

# 基于一对一极限学习机的人脸表情识别方法<sup>①</sup>

张小庆, 于威威

(上海海事大学 信息工程学院, 上海 201306)

**摘要:** 为了克服极限学习机(ELM)稳定性差、识别率不高的缺陷, 利用支持向量机(SVM)一对一投票式分类算法准确度高的优势, 提出一种改进的表情识别方法. 该方法将一对一分类算法和 ELM 算法相结合形成一个新的算法即 OAO-ELM(One-Against-One-Extreme Learning Machine), 首先, 对样本采用一对一的分类并利用 ELM 训练成一个弱分类器, 然后, 将这些弱分类器组合成一个最终的强分类器. 预测结果, 采用投票方式. 用 Gabor 滤波提取表情特征, 由于提取后特征维度很高, 冗余大, 引入主成分分析(PCA)来降维. 基于 JAFFE 数据库实验结果表明, 该算法在人脸表情识别上具有较高分类识别率和稳定性.

**关键词:** 极限学习机; 支持向量机; 一对一; Gabor 滤波; 主成分分析

## Facial Expression Recognition Method Based on One-Against-One Extreme Learning Machine

ZHANG Xiao-Qing, YU Wei-Wei

(Information Engineering College, Shanghai Maritime University, Shanghai 201306, China)

**Abstract:** By using a vote of one-against-one Support Vector Machine advantages of high classification algorithm accuracy, an improved expression recognition method was proposed in order to modify the Extreme Learning Machine's disadvantage of bad stability and poor classification accuracy. The method combines one-against-one classification algorithm with Extreme Learning Machine, which are consist of a new algorithm—OAO-ELM. First, the algorithm uses the ELM process classification as weak classifier when training sample by one-against-one. Then, these weak classifiers are combined into the finally strong classification. Prediction the results of classification, by votes to the class. Gabor facial expressional features, since the high-dimensional Gabor features are redundant; The dimensional principal component analysis is used to select these features. Experimental results based on the JAFFE database show that it obtains higher accuracy and better stability.

**Key words:** extreme learning machine; support vector machine; one-against-one; Gabor filter; PCA

## 1 引言

人脸表情包含着非常丰富的信息, 是人们日常交流情感有效、直接的方式. 随着人机交互、模式识别等技术的迅速发展, 人脸表情识别现已成为一个国际的研究热点. 人脸表情识别涉及的邻域有计算机视觉、图像处理、心理学、模式识别等, 并在网络教学、安全驾驶、人机交互等方面都有广泛和重要的应用. 美国心理学家 Ekman<sup>[1]</sup> 和 Friesen 对人的面部表情进行了分类, 共 6 种基本表情情况: 高兴(happiness)、愤怒(anger)、厌恶(disgust)、恐惧(fear)、悲伤(sadness)

和惊奇(surprise). 人脸表情识别主要过程可以分为预处理、特征提取和分类 3 个步骤.

最常用的人脸表情分类方法主要有 K 最近邻(KNN, k-Nearest Neighbor)算法、神经网络算法、支持向量机(SVM, Support Vector Machine)算法等. 其中 KNN 算法简单有效, 但是当不同类的样本数量不均衡时, 识别性能较差; ELM(ELM, Extreme Learning Machine)算法训练速度快, 泛化性能好, 但是隐含层参数均随机产生这就会造成分类结果的每次都不一样, 影响整体的识别率和稳定性; SVM 算法<sup>[2]</sup>准确度高,

<sup>①</sup> 收稿时间:2015-01-20;收到修改稿时间:2015-03-18

泛化性能好,但训练算法速度慢。

一对一分类算法是 MSVM 分类方法中的一种,由参考文献[3],可以验证出一对一分类算法整体性能较高,适合处理样本类别较少的,并且能达到很好的分类效果。因此,本文针对 ELM 存在的缺点,提出一种基于一对一投票式的极限学习机算法(OAO-ELM)。比单独的极限学习机(ELM),具有更高的识别率,且稳定性更好;与 SVM 相比,识别的速度快。

为了验证 OAO-ELM 算法的有效性,本文首先采用 Gabor 滤波特征提取方法;其次,利用 PCA 对特征进行选择;最后 KNN、SVM、ELM、OAO-ELM 的分类算法进行比较验证。

## 2 人脸表情特征提取

目前,人脸表情的特征提取的方法可以分为局部特征提取<sup>[4]</sup>和整体特征提取<sup>[5]</sup>。局部特征提取面部特征的位置、大小以及相对位置,但该方法在很大程度上减少了需要处理的数据,而且代表人脸图像的特征点也有限,导致一些重要的表情识别和分类信息丢失;而整体特征提取是从整个人脸图像出发,提取人脸整体的统计特征。因此,本文采用广泛应用的基于整体特征的 Gabor 小波算法。

### 2.1 二维 Gabor 小波特征提取

二维 Gabor 小波与人的视觉皮层简单细胞反射区有着相同的特性,它能够获得较高的空域和频域的分辨率,同时具有很好的空间局限性和方向选择性,以及正交相位关系的特性。二维 Gabor 滤波器<sup>[6]</sup>是一个复函数,它由实部和虚部组成,其滤波器的核函数形式定义如下<sup>[7]</sup>:

$$\psi_{u,v} = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left[-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right] \left[ \exp(i k_{u,v} z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (1)$$

其中,  $u$  和  $v$  分别表示 Gabor 滤波器的方向和尺度;  $k_{u,v}$  表示滤波器的中心频率,并且  $k_{u,v} = k_v \exp(i\varphi_u)$ , 而  $k_v = k_{max}/f^v$ ,  $\varphi_u = u\pi/8$ ,  $k_{max}$  表示最大频率,  $f$  表示频域里该函数的空间因子;  $\sigma$  表示高斯半径;  $z = (x, y)$  表示图像像素坐,高斯函数函数可以表示为  $\exp\left(-\|k_{u,v}\|^2 \|z\|^2 / 2\sigma^2\right)$ ; 函数  $\exp(i k_{u,v} z)$  表示振荡函数; 而函数为  $\exp(-\sigma^2/2)$  表示滤波器对光照时能有一定的容忍度。

本文先对原图像采用双线性插值法及灰度归一化后的大小是  $128 \times 128$  的像素,接着采用 Gabor 特征的下采样方法,对每幅图进行下采样的处理,采样因子  $\rho=8$ ,这样得到的每幅图的维数为  $256(16 \times 16)$ 。最后,采用 Gabor 小波的算法对每幅图像进行特征提取。

文献[8]的实验结果,本文采用 5 个频率和 8 个方向的 Gabor 滤波器提取人脸特征。根据式(1)的定义可知, Gabor 小波变量  $u=0, \dots, 4; v=0, \dots, 7$ , 得到 40 个 Gabor 小波系数图。

由于本文采用 5 个频率、8 个方向的 Gabor 小波函数,得到 40 个 Gabor 小波系数图,这样得到的图像 Gabor 特征维数是原灰度图像维数的 40 倍。如果直接利用 Gabor 算法得到的特征集合后进行分类,就会容易陷入“维数灾难”的影响。因此,有必要采用主成分分析(PCA)方法来进行特征降维<sup>[9]</sup>。

设人脸表情图像由 Gabor 滤波提取后的分辨率为  $m \times n$ , 则第  $i$  个训练样本可以表示为  $m \times n$  的图像,提取后的矩阵  $X_i (i=1, \dots, N)$ ,  $N$  表示训练样本的个数。则训练样本的均值为:

$$\bar{X} = \sum_{i=1}^N X_i / N \quad (2)$$

对训练样本的总体散布矩阵进行估计如式(3):

$$S_x = \frac{1}{N} \sum_{i=1}^N (X_i - \bar{X})^T (X_i - \bar{X}) \quad (3)$$

通过计算  $S_x$  的特征值和特征向量,选取前  $N-1$  个较大的特征值所对应的特征向量  $\alpha$  作为投影轴,即  $Y = [\alpha_1, \alpha_2, \dots, \alpha_l]$ 。  $X_i$  的特征矩阵  $Z_i \in R^{m \times l}$ , 可以表示为

$$Z_i = X_i \quad (4)$$

## 3 各模块的算法设计与实现

### 3.1 极限学习机算法

极限学习机是由 Huang G.B. 等人<sup>[10]</sup>于 2006 年提出,它是一种新型神经网络算法,即是一种快速单隐层神经网络训练算法。该算法隐含层参数均随机产生,输出的权重则通过最小二乘法直接计算得出。与 SVM、传统神经网络相比,ELM 算法<sup>[11]</sup>表现出更快的学习速度和更好的泛化能力,这是由于该算法的隐含层无需迭代,且整个训练过程是全自动的无监督的。

文献[12]给出  $N$  个不同样本的矩阵  $(x_i, y_i)$ ,  $x_i = [x_{i1}, x_{i2}, \dots, x_{in}] \in R^n$ ,  $y_i = [y_{i1}, y_{i2}, \dots, y_{im}] \in R^m$

其中  $x_i$ : 输入样本矩阵,  $y_i$ : 输出样本矩阵,  $i=1,2,\dots,N$ . 隐含层节点的数目为  $L$ , 激活函数为  $g(x)$ , 则 ELM 算法的网络结构可表示为:

$$\sum_{i=1}^L \beta_i g_i(x_j) = \sum_{i=1}^L \beta_i G(a_i, b_i, x) = y_i \quad (5)$$

其中,  $j=1,2,\dots,N$ ,  $a_i$  和  $b_i$  为隐含层节点参数,  $a_i = [a_{i1}, a_{i2}, \dots, a_{im}]$  表示输入节点与第  $i$  个隐含层节点的连接权重向量,  $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]$  表示第  $i$  个隐含层节点与输出节点的连接权重向量,  $b_i$  表示第  $i$  个隐含层节点的偏移值.

由(5)式可知,  $G(a_i, b_i, x)$  为隐含层神经元的输出, 针对加法型的隐含层节点  $G$  为

$$G(a_i, b_i, x) = g(a_i \cdot x_i + b_i) \quad (6)$$

由(5)式表示的  $N$  个等式用矩阵表示为:

$$H\beta = Y \quad (7)$$

其中,

$$H = \begin{bmatrix} g(a_1 \cdot x_1 + b_1) & \cdots & g(a_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(a_1 \cdot x_N + b_1) & \cdots & g(a_L \cdot x_N + b_L) \end{bmatrix}_{N \times L}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}$$

$H$  表示神经网络的隐含层输出矩阵, 由 Moore-Penrose 广义逆定理, 首先将  $H$  进行满秩分解得:

$$H = FG \quad (F \in C_r^{N \times r}, G_r^{r \times L})$$

通过奇异值分解求得  $H^+$ :

$$H^+ = G^H (GG^H)^{-1} (F^H F)^{-1} F^H$$

可以得出:  $\beta = H^+ T$

ELM 算法描述过程如下:

**ELM 算法:**

Input: 给定训练样本  $\{(x_i, y_i)\} \subset R^n \times R^m$ , 其中  $i=1,\dots,N$ , 隐含层输出函数  $G(a,b,x)$  和隐含层节点个数  $L$ .

a) 随机生成隐含层节点参数

$(a_j, b_j), j=1,\dots,L;$

b) 计算隐含层输出矩阵  $H$ ;

c) Output: 网络最优权  $\beta = H^+ T$ .

**3.2 一对一 SVM 分类算法(One-against-one, OAO)**

$$\min_{w^{ij}, b^{ij}, \xi^{ij}} \frac{1}{2} w^{ij} \cdot w^{ij} + c \sum_{i=1}^l \xi_i^{ij} \quad \text{一对一 SVM 分类方法,}$$

由 Knerr 提出<sup>[13]</sup>. 该方法是在  $k$  类问题中进行两两组合, 构造  $k(k-1)/2$  个二分类器, 每个分类器只对两类数据进行训练. 训练区分第  $i$  类和第  $j$  类样本之间的支持向量机, 需要解决如下最优化问题:

$$\text{s.t.} \begin{cases} w^{ij} \cdot \phi(x_i) + b^{ij} \geq 1 - \xi_i^{ij} & \text{if } y_i = i \\ w^{ij} \cdot \phi(x_j) + b^{ij} \leq -1 + \xi_j^{ij} & \text{if } y_j = j \\ \xi_i^{ij} \geq 0, & i, j = 1, 2, \dots, l \end{cases} \quad (8)$$

用第  $i$  类和第  $j$  类的数据结构分类器的决策函数为,

$$f_{i,j}(x) = (w^{i,j})^T \phi(x) + b^{i,j} \quad (9)$$

构造完  $k(k-1)/2$  个分类器后, 对未知样本进行预测采用投票法. 该方法首先将类中每个类别的初始票数均设置为 0; 然后对样本  $x$ , 依次带入  $k(k-1)/2$  个分类器进行判别, 如果分类器  $C_{i,j}$  认为样本  $x$  属于类别  $i$ , 则将类别  $i$  的票数加 1, 否则就认为  $x$  属于类别  $j$ , 则将类别  $j$  的票数加 1; 最后, 对  $k(k-1)/2$  个分类器都判别完后, 票数最高的那个类别就是样本  $x$  所属的类别.

**3.3 OAO-ELM 算法**

由于 ELM 算法与传统的神经网络及 SVM 相比训练速度非常快、需要人工干扰较少、泛化能力强等优点. 因此本文将 ELM 作为表情分类的分类器, 并与一对一的 SVM 算法相结合, 形成一个新的分类算法即 OAO-ELM. 该算法主要分为两个步骤, 第一步是对训练集进行一对一的分类, 得到  $p = k(k-1)/2$  个不同的训练子集, 利用这些训练子集在 ELM 分类器上训练, 得到  $p$  个不同的分类器. 第二步, 是利用多数投票策略对未知样本预测.

具体算法的描述如下:

**训练过程:**

输入: 训练样本集  $S$ ;

输出: 样本集成的分类器的模型  $M$ .

①对训练样本集  $S$  进行分类:  $S = \{S_1, S_2, \dots, S_k\}, (k$ : 样本的类别数目);

② for  $i=1$  to  $k-1$ , for  $j=i+1$  to  $k$ ;

③每次只选择2类样本,如  $S_i, S_j$ , 在 ELM 上训练得到成员神经网络分类器  $m_i$ ;

④保存集成的分类器模型  $M = \{m_1, m_2, \dots, m_p\}$ .

**预测过程:**

输入: 测试样本  $t$  集成分类模型  $M = \{m_1, m_2, \dots, m_p\}$

输出: 测试样本  $t$  的类别号  $i$

① for  $i=1$  to  $k$ ;

②计算样本  $t$  在  $m_i$  中的类别输出  $c_i$ ;

③对  $C = [c_1, c_2, \dots, c_k]$ , 进行投票式判别, 输出样本的类别号, 得票最多的判给那一类或者票数相同就判给类别号小的, 作为最终的分类结果.

**4 实验的描述与结果分析**

当前, 表情分类<sup>[14]</sup>方法有多种, 常见有 SVM、基于 K 近邻法则(KNN)等. 本实验在 Gabor 滤波提取特征的基础上, 比较 KNN、SVM、ELM 与本文提出的 OAO-ELM 方法在表情识别上的正确率.

**4.1 实验条件**

本文实验选择 JAFFE(Japanese Female Facial Expression)日本女性的人脸表情库进行训练与测试, 数据库包含7种表情组成213张图片, 例如生气, 厌恶, 害怕, 高兴, 中性, 悲伤和吃惊构成10位日本女性图片, 每人每种表情2-4幅, 表情库如图1所示. 从中选择每人每种表情中的一张表情图像共70张作为测试样本, 其余为训练样本, 循环3次, 取平均值作为识别率的结果.



图1 JAFFE 数据库下的部分人脸表情库

**4.2 实验流程图**

实验中通过预处理包括对表情图像进行旋转, 灰度归一化和双线性插值法后的大小为  $128 \times 128$ , 接着采用 Gabor 滤波下采样和特征提取后的特征维数为  $16 \times 16 \times 40$ , 相比直接采用 Gabor 小波维数降低了64倍, 因此可以降低数据处理时间和计算的复杂性. 接着对上述的得到的特征向量采用 PCA 进一步进行降维处

理. 最后, 对获得的表情特征向量进行分类, 所得到的实验结果如下图所示. 其中本文 ELM 和 OAO-ELM 的隐含层神经元的个数设置在 100-400 之间, 选择其中隐含层节点数识别率最高的作为实验结果的比较. SVM 选择的核函数为高斯函数, 惩罚因子 C 取 1000. 实验中的流程图如图2所示.

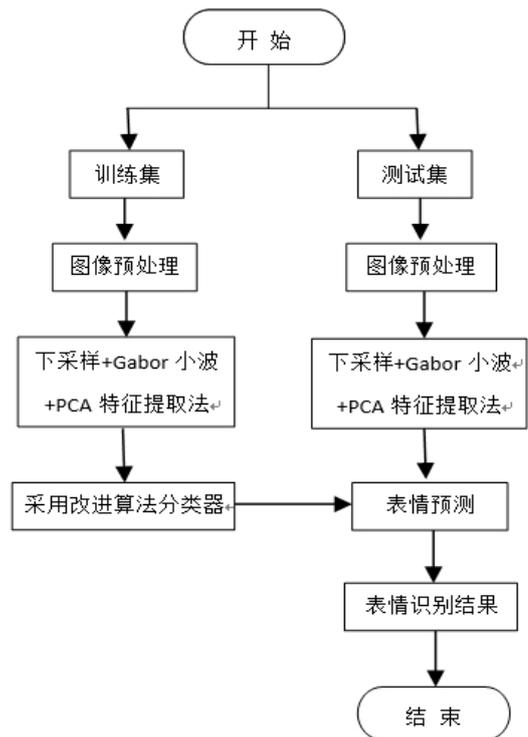


图2 表情识别流程图

**4.3 结果分析**

表1 JAFFE 数据库下不同方法的识别率(%)比较

表情	KNN	SVM	ELM	本文方法
生气	83.3	93.3	93.3	<b>96.7</b>
厌恶	80	93.3	90	93.3
恐惧	<b>80</b>	90	86.7	90
开心	90	100	96.7	100
中性	90	100	100	100
伤心	86.7	96.7	93.3	<b>100</b>
惊讶	83.3	96.7	93.3	96.7
平均识别率	84.7	95.7	93.3	<b>96.6</b>

通过实验可以发现, 表1说明了 OAO-ELM 的识别率高于 KNN、ELM, 与 SVM 识别率相比略高些. 同时可以看到, 集成后的 OAO-ELM 比相对单一的 ELM 识别率有明显的提高, 是因为它对训练集重构了, 使

得训练集之间有较大的差异性,同时各个训练集本身保持了原有样本集的大部分信息.其中开心,中性和伤心的识别率高于其他表情,恐惧的识别率最低,这是由恐惧与惊讶、生气表情中的部分图像相似造成误判的比较多.

表 2 JAFFE 数据库下不同方法的表情分类识别时间(s)比较

算法	KNN	SVM	ELM	本文方法
第一次	1.12	0.81	0.08	0.49
第二次	1.15	0.87	0.09	0.55
第三次	1.09	0.91	0.11	0.61
平均识别时间	1.12	0.86	0.09	0.55

从表 2 的分类识别的时间可以看出, KNN 的时间最长, OAO-ELM 的时间比 SVM 的识别消耗时间要少一些, ELM 识别时间最短. 而 OAO-ELM 识别却比单独的 ELM 的识别时间要花费很多, 这是因为它是由多个分类器组成, 分类时速度会慢些.

为了看出 ELM 与 OAO-ELM 这两个算法的稳定性, 求得上述实验过程中 10 次重复实验结果的标准差, 结果如图 3 所示. 从图 3 可以看出, OAO-ELM 算法的稳定性要远远优于 ELM.

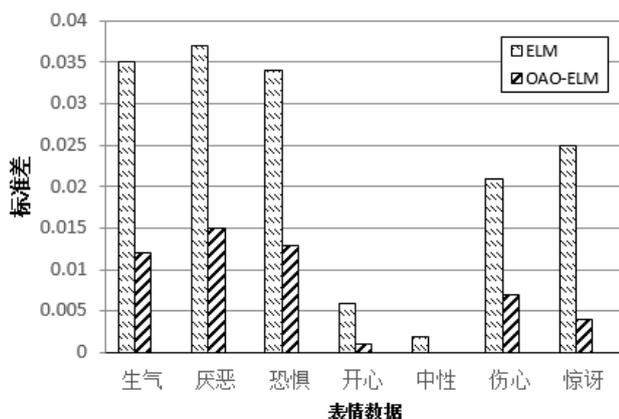


图 3 不同表情数据下的标准差

## 5 结论

本文提出一种基于一对一极限学习机算法—OAO-ELM. 该算法在对原有训练集数据重构的基础上, 使集成的极限学习机产生稳定性更强的分类器. 实验结果证明该算法相比 KNN、SVM 以及单独的 ELM 算法能获得更高的分类率, 同时该算法比单独的 ELM 具有更好的稳定性, 但是该算法的时间复杂度比单独的 ELM 要高出许多, 略高于 SVM. 因此, 该算法的计算复杂性有待于进一步的提高, 是下一步的研究

重点.

## 参考文献

- Ekman, Pual. Facial expression and emotion. American Psychologist, 1993, 48(4): 384–392.
- 丁世飞, 奇丙娟, 谭红艳. 支持向量机理论与算法研究综述. 电子科技大学学报, 2011, 40(1): 2–7.
- 余辉, 赵晖. 支持向量机多类分类算法新研究. 计算机工程与应用, 2008, 44(7): 185–189.
- Pantic M, Rothkrantz L. Facial action recognition for facial expression analysis from static face images. IEEE Trans. on Systems, Man and Cybernetics: Part B, 2004, 34(3): 1449–1461.
- Chellappa R, WILSON C L, SIROHEY S. Human and machine recognition of faces: a survey. Proc. of the IEEE.[S. I.]. IEEE. 1995. 705–740.
- 宋伟, 赵清杰, 宋红, 樊茜. 基于关键块空间分布与 Gabor 滤波的人脸表情识别算法. 中南大学学报(自然科学版), 2013, 44(2): 239–243.
- 赵三平, 苏丽秋. 基于 Gabor 变换的图像纹理识别. 通信技术, 2008, 41(12): 259–261.
- 孔锐, 张冰. Gabor 滤波器参数设置. 控制与决策, 2012, 27(8): 1277–1280.
- 张晓璐. 基于 PCA 的人脸识别技术的研究. 辽宁科技学院学报, 2014, 4: 38–42.
- Huang GB, Zhu QY, Siew CK. Extreme learning machine: a new learning scheme of feed forward neural networks. Proc. of the International Joint Conference on Neural Networks (IJCNN 2004). Budapest, Hungary. July 2004. 985–990.
- Huang GB, Zhu QY, Siew CK. Extreme learning machine: theory and applications. Neurocomputing, 2006, 70: 489–501.
- Zheng WB, Qian YT, Lu HJ. Text categorization based on regularization extreme learning machine. Neural Computing & Applications, 2012.
- Weston J, Watkins C. Multi-Class support Vector machines. Proc. of ESANN99. USA. Factor Press. 1999.
- 蒋斌, 贾克斌, 杨国胜. 人脸表情识别的研究进展. 计算机科学, 2011, 38(4): 25–30.