的相处和交流, 人脸表情识别在过去几十年得到广泛研究^[1-2], 并且成为一个热点。

伴随着压缩感知的发展, 国内外研究者已经在



收稿日期: 2015-12-02 修回日期: 2016-01-25;
基金项目: 国家自然科学基金(61262019);
作者简介: 黎明(1965-), 男, 江西南昌, 博士, 教授,
研究方向为智能计算, 图像处理与模式识别; 彭秀姣
(1990-), 女, 江西九江, 硕士生, 研究方向为图像处理与模式识别。

<http://www.china-simulation.com>

稀疏表示上进行了广泛的研究, 各方面研究也越来越完善。稀疏表示模型已被大量应用在图像处理在许多方面, 比如图像压缩、模式识别、信号检测等。

稀疏表示即使用稀疏的、少量的数据表示一个完整的信号。目前, 研究者们将稀疏表示的相关理论用在模式识别的分类中^[3-6], 该方法的依据是: 同一个特征空间中的所有特征向量属性是相同的, 可以相互线性表示, 即待测样本可以被其所属类别中的训练样本线性表示。基于稀疏表示的分类方法 (Sparse Representation based Classification, SRC) 是模拟生物视觉感知系统中神经元响应的稀疏特性提出的分类方法, 具有生物视觉系统一些特点即对遮挡具有一定的鲁棒性^[7]。例如, Wright 等^[7]通过模拟生物视觉响应机制, 提出基于稀疏表示的分类方法, 并证明了 SRC 能够解决人脸识别之中的遮挡问题; Cotter 等^[8]将 SRC 用于人脸表情识别中, 实验结果表明, SRC 相比近邻法、SVM 等分类方法在人脸表情图片有遮挡时可以取得很好的效果。

基于稀疏表示的分类方法关键之一就是字典的学习。欧阳琰等^[9]直接将图像的表情特征作为字典元素, 然后以该字典为基础进行表情分类, 虽然也取得了很好的实验效果, 但是所得字典可能存在信息冗余、噪声、稀疏编码耗时长以及不利于分类的信息等问题, 为此, 研究者们开始致力于字典的构建方法。Mohammadi 等^[10]提出基于 PCA 的字典构建方法, 将每一个人的表情图像减去各自的中性表情得到差分图像, 用 PCA 提取每一个子类的主成份脸作为字典元素, Cohn-Kanade 数据库上的实验结果表明此方法相比已经存在的字典学习方法需要更少的训练样本同时识别率提高了 6%。Aharon 等^[11]提出一种新的算法—KSVD 来设计过完备字典, 该方法主要是考虑训练样本能否恰当的被字典中的元素稀疏表示, 但该算法得到的字典并不适合分类。针对 KSVD 的缺点, 研究者也提出了各种字典学习方法^[12-14]。章强等^[15]提出 D-KSVD 的字典学习方法, 引进分类判别误差, 该字典表示的学习方法具有很好的表示能力同时也适用于分

类的。但是, 上面这些字典学习方法都是对所有类学习一个通用的字典, 并没有考虑字典中同类元素之间的关系、不同类之间的关系。为此, Zhang 等^[16]提出了 Fisher 判别字典学习, 在字典学习过程中, 加入 Fisher 判别准则, 在分类同时考虑重构残差和稀疏编码系数的判别信息进行分类。

虽然上面这些字典学习方法的分类效果比较理想, 但是这些字典学习方法中字典的大小会随着训练样本的增加而越来越大, 这样在稀疏编码时耗时也会随训练数据的增大而增加; 针对以上字典学习的不足, 本文提出了由 HOG 特征构成训练集, 在稀疏字典学习过程对训练集采用 Fisher 判别和改进的 K-均值聚类分析的字典学习方法, 最后用残差加权稀疏表示的方法进行表情分类。

1 人脸面部表情识别

1.1 基于 HOG 的面部表情特征提取

计算机视觉和图像处理中常用来进行目标检测的一种方法是方向梯度直方图 (Histogram of Oriented Gradient, HOG), 该方法也可用于图像的特征提取, HOG 特征是由计算和统计图像局部区域的梯度方向直方图构成的, HOG 特征不仅能够提取人脸中大量的形状信息来反应面部运动单元的变化, 而且由于 HOG 是在图像的局部区域上进行操作, 所以它对图像几何的和光学的变化具有鲁棒性^[17]。

HOG 特征能提取足够的表情信息, 并将表情相关的信息描述出来, 同时 HOG 具备了对不同光照的处理, 可以避免图像采集过程中光照对图像的影响, 因此采用 HOG 来提取图像的特征更适合 SRC 进行表情分类。

HOG 特征提取能够很好的运用梯度或边缘的密度分布将局部目标的表象和形状描述出来^[17]。具体的实现方法是: 首先将图像分成一小块一小块相连的区域, 把每一小块区域叫细胞单元, 然后采集每一小块区域中各像素点的梯度和边缘的方向直方图, 最后把这些直方图组合起来构成每个细胞

单元的特征^[18-19]。为了提高性能,还可以把每个细胞单元的直方图在图像中更大的范围内(比如数个细胞单元组成一个大块 block)进行对比度归一化,所采用的方法是:先计算每个区间(block)中各直方图的密度,然后根据所计算的密度在区间内对每个细胞单元做归一化。经过区间归一化后,可以降低图像光照变化和阴影对实验的影响。

HOG 面部表情特征的提取方法如图 1 所示。

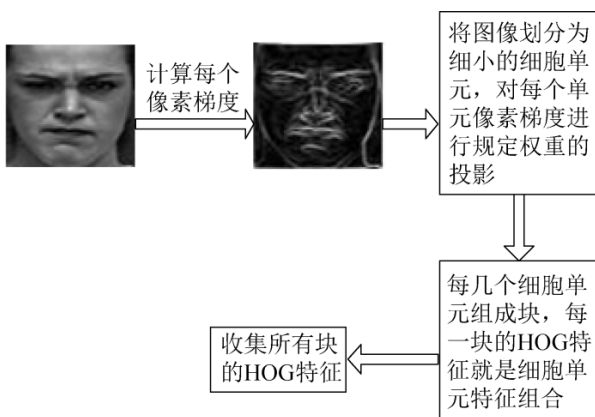


图 1 人脸面部表情 HOG 特征提取

Fig.1 Extraction of facial expression HOG features

1.2 基于稀疏表示分类的面部表情识别

稀疏表示分类的面部表情识别是用已标记的 6 类不同表情的训练样本来正确的判断待测样本属于哪一类,是将待测表情图像看作字典中训练图像集的线性组合,通过寻找最稀疏解来进行分类。字典的学习方法直接影响稀疏表示分类应用过程中的效果,学习所得的字典里面信息越丰富,字典中不同类间的差别越明显,基于稀疏表示的人脸表情识别就能取得好的分类结果^[20-21]。在人脸表情分类中,假设每一类的样本数为 $n_i(i=1, 2, 3, 4, 5, 6)$,所以,可以用训练集中的样本构建字典 D :

$$D = \{D_1, D_2, D_3, D_4, D_5, D_6\} \quad (1)$$

式中: $D_i = \{d_1^i, d_2^i, \dots, d_{n_i}^i\}$ 。

上面的 D_i 表示所有属于第 i 类的训练样本。假设测试样本为 y , 根据 SRC 原理, 可以将 y 表示为字典中元素的线性组合:

$$y = x_1^1 d_1^1 + x_2^1 d_2^1 + \dots + x_{n_1}^1 d_{n_1}^1 + \dots + x_1^6 d_1^6 + \dots + x_{n_6}^6 d_{n_6}^6 \quad (2)$$

理想情况下, y 仅仅只能被字典中第 i 类的样本线性表示, 即 D 中与 y 不同类的数据对应的系数应该为 0, 则上式可以写成

$$y = D x_0 \quad (3)$$

式中: $x_0 = [0, 0, \dots, x_1^i, x_2^i, \dots, x_{n_i}^i, \dots, 0, 0]$ 。

由于 x_0 是稀疏的, 所以可以通过求解 l_1 范数问题来得到 x_0 , 如下所示:

$$\hat{x} = \arg \min \|x\|_1 \text{ s.t. } D x = y \quad (4)$$

在具体应用中, y 不可能完全由字典中同类的样本线性表示, 字典中与 y 不同类的样本所对应的稀疏编码系数也可能是非零的。因此, 在判断时需比较稀疏编码重构后的图像与原图像的差值, 对应于重构误差最小的类别为该样本的所属类别, 即:

$$\text{identify}(y) = \arg \min_i \|y - D \delta_i(x)\|_2 \quad (5)$$

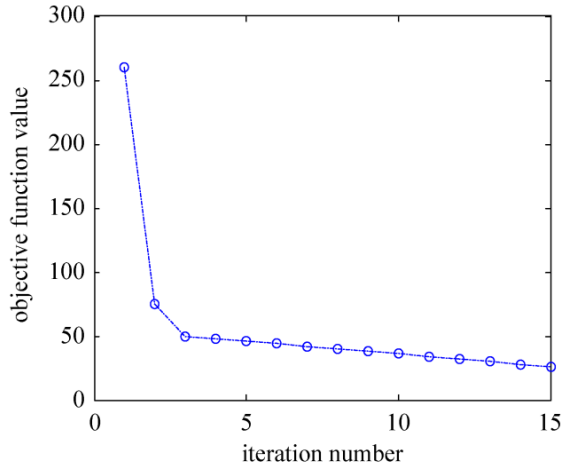
式中: $\delta_i(x) = [0, 0, \dots, x_1^i, x_2^i, \dots, \dots, 0, 0]$ 。

1.2.1 基于稀疏表示的聚类字典的构建

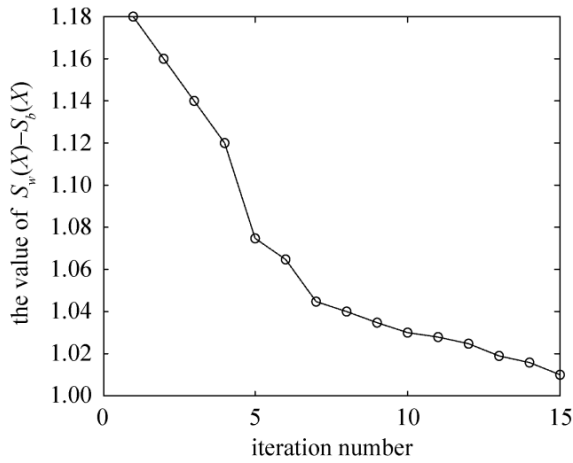
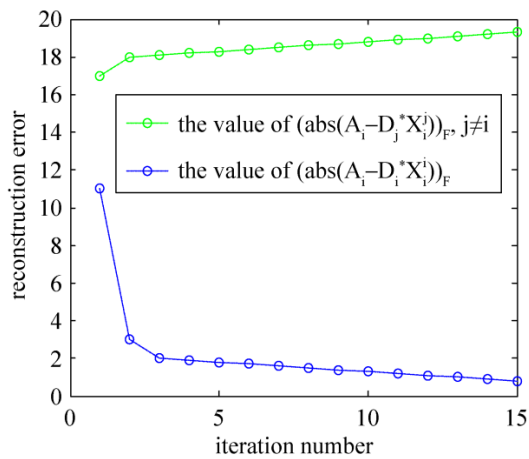
稀疏表示本质上是一种建模方法, 是指待测信号能被字典中的一些元素线性表示。因此, 字典的选择是稀疏表示的关键部分。字典的学习方法一般分为两种: (1) 基于一些数学模型来构建字典; (2) 通过训练样本和最小化目标函数, 从而得到最优化后的字典。基于数学模型得到的字典能适用于所有类型的图像, 但是对于特定类型的图片, 例如人脸表情图片, 却不是最佳的, 因此普遍采用第二种字典学习方法。

为改善以往字典学习方法性能, 通过训练样本及最小化目标函数, 提出基于聚类的字典学习方法。

本文先提取六类表情训练样本的 HOG 特征构成训练集 A , $A = [A_1, A_2, \dots, A_6]$, 并在该训练集上进行字典学习, 要想获得适合分类的字典, 则需对字典进行相应的约束, 除了要求字典具有较强的表示能力同时还因对不同表情类的样本具备较强的区分能力, 为此构建了如下的目标函数。其中, 目标函数 $J(D, X)$ 值、 $S_W(X) - S_B(X)$ 值、 $r(A, D, X)$ 的值随着迭代次数的增加而变化的曲线如图 2 所示。



(a) 目标函数的值随着迭代次数变化的曲线图

(b) $S_W(X) - S_B(X)$ 值的变化曲线图(c) $r(A, D, X)$ 中相关表达式的变化曲线图图2 $J_{(D, X)}$ 中相关表达式的值随着迭代次数的变化曲线
Fig.2 Change curve of the correlation expression in $J_{(D, X)}$ with the number of iterations

$$J_{(D, X)} = \arg \min \{ r(A, D, X) + \lambda_1 (S_W(X) - S_B(X)) + \lambda_2 \|X\|_1 \} \quad (6)$$

公式(6)中, λ_1, λ_2 是常量, $r(A, D, X)$ 代表字典对训练集的表示能力。其中 $D = [D_1, D_2, \dots, D_6]$ 表示的是六类表情构成的类别字典, D_i 是第 i 类的子字典; $A = [A_1, A_2, \dots, A_6]$ 是六类表情对应的训练集, A_i 是第 i 类训练样本的子集; $X = [X_1, X_2, \dots, X_6]$ 是六类表情对应的稀疏编码系数, X_i 是相应的编码系数。于是第 i 类表情图片的训练集 A_i 应该能被字典 D 线性表示, 即

$$A_i \approx DX_i = D_1 X_i^1 + \dots + D_i X_i^i + \dots + D_6 X_i^6 \quad (7)$$

由于 D_i 是第 i 类对应的字典, 理想情况下 A_i 应该能被 D_i 而不是 $D_j (j \neq i)$ 线性表示, X_i^j 应该接近零, 因此根据上文所述可以得到式(8)。

$$r(A, D, X) = \sum_{i=1}^6 \left(\|A_i - DX_i\|_F^2 + \|A_i - D_i X_i^i\|_F^2 + \sum_{j=1, j \neq i}^6 \|D_j X_i^j\|_F^2 \right) \quad (8)$$

为了改善字典中各类元素之间以及同类元素之间的关系, 可以在稀疏编码过程中在系数上加入 Fisher 判别准则, 根据 Fisher 判别准则可以写出各类表情稀疏编码系数的类内离散度矩阵和类间离散度矩阵, 如公式(9)、(10)所示, 其中 m_i 是第 i 类的稀疏编码系数 X_i 的平均值, m 是稀疏编码系数 X 的平均值, n_i 是第 i 类样本的个数, $S_W(X)$ 是每一类表情稀疏编码系数的总类内离散度矩阵, $S_B(X)$ 是六类表情的类间离散度矩阵。

$$S_W(X) = \sum_{i=1}^6 \sum_{x_k \in X_i} (x_k - m_i)(x_k - m_i)^T \quad (9)$$

$$S_B(X) = \sum_{i=1}^6 n_i (m_i - m)(m_i - m)^T \quad (10)$$

为了使同类表情对应的字典元素更紧凑, 不同类字典元素之间差别更明显, 可使稀疏编码系数有小的类内离散度大的类间离散度, 即 $S_W(X) - S_B(X)$ 取最小值。

为了减少稀疏编码计算时间, 本文在字典学习过程进行改进的 K -均值聚类分析, 先求出每类表情子类各自的均值作为各自聚类中心, 找出每个子字典中前 k 个最接近的元素, 因此, 由 6 类表情样本训练学习后的字典大小不会太大。

结合 K -means 聚类算法, 针对聚类中心选择的

随机性缺点, 本文选取每一类的均值作为聚类中心, 计算每一类样本到均值中心的距离, 找出前 k 个离均值中心最近的样本作为最终的选择。对 Fisher 判别学习后的字典中的每一个子类对应的字典进行改进的 K -means 聚类分析。首先计算每个子类的均值, 把此均值作为聚类中心; 然后对每一个字典元素, 计算其与聚类中心的距离来, 选出前 k 个离均值中心最近的元素作为最终学习得到的字典。通过上述字典学习方法得到的字典不仅可以很好的线性表示待测样本同时可以减少稀疏编码的时间。

上面所述字典学习方法(Improved K-means Cluster and Fisher Discrimination Dictionary Learning, IKC-FDDL)可以分为3个方面的学习: 固定字典 D 更新系数 X 、固定系数 X 更新字典 D 、改进的 K-均值聚类字典学习。

目标函数 $J_{(D,X)}$ 对 (D,X) 不是联合凸的, 但是当固定其中一个时, J 关于 D 或 X 是凸的。因此目标函数 $J_{(D,X)}$ 可以分为两个子问题: (1)固定 D , 更新 X ; (2)固定 X , 更新 D 。通过迭代获得表示能力、区分能力较强的字典 D 。

首先, 假设字典 D 是固定的, $J_{(D,X)}$ 转化为求解稀疏编码系数 X , 分类计算每一类对应的系数 X_i , $i=1,2,\dots,6$, 当计算 X_i 时, 固定 X_j , $j \neq i$ 。因此公式(6)中的目标函数可简化为公式(11), 通过求解公式(11)得到每一个 X_i 。

$$J_{(X_i)} = \arg \min \{r(A_i, D, X_i) + \lambda_1(S_W(X_i) - S_B(X_i)) + \lambda_2 \|X_i\|_1\} \quad (11)$$

假设前面得到的系数 X 是固定的, 分类更新 D_i , 计算 D_i 时, 固定 D_j , $j \neq i$, 此时公式(6)中的目标函数可简化为:

$$J_{(D_i)} = \arg \min \{r(A, D, X)\} = \left\{ \|A - D_i X^i - \sum_{j=1, j \neq i}^6 D_j X^j\|_F^2 + \|A_i - D_i X^i\|_F^2 + \sum_{j=1, j \neq i}^6 \|D_i X_j^i\|_F^2 \right\} \quad (12)$$

最后, 计算上面所得字典 D 中每个类所对应字典 D_i 的均值, 此值作为每类的聚类中心, 找出前 k 个离均值中心最近的样本作为最终的字典

元素选择, k 的选取由每个表情训练样本的个数决定, 经过上面3个步骤可以得到最终的字典 $D_{ultimate}$ 。

IKC-FDDL 过程可以归纳为如下几个步骤:

步骤 1: 随机产生每一类字典对应的各个元素来初始化字典 D ;

步骤 2: 固定 D , 通过迭代过程求解稀疏编码系数 X ;

步骤 3: 固定步骤 2 得到的系数 X , 求解字典 D ;

步骤 4: 返回步骤 2, 直到 $J_{(D,X)}$ 达到最小值或者达到最大的迭代次数;

步骤 5: 对步骤 4 得到的字典中的每一类字典进行改进的 K-均值聚类字典学习, 优化每一类字典中元素的个数, 得到字典 $D_{ultimate}$ 。

1.2.2 基于稀疏表示的面部表情分类

通过稀疏表示分类算法进行分析, 可以发现: 在残差计算公式中, 残差的值仅随稀疏编码系数的变化而变化的, 即真正决定残差大小的因素是系数向量 x , 所以本文对残差用各类系数的 L_2 范数进行加权, 得到新的分类方法, 如下所示:

$$identify(y) = \arg \min_i \frac{\|y - D_{ultimate} \delta_i(x)\|_2}{\|\delta_i(x)\|_2} \quad (13)$$

综上所述, 基于改进的 K-均值聚类和 Fisher 判别字典学习与稀疏表示的人脸表情识别方法如下所示:

1) 输入一组标记好六类表情的训练样本 $A = \{A_1, A_2, A_3, A_4, A_5, A_6\}$;

2) 输入一个测试样本 y , 分别提取训练样本和测试样本的 HOG 特征, HOG 特征的提取可以看作是样本空间到特征空间的投影;

3) 通过 IKC-FDDL 方法对提取 HOG 特征后的训练样本进行字典学习, 得到字典 $D_{ultimate}$;

4) 计算稀疏编码系数 $\hat{x} = \arg \min \|x\|_1 \text{ s.t. } D_{ultimate} x = y \quad (14)$

5) 决定测试样本所属的类别

$$identify(y) = \arg \min_i \frac{\|y - D_{ultimate} \delta_i(x)\|_2}{\|\delta_i(x)\|_2} \quad (15)$$

6) 重复步骤 2 到步骤 5 直到所有的测试样本都被分类。

2 实验结果

为了评估和验证本文所提方法的识别性能, 本文采用 Cohn-Kanade 人脸表情数据库进行实验, 每种表情图片随机选择 30 幅训练样本构建字典, 剩下的用来做测试样本, 验证本文所提方法的性能。实验配置是 Intel Core i5, 主频 2.53 GHz, 实验仿真软件是 matlab2010。

2.1 Cohn-Kanade 数据库介绍

Cohn-Kanade 数据库是 CMU 机器人研究所和心理系于 2000 年共同建立的人脸表情数据库, 它包括 100 个成年人(包括女性, 非裔美洲人, 拉丁美洲人, 亚洲人)近 500 张不同表情序列^[22]。在预处理阶段, 将每张图像大小归一化为 64×64 。经预处理后的图像如图 3 所示, 从左到右依次为: 生气、厌恶、恐惧、高兴、悲伤、惊奇。在每类表情中随机选取 20 个人, 每个人 3 种不同强度的表情作实验的数据集。在此数据集中, 每类随机选取 10 个人的三张表情作为训练集, 剩下的作为测试集。



图 3 预处理后的 Cohn-Kanade 库的人脸表情图像示例
Fig. 3 An example of the facial expression image of the pre processed Cohn-Kanade Library

2.2 HOG 参数选取

文献[9]指出有两个重要的参数直接影响 HOG 特征的提取。一个是图像分块时单元格的选取, 另一个是直方图的方向选取。为了获得最佳参数值, 本文针对文献[9]的方法进行了实验, 实验结果如图 4 所示, 由图 4 可知, 当图像分割成 961 个单元格和直方图方向分为 9 时, 人脸表情识别方法能取得较高的识别率, 因此下面的实验都是在这个最佳参数下进行的。

2.3 无遮挡图片的实验结果

在经典的稀疏编码中, 随着训练样本的增加字典的大小也会增大, 这样稀疏编码时所消耗的时间也会增加, 算法的实践性差, 因此, 字典学习方法将影响算法的运行时间。表 1 显示了所有测试图片稀疏编码时所消耗的时间, 从表 1 可以看出本文提出的方法在保证不降低识别率的前提下可以减少稀疏编码所耗的时间。

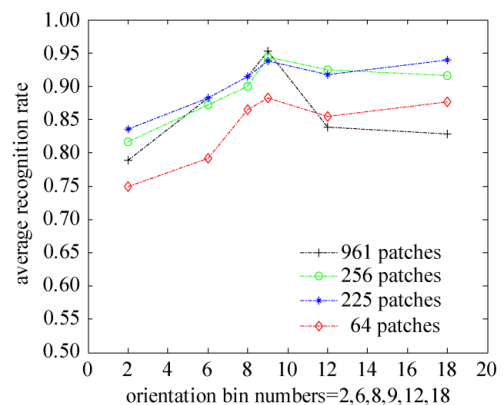


图 4 不同 HOG 参数下的表情正确识别率
Fig. 4 Correct recognition rate of facial expressions under different HOG parameters

表 1 不同的字典学习方法所耗时间对比

Tab.1 Time comparison of different dictionary learning methods

学习方法	时间/s	识别率/%
FDDL+SRC ^[16]	119.22	94.56
本文	71.01	95

在无遮挡图片的实验中, 给出了本文在最优的参数及改进的稀疏编码分类下对无遮挡的面部表情图像的识别结果, 如表 2 所示, 表 2 中左右两个子表中主对角线上的值表示每类表情的正确识别率, 非主对角线上的值表示表情之间的错判率。由表 2 可以看出本文方法在处理高兴、惊讶等表情时能够取得非常显著的效果。

实际进行表情分类时, 待分类的表情数据跟已有的训练数据之间不一定有某种关联, 表 2 的结果是在这样的情况下进行实验的, 即训练样本和测试样本取自不同人的脸表情数据, 训练样本和测试样本之间是无联系的, 这样的实验测试具有实用性。

表 2 经典 SRC 和改进的 SRC 对无遮挡表情图像的识别率
Tab. 2 Recognition rate of unshielded expression images by classical SRC and improved SRC /%

表情类型	HOG+SRC 下表情识别率						HOG+改进的 SRC 下表情识别率					
	Ang	Dis	Fea	Happ	Sad	Sur	Ang	Dis	Fea	Happ	Sad	Sur
Anger	80	10			10		90				10	
Disgust	10	83.33			6.67		3.33	86.67			10	
Fear			93.33	6.67					93.33	6.67		
Happiness				100						100		
Sadness	3.33				96.67		3.33				96.67	
Surprise						100						100

从图 5 中可看出生气和厌恶表情之间的相似度比较高,同时表 2 中两种方法实验结果的对比也显示这两个表情之间会相互错判,因此为了提高实验的识别率,本文采用基于残差加权的稀疏编码方法,相比经典的方法,此方法可以减少相似表情之间的错判率。



图 5 前三张是生气表情、后三张为厌恶表情
Fig. 5 The first three were angry expressions and the other three were disgust expressions

同时将本文中的特征提取方法同其它已经存在的特征提取方法在 SRC 下进行实验对比,通过表 3 中的结果对比可以发现特征提取方法对基于稀疏表示的表情分类方法有着非常大的影响,只有提取适合 SRC 的特征才能获得最佳的识别率。

2.4 图片有遮挡时的实验结果

在这一小节测试是在图片有遮挡时的识别效果。我们采用两种遮挡方法,即眼睛和嘴巴分别遮挡下的实验,并且对眼睛和嘴部进行 3 种不同程度的遮挡: small, medium, large, 如图 6 所示。表 4 给出了眼睛和嘴巴遮挡下的最终识别结果。

表 3 HOG 特征与其它特征提取方法的对比
Tab. 3 Comparison of HOG features and other feature extraction methods /%

方法	识别率	方法	识别率
Gabor+SRC	82.54	Laplacian+SRC	71.48
Eigen+SRC	79.34	本文	95
Fisher+SRC	34.93		



图 6 人脸指定遮挡的示意图
Fig. 6 An example of face specified occlusion

从表 4 第 3 行数据可以看出,用 HOG 特征并结合改进的 SRC 的分类方法对有遮挡下的表情识别率比其他方法有显著提高。

表 4 图片有遮挡下的正确识别率
Tab. 4 Correct recognition rate under occlusion

表情识别方法	E-small	E-medium	E-Large
Gabor+SRC	76.32	68.23	60.23
HOG+kNN	85	83.89	81.67
本文	93.33	91.11	89.44
表情识别方法	M-small	M-medium	M-large
Gabor+SRC	72.12	63.12	59.36
HOG+kNN	80.56	77.22	75.56
本文	94.44	92.78	91.11

3 结论

本文针对以往的字典学习方法中未考虑字典的大小这一因素,提出改进的 K-均值聚类和 Fisher 判别字典学习与稀疏表示的人脸表情识别方法。该方法首先使用 Fisher 判别对训练集进行字典的初始学习,然后针对初始学习后的每一类表情的子字典采用改进的 K-均值算法找出前 k 个最近邻的元素作为新的子字典,最后将每一类人脸表情对应的子字典组合成基于稀疏表示分类时所用的字典。Cohn-Kanade 数据库上的实验结果表明本文的方法在取得较好的识别率的同时所耗的时间较少,在

稀疏表示分类时采用残差加权的思想, 减少了相似表情的错判率。

参考文献:

- [1] 薛雨丽, 毛峡, 郭叶. 人机交互中的人脸表情识别研究进展[J]. 中国图象图形学报, 2009, 14(5): 764-772. DOI:10.11834/jig.20090503.
XueYuli, MAO Xi, GUO Ye. The Research Advance of Facial Expression Recognition in Human Computer Interaction[J]. Journal of Image & Graphics, 2009, 14(5): 764-772. DOI:10.11834/jig.20090503.
- [2] 蒋斌, 夏克斌, 杨国胜. 人脸表情识别研究进展[J]. 计算机科学, 2011, 38(4): 25-31.
JIANG Bin, XIA Kebin, YANG Guosheng. Research Advance of Facial Expression Recognition[J]. Computer Science, 2011, 38(4): 25-31.
- [3] 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 Minghan, 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.
- [4] Liu W, Song C, Wang Y. Facial expression recognition based on gabor features and sparse representation[C]// 2012 12th International Conference on Control Automation Robotics & Vision (ICARCV), IEEE, 2012: 1402-1406.
- [5] Zhen W, Zi L Y. Facial expression recognition based on local phase quantization and sparse representation[C]// 2012 Eighth International Conference on Natural Computation (ICNC), IEEE, 2012: 222-225.
- [6] Huang M W, Wang Z W, Ying Z L. A new method for facial expression recognition based on sparse representation plus lbp[C]//2010 Third International Congress on Image and Signal Processing (CISP), IEEE, 2010, 4: 1750-1754.
- [7] Wright J, Yang A Y, Ganesh A. Robust face recognition via sparse representation[C]. IEEE Trans. Pattern Anal. Mach. Intell, 2009, 31(2): 210-227.
- [8] Cotter S F. Sparse representation for accurate classification of corrupted and occluded facial expressions[C]//2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), IEEE, 2010: 838-841.
- [9] Ou Y Y. Accurate and robust facial expressions recognition by fusing multiple sparse representation-based classifiers[J]. Neurocomputing (S0925-2312), 2014, 73(3): 71-78.
- [10] Mohammadi M R, Fatemizadeh E, Mahoor MH. PCA-based dictionary building for accurate facial expression recognition via sparse representation[J]. J. Vis. Commun. Image R(S1047-3203), 2014, 25(5): 1082-1092.
- [11] Aharon M, Elad M, Bruckstein A. K-svd: an algorithm for designing overcomplete dictionaries for sparse representation[C]. IEEE Trans. Signal Process, 2006, 54(11): 4311-4322.
- [12] Mairal J, Bach F, Ponce J. Learning discriminative dictionaries for local image analysis[C]// CVPR, 2008.
- [13] Yang M, Zhang L, Yang J. Metaface learning for sparse representation based face recognition[C]//2010 17th IEEE International Conference on Image Processing (ICIP), IEEE, 2010: 1601-1604.
- [14] Rubinstein R, Bruckstein A M, Elad M. Dictionaries for sparse representation modeling[C]. Proc. IEEE, USA: IEEE, 2010, 98(6): 1045-1057.
- [15] Zhang Q, Li B. Discriminative k-svd for dictionary learning in face recognition[C]// 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), USA: IEEE, 2010: 2691-2698.
- [16] Yang M, Zhang L, Feng X. Fisher discrimination dictionary learning for sparse representation[C]//2011 IEEE International Conference on Computer Vision (ICCV), IEEE, 2011: 543-550.
- [17] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, Montbonnot, France, USA: IEEE Press, 2005: 886-893.
- [18] 罗元, 崔叶, 王艳. 基于融合 DCT 和 LBP 特征提取的面部表情识别[J]. 半导体光电, 2014, 35(2): 330-334.
LUO Yuan, CUI Ye, WANG Yan. Face Expression Recognition Based on Fusion Feature of DCT and LBP[J]. Semiconductor Optoelectronics, 2014, 35(2): 330-334.
- [19] 李雅倩, 李颖杰, 李海滨. 融合全局与局部多样性特征的人脸表情识别[J]. 光学学报, 2014, 34(5): 1-7.
Li Yaqian, Li Yingjie, Li Haibin. Fusion of Global and Local Various Feature for Facial Expression Recognition [J]. Journal of Acta Optica Sinica, 2014, 34(5): 1-7.

(下转第 44 页)