

# 采用改进高斯核的 MLS-SVM 人脸表情识别算法

卢照敢<sup>1,2</sup> 许春梅<sup>3</sup> 孙楠<sup>1</sup> 苗许娜<sup>1</sup>

(河南财经政法大学计算机与信息工程学院 郑州 450002)<sup>1</sup>

(西安电子科技大学电子工程学院 西安 710071)<sup>2</sup> (河南理工大学计算机科学与技术学院 焦作 454000)<sup>3</sup>

**摘要** 针对用于支持向量机的低维输入数据空间向高维特征空间的映射,通过黎曼测度张量扩大了支持向量机的线性可分边界,进一步提高了支持向量机分类的准确性。考虑到 MLS-SVM 的多分辨逼近效果和改进高斯核函数对支持向量机分类准确度的提升,企图努力给出一种基于两者优点的人脸表情识别算法,以反映人类在自然界中的认知过程,提出了采用改进高斯核的 MLS-SVM 人脸表情识别算法。实验结果表明,其人脸表情识别性能通过修改高斯核函数获得了较大的提升。

**关键词** 支持向量机,人脸表情识别,高斯核,最小二乘算法

中图法分类号 TP391.4 文献标识码 A

## Improved Face Emotion Identification Algorithm by MLS-SVM with Modified Gauss Kernel Function

LU Zhao-gan<sup>1,2</sup> XU Chun-mei<sup>3</sup> SUN Nan<sup>1</sup> MIAO Xu-na<sup>1</sup>

(College of Computer and Information Engineering, Henan University of Economics and Laws, Zhengzhou 450002, China)<sup>1</sup>

(School of Electronic Engineering, Xidian University, Xi'an 710071, China)<sup>2</sup>

(School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454000, China)<sup>3</sup>

**Abstract** Aiming at the map of the low dimension data space for support vector machines (SVM) to high dimension feature space, the Riemann measure tensor was used to extend the SVM linear disjunctive borders, and the SVM classification accuracy was improved. In fact, the good multiple resolution approximation accuracy of MLS-LSM has the similar process to that of the face emotion identification of human eyes. Therefore, one face emotion identification algorithm with the improved MLS-SVM by Gauss kernel function was proposed in this paper, which could improve the face emotion identification accurate. At last, the numerical evaluation results show that the accurate of face emotion identification is improved with the comparison to two classification approaches.

**Keywords** Support vector machine, Face emotion identification, Gauss kernel function, Least square algorithm

## 1 引言

人脸表情可以反映人的情感、精神活动、社交互动以及身体信号。因此,多年以来,自动的人脸表情识别成了一个相对比较活动的研究领域,并在人脸情感、人机交互和图像检索等方面得到了应用<sup>[1]</sup>。人脸表情识别实际上是对人脸图像的表情特征进行分类的结果,需要事先用大量的人脸表情图像及其对应的表情对识别算法进行训练,再根据训练结果获得的识别模式对其他的人脸图像表情进行识别<sup>[2]</sup>。很显然,人脸表情识别的表情训练过程实际上是一个复杂的多维数据逼近过程<sup>[3]</sup>,几乎无法用现有的数学方法进行建模。与此同时,除了人脸识别训练过程中的上述问题外,影响人脸表情识别效果的另一个方面就是人脸表情图像的分类算法<sup>[4]</sup>。也就是说,为了提高当前人脸表情识别的精确度,需要从人脸表情训练算法和分类算法入手来解决所存在的问题。众所周知,支持向量机作为根据样本训练数据的学习方法,已经在科学与

工程领域得到了很广泛的应用<sup>[5]</sup>。作为一种以统计学习理论为支撑的学习算法,支持向量机更广泛地被认为是一种用核函数替代的学习算法,具有许多的优点,如不存在局部极小点,简化了损失函数最优化中的复杂学习过程。由此,支持向量机已经成功地应用于人脸识别和文字分类等信号处理领域。

在汪烈军的博士论文中,通过研究当前支持向量机逼近算法的性能及不足,引入了基于多分辨率的最小二乘逼近算法(Least Square algorithm, with Multiple resolution analysis, MLS),以解决当前人脸表情训练数据逼近度差的问题<sup>[6]</sup>。这里,本文拟通过研究支持向量用于解决分类问题的途径与不足,改进支持向量机实现分类的方法,提高人脸表情图像分类的准确度。

## 2 当前人脸表情识别的研究现状

实际上,人脸表情识别分析的历史可以追溯到达尔文在

本文受国家青年科学基金项目(61309033)资助。

卢照敢(1977—),男,博士后,讲师,主要研究方向为信号与信息处理、数字图像处理, E-mail: luzhaogan@126.com; 许春梅(1979—),女,硕士,讲师,主要研究方向为数字图像处理、视频处理与通信等; 孙楠(1983—),男,硕士,讲师,主要研究方向为数据挖掘、图像处理、优化算法等; 苗许娜(1977—),女,博士,讲师,主要研究方向为 Ad hoc 组播路由算法、认知无线网络。

19 世纪所做的工作<sup>[7]</sup>。在 1872 年,达尔文发现了人脸表情的一般性及其与人们思想活动的内在联系,与其他的身体变化有关。到了 1972 年,Ekman 和 Friesen 将人类的表情划分成了 6 个基本情感,即高兴、悲伤、惊讶、害怕、厌恶和生气等<sup>[8]</sup>,并提出了一个人脸活动编码系统(facial action coding system, FACS)。而在 1978 年, Suwa 对人脸图像序列中的自动人脸表情识别分析做了初步研究,并为后续的人脸表情识别提供了理论支撑<sup>[9]</sup>。人们对人脸表情自动识别的研究,正是得益于 Mase 和 Pentland 在 1990 年的开创性工作<sup>[10]</sup>,才取得了突飞猛进的发展。人脸特征提取是人脸表情识别的关键步骤,常用的人脸特征提取方法包括主成分分析法(Principal Component Analysis, PCA)、人脸模式法<sup>[11]</sup>和 Fisherface<sup>[12]</sup>等方法。一般来说,这些方法可划分为两类,即局部特征提取方法和全局特征提取方法<sup>[13]</sup>,前者通过检测搜索人脸的嘴、眉毛等几何结构对人脸的表情进行识别,而后者将人脸图像看作一个整体矩阵并提取其全局的特征。

在提取了人脸图像的特征之后,需要用分类算法对其按特征进行分类,获得不同的人脸表情识别结果。根据 Fasel 和 Luettn 的研究结果<sup>[7]</sup>,人脸表情分类算法可分为基于空时模式的分类算法和基于空间的分类算法两类,前者包括隐马尔可夫模型(Hidden Markov models, HMM)、递归神经网络法(recurrent neural networks, RNN)及空时表情能量模板等。例如,在文献<sup>[14]</sup>中介绍了多种基于 HMM 的分类算法,并实现了与图像运动信息提取方法互相融合。而基于空间的分类算法包括基于规则的神经网络法和支持向量机方法等,这些方法一般直接用于人脸图像识别,或者结合基于 PCA 和 Gabor 小波的人脸特征提取结果来实现对人脸表情的识别<sup>[15]</sup>。

如上所述的人脸表情识别算法,多注重于人脸表情图像的分类,但实际的分类准确度在更大程度上取决于人脸表情图像的分类训练模式,只有在正确的人脸图像分类训练模式上,好的人脸表情图像分类算法才能够实现准确度更高的人脸表情图像识别。无论是人脸表情分类算法的训练过程还是分类过程,其本质上都是信号逼近过程,本文将 MLS-SVM 中的核函数取为用黎曼测度张量改进的高斯核函数,以提高人脸表情识别过程中的训练效果和分类准确度。然而,众所周知,图像数据是二维数据,无法直接用于人脸表情图像的表情分类,需要将其转换为低维的特征数据才能够用于人脸表情识别,因此人脸图像的表情特征提取是一个至关重要的环节。接下来,本文将首先给出人脸表情特征的提取方法,然后给出基于改进高斯核的 MLS-SVM 人脸表情识别方法。

### 3 人脸表情特征提取

人脸表情可反映当事人的行为信息,是人体情感的体现。为了实现对人脸表情的识别,正确恰当的人脸图像表情特征提取至关重要,会使其后的人脸表情识别准确率大大提高。这里,本文采用了基于 Gabor 小波变换和分形维的人脸表情特征提取方法<sup>[16]</sup>。

(1) 该方法首先需要人脸图像进行灰度处理,将其分割成 2 个子区域,即包含眼睛和眉毛的区域与包含嘴巴的区域,它们的具体表现形式对人脸表情的解读具有很重要的影响。

(2) 然后,将分割后的区域分别进行归一化,获得灰度均

衡化和尺度归一化的灰度分割图像,其实现思路是用它们的灰度直方图调整其均值和方差,将不同人脸图像包含眼睛和眉毛的区域与包含嘴巴的区域转化为相同大小的子图像。

(3) 接下来,为了获得人脸表情的特征提取精度,还需要将归一化的子区域图像网格化,以获得反映人脸表情的关键特征点,如眼角、嘴巴和眉毛等。实际中,为了使人脸表情特征的关键特征点包含在多个网格中,本文将区域子图像划分为  $20 \times 20$  大小的网格。

(4) 最后,本文将所有区域子图像的所有网络进行小波变换,取小波系数的平均幅度和方差作为该小网格区域的特征矢量。与此同时,还将每个小网格图像区域的分形盒维数和差分分形维数作为其相应额外的表情特征矢量。

在上述的人脸表情特征提取过程中,需要使用的 Gabor 小波变换可由下式表示:

$$g_{u,v}(x,y) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2(x^2+y^2)}{2\sigma^2}\right) \left(\exp\left(ik \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \quad (1)$$

$$k = \begin{pmatrix} k_x \\ k_y \end{pmatrix} = \begin{pmatrix} k_v \cos\varphi_u \\ k_v \sin\varphi_u \end{pmatrix} \quad (2)$$

其中,  $i$  是所用的复数因子,  $v$  是 Gabor 滤波器系数的长度,  $u$  表示 Gabor 小波函数的方向,  $k$  表示对每个小块所使用 Gabor 小波方向的总个数,一般取值为 6 个方向,参数  $\sigma/k$  的取值显示了所使用高斯核函数的宽度,取  $\sigma = \sqrt{2\pi}$ 。实际操作时,鉴于人脸表情图像均是高频信息,需要取高频小波函数来与需要变换的区域进行卷积。

为了计算子区域网格的分形维数,假设每个小网格的大小为  $s \times s$ ,其像素  $(x,y)$  处的灰度值用  $z$  表示,则该小网格的  $z$  方向形成一个  $s \times s \times s'$  大小的立方体,其中  $s' = s \times G/M$ ,  $G$  是整个人脸图像的灰度等级,而  $M$  是整个人脸图像的大小。由此,本文进一步假设整个人脸图像的第  $(i,j)$  个网格灰度最大值与最小值为  $K$  和  $L$  ( $0 \sim 255$ ),用  $n_r(i,j)$  表示覆盖该网格所需要的盒子数,即:

$$n_r(i,j) = K - L + 1 \quad (3)$$

由此,覆盖整个人脸图像所需要的总盒子数  $N_r$  可由下式获得:

$$N_r = \sum_{i,j} n_r(i,j) \quad (4)$$

这样,取不同的  $r$  值,可相应获得覆盖整个人脸图像所需要的盒子总数  $N_r$  值,进一步利用最小二乘法对集合  $\{\log_n N_r, \log_n(1/r)\}$  进行线性回归,其斜率就是覆盖整个人脸图像所需要的总盒子数  $D$ 。

接下来,本文进一步给出人脸图像表情特征分形维数的计算公式:

$$gd(k) = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{M-k-1} |f(i,j) - f(i,j+k)|}{4M(M-k-1)} + \frac{\sum_{i=0}^{M-k-1} \sum_{j=0}^{M-1} |f(i,j) - f(i+k,j)|}{4M(M-k-1)} + \frac{\sum_{i=0}^{M-k-1} \sum_{j=0}^{M-k-1} |f(i,j+k) - f(i+k,j)|}{4M(M-k-1)} + \frac{\sum_{j=0}^{M-k-1} \sum_{i=0}^{M-k-1} |f(i,j+k) - f(i+k,j)|}{4M(M-k-1)} \quad (5)$$

其中,  $k$  是尺度因子, 其变化范围从 1 到选择的最大值。同样, 按照最小二乘法对  $gd(k)$  进行线性回归, 得到相应的斜率  $H$ 。

#### 4 采用改进高斯核 MLS-SVM 人脸表情识别算法

实际中支持向量机的性能依赖于核函数, 根据核函数的黎曼几何结构, Amari 提出了一种改进高斯核函数的支持向量机, 以提高经典支持向量机的性能, 其思想是提高等角投影中边界的空间分辨率<sup>[17]</sup>, 以达到提高了两类数据集的可分性。假设核函数可以表示为:

$$K(x, x_i) = p(x)p(x_i)K(x, x_i) \quad (6)$$

则在特殊情况下支持向量机 SVM 所用的核函数是高斯核函数, 也就是

$$K(x, x_i) = \exp(-\|x - x_i\|^2 / \sigma^2) \quad (7)$$

这里, 参数  $\sigma$  是核函数的宽, 经过理论证明其相应的黎曼测度张量<sup>[17]</sup>为:

$$g_{ij}(x) = \frac{1}{\sigma^2} \delta_{ij} \quad (8)$$

通过改进高斯核函数, 其黎曼测度张量进一步变为:

$$\tilde{g}_{ij}(x) = p_i(x)p_j(x) + p^2(x)g_{ij}(x) \quad (9)$$

为了确保  $p(x)$  在支持向量  $SV$  的周围拥有较大的数值, 可根据一个样本构造如下的结果, 即:

$$p(x) = \sum_{i \in SV} a_i \exp(-\|x - x_i\|^2 / 2\tau^2) \quad (10)$$

其中,  $\tau$  是一个自由参数, 而式(10)实现的是对所有支持向量上的求和。这样, 将式(10)和式(6)得到的高斯核函数, 用于支持向量机的低维输入数据空间向高维特征空间的映射, 扩大了线性可分区域的边界, 相对于修改前的高核函数来说, 进一步提高了支持向量机分类的准确性。

实际上, 人类对事情的认识过程是一个由粗到细的渐变过程, 而对人脸表情的判读更是如此。考虑到 MLS-SVM 的多分辨逼近效果和改进高斯核函数对支持向量机分类准确度的提升, 本文企图努力给出一种基于两者优点的人脸表情识别算法, 以反映人类在自然界中的认知过程。下面给出了将两者的优点应用于人脸表情识别的过程。

(1) 首先, 需要对已确定人脸表情的图像进行人脸表情特征提取, 其具体的提取算法如上小节内容所示。

(2) 接下来, 用所有训练人脸图像的表情特征对来训练改进高斯的 MLS-SVM 支持向量机, 获得能够识别人脸表情的分类器, 需要针对每个尺度对的 MLS-SVM 支持向量机。

(3) 最后, 对需要表情识别的人脸图像进行表情特征提取, 将其输入到改进高斯核的 MLS-SVM 支持向量分类器, 进行人脸表情识别。

这里, 本文用分类精确率 (Classification Accuracy Rate, CAR) 衡量人脸表情识别的精度, 并用下面的实验测试结果验证该人脸表情识别的性能。

#### 5 实验结果

为了测试本文提出的基于修改高斯核的 MLS-SVM 人脸表情识别算法, 这里将 JAFFE 人脸图像数据库作为测试数据, 用该方法对其进行分类并与实际数据作对比。由此, 获得相应的分类性能反映了该算法的优点。

在本文的实现中, 选择了 10 个年轻日本女性的人脸图像, 她们具有 7 个基本的表情, 如图 1 所示, 人脸训练图像中

每种表情有 3 张图像, 在该数据库中包含 210 幅人脸表情图像, 每幅图像的大小为  $256 \times 256$ 。本文选择了其中的 70 幅图像作为训练人脸图像集合, 余下的 210 幅人脸表情图像用作测试数据, 每种表情具有 30 幅测试图像。经过人脸表情特征提取之后, 人脸图像样本图像可用  $d$  维空间的特征矢量表示, 通过 3 倍取的交叉有效性实验后, 本实验中的维数  $d$  取为 28。在这里的实验验证中, 高斯核函数按式(7)进行改进, 其核宽度  $\sigma$  等于最优值 1, 即  $\sigma=1.0$ 。根据 Amari 的分析结果, 参数  $\tau$  的最优值在  $\sigma/\sqrt{d}$  周围, 则本文取  $\tau$  值为 0.18, 而支持向量机的正则化因子设置为  $C=10^5$ 。



图 1 JAFFE 数据库中的 7 个人脸表情图像

本文在表 1 中给出了修改高斯核的支持向量机人脸表情识别结果, 其分类精确率 (Classification Accuracy Rate, CAR) 通过改进高斯核函数后获得了较大的提高。与此同时, 本文还进一步实现了模糊支持向量机 (Fuzzy-SVM)<sup>[18]</sup>、高斯核支持向量机 (GSVM)<sup>[16]</sup> 等两种经典的人脸表情分类算法, 将其与本文提出的基于修改高斯核 MLS-SVM 人脸表情识别算法 (GMLS-SVM) 的人脸表情分类结果进行对比, 如表 1 所列。

表 1 不同人脸表情分类算法对 JAFFE 人脸图像的分类结果

序号	表情	Fuzzy-SVM	GSVM	GMLS-SVM
1	Angry	91.3%	93.3%	100.0%
2	Disgust	89.3%	93.3%	100.0%
3	Fear	87.0%	90.0%	86.7%
4	Happy	88.3%	96.7%	90.0%
5	Neutral	97.0%	100.0%	100.0%
6	Sad	88.3%	93.3%	93.3%
7	Surprise	86.7%	96.7%	96.6%
平均分类精度		89.7%	94.8%	95.7%

很显然, 本文提出的基于修改高斯核 MLS-SVM 人脸表情识别算法可以获得较为理想的人脸表情分类效果, 相对于高斯核支持向量机实现的人脸表情分类结果提高了 1%, 而相对于 Fuzzy-SVM, 分类效果提高了将近 6%。

为了显示参数  $\tau$  选择的重要性, 本文用两种不同  $\tau$  参数的修改高斯核支持向量机对 JAFFE 人脸图像进行测试, 其结果对比如表 2 所列, 结果显示当  $\tau$  的取值接近最优值 0.18 时, 平均分类精确度较高, 而无论参数  $\tau$  的值过大或过小, 最终的人脸表情识别效果均比较差。

表 2 用不同参数  $\tau$  值的修改高斯核支持向量机对人脸表情识别的结果对比

序号	表情	$\tau=0.18$	$\tau=2.98$
1	Angry	100.0%	86.6%
2	Disgust	100.0%	90.0%
3	Fear	86.7%	83.3%
4	Happy	90.0%	96.6%
5	Neutral	100.0%	90.0%
6	Sad	93.3%	83.3%
7	Surprise	96.6%	100.0%
Avg. CAR		95.7%	90.0%

结束语 本文通过研究当前支持向量机逼近算法的性能

(下转第 142 页)

了USAN区域的大小和形状信息。实验结果表明,M-USAN算子能够有效抗噪声干扰且具有较高的边缘定位精度。

### 参 考 文 献

[1] 冈萨雷斯,等. 数字图像处理(第二版)[M]. 阮秋琦,等译. 北京: 电子工业出版社,2003

[2] Canny J. A computational approach to edge detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986,8(6):679-698

[3] 张喜平,薛立霞,杨帆,等. 基于模糊推理的边缘检测新算法[J]. 计算机工程与应用,2011,47(7):208-210

[4] L De-Sian, C Chien-Chang. Edge detection improvement by ant colony optimization[J]. Pattern Recognition Letters, 2008, 29: 416-425

[5] 张闯,王婷婷,孙冬娇,等. 基于欧式距离图的图像边缘检测[J]. 中国图象图形学报,2012,18(2):176-183

(上接第134页)

及不足,引入了基于多分辨率的最小二乘逼近算法,以解决当前人脸表情训练数据逼近度差的问题。同时,通过研究支持向量用于解决分类问题的途径与不足,改进支持向量机实现分类的方法,提高人脸表情图像分类的准确度。

### 参 考 文 献

[1] 胡步发,陈炳兴,黄银成. 基于表情子空间多分类器集成的非特定人人脸表情识别[J]. 计算机应用,2011,31(3):736-740

[2] 蒋斌,贾克斌,杨国胜. 人脸表情识别的研究进展[J]. 计算机科学,2011,38(4):25-31

[3] 汪烈军. 基于支持向量机的图像信号处理技术研究[D]. 西安: 西安交通大学,2012

[4] Fasel B, Luetttin J. Automatic facial expression analysis: a survey [J]. Pattern Recognition, 2003, 36(1): 259-275

[5] Ekman P, Friesen W V. Constants across cultures in the face and emotion[J]. Journal of Personality and Social Psychology, 1971, 17(2): 124-129

[6] 洪伟铭. 组合方法在人脸表情识别中的应用研究[J]. 计算机仿真,2011,28(2):337-340

[7] 郑秋梅,吕兴会,时公喜. 基于多特征集成分类器的人脸表情识别[J]. 中国石油大学学报(自然科学版),2011,35(1):174-177

[8] Suwa M, Sugie N, Fujimora K. A preliminary note on pattern recognition of human emotional expression[C]// Proc of the 4th Int Joint Conference on Pattern Recognition (IJCP). Kyoto, Japan, 1978: 408-410

[9] Mase K, Pentland A. Recognition of facial expression from optical flow [J]. IEICE Trans, 1991, E74(10): 3474-3483

[6] 李海华,范娟. 基于梯度的自适应边缘检测算法研究[J]. 河南理工大学学报(自然科学版),2013,32(1):76-79

[7] Smith S M, Brady J M. SUSAN—A new approach to low level image processing[J]. Int. J. Comput. Vis., 1997, 1(23): 45-78

[8] Ewaryst R. SUSAN edge detector reinterpreted, simplified and modified[C] // 2007 Int. Workshop on Multidimensional Systems. 2007: 69-74

[9] Perez M M, Dennis T J. An adaptive implementation of the SUSAN method for image edge and feature detection[C] // International Conference on Image Processing (ICIP'97). 1997: 26-29

[10] 李颖杰,杨华. 基于改进SUSAN算法的红外图像边缘增强方法[J]. 光电技术应用,2009,24(4):64-66

[11] Prokoprij, Reeves A P. A survey of moment-based object representation and recognition [J]. Graphical Models and Image Processing, 1992, 54(5): 438-460

[10] Turk M A, Pentland A P. Face recognition using eigenfaces [C] // Proc. IEEE Conf. on Computer Vision and Pattern Recognition. 1991: 586-591

[11] Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs. fisherfaces: recognition using class specific linear projection[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 1997, 19(7): 711-720

[12] Ying Zi-lu, Li Jing-wen, Zhang You-wei. Facial expression recognition based on classifier Combinations[C] // Proceedings of 8th IEEE Int. Conf. Signal Processing (ICSP). Guilin, China, 2006: 367-372

[13] Otsuka T, Ohya J. Spotting segments displaying facial expression from image sequences using HMM[C] // IEEE Proc of the 2th Int. Conf. on Automatic Face and Gesture Recognition. Nara, Japan, 1998: 442-447

[14] Lyons M J, Akamatsu S, Kamachi M, et al. Coding facial expressions with Gabor wavelets[C] // IEEE Proc of the 2th Int. Conf. on Automatic Face and Gesture Recognition. Nara, Japan, 1998: 200-205

[15] 叶吉祥,胡秀明. 基于Gabor小波与分形维的人脸表情特征提取[J]. 计算机工程,2008,34(4):226-230

[16] 董玉龙,姜威. 基于提升小波和FLD的人脸表情识别[J]. 光学技术,2012,38(5):579-582

[17] Amari S. Improving support vector machine classifiers by modifying kernel functions[J]. Neural Networks, 1999, 12: 783-792

[18] Marc A, Michel B. Image coding using wavelet transform [J]. IEEE Transactions on Image Processing, 1992, 1(2): 205-220