基于 SLPP 与 MKSVM 的痛苦表情识别

张 伟¹, 黄 炜², 夏利民¹, 罗大庸¹

(1. 中南大学信息科学与工程学院, 长沙 410075; 2. 长沙航空职业技术学院计算机系, 长沙 410124)

摘 要: 为提高痛苦表情识别的准确率, 提出一种基于监督保局投影(SLPP)与多核线性混合支持向量机(MKLMSVM)的识别方法。引入先验类标签信息的 SLPP 获取痛苦表情特征, 以解决保局投影方法在未使用先验类标签信息的情况下忽略类内局部结构的问题, 并采用 MKLMSVM 实现痛苦表情的分类。实验结果表明, 该方法的识别准确率可达 88.56%, 明显优于主动外观模型方法, 与一般的支持向量机分类相比, 可以提升决策函数的可解释性及分类性能。

关键词: 痛苦表情识别; 监督保局投影; 先验类标签; 多核支持向量机; 多核线性混合; 主动外观模型

Pain Expression Recognition Based on SLPP and MKSVM

ZHANG Wei¹, HUANG Wei², XIA Li-min¹, LUO Da-yong¹

(1. College of Information Science and Engineering, Central South University, Changsha 410075, China;

2. Department of Computer, Changsha Aeronautical Vocational and Technical College, Changsha 410124, China)

【Abstract】 In order to improve the accuracy rate of pain expression recognition, a method is proposed based on Supervised Locality Preserving Projections(SLPP) and Multiple Kernel Linear Mixture Support Vector Machines(MKLMSVM). The SLPP using prior class label information is adopted for extracting feature of pain expression, which can solve the problem that LPP ignores the within-class local structure without the use of the prior class label information, and then MKLMSVM is employed for recognizing pain expression. Experimental results demonstrate that the accuracy of the proposed approach can reach 88.56%, and is significantly better than the Active Appearance Models(AAM), compared with normal Support Vector Machine(SVM), which can improve the interpretability of decision function and classifier performance.

【Key words】 pain expression recognition; Supervised Locality Preserving Projections(SLPP); prior class label; Multiple Kernel Support Vector Machines(MKSVM); multiple kernel linear mixture; Active Appearance Models(AAM)

DOI: 10.3969/j.issn.1000-3428.2013.12.042

1 概述

痛苦表情是因为主体对内外环境异常的认知而呈现在面部的一种生理神经性反射, 这种反射与认知的客体、认知的主体以及环境等因素密切相关, 所以痛苦表情的呈现方式不仅是多样化的, 在某些情况下甚至存在难以被外部主体辨识和认知的问题。如在临床实践中, 医护人员一般采用问询或视觉模拟评估(Visual Analog Scale, VAS)的方法完成患者痛苦程度的判断, 但是这类基于经验的方法存在较多的局限, 如问询类方法受到患者表达能力的限制而无法完成诸如婴幼儿及昏迷类患者痛苦程度的判定; 而基于 VAS 的判断方法则受到光照、年龄、性别、种族等多种因素的干扰。

为了克服 VAS 等基于经验的人工判别方法的缺点, 基于机器学习的痛苦表情识别技术获得了长足发展, 其中, 基于保局投影(Locality Preserving Projections, LPP)的表情特征提取以及基于支持向量机(Support Vector Machines, SVM)的表情分类方法得到了广泛应用。文献[1]提出一种基于 Contourlet 变换与 LPP 相结合的人脸表情识别方法。该方法将人脸表情图像分割为左眼、右眼和嘴 3 个部分, 利用 Contourlet 变换对局部表情图像和原始图像进行处理, 得到图像的低频分量和高频分量, 结合局部表情图像的低频分量与原始图像的高频分量, 采用 LPP 算法提取表情特征, 并利用 SVM 进行分类。文献[2]开发了一种自动识别痛苦表情的方法, 该方法采用主动外观模型(Active Appearance Models, AAM)从面部图像解耦轮廓和外观信息, 进而获得

基金项目: 国家博士点基金资助项目(20090162110057); 湖南省科技计划基金资助项目(2011GK3213)

作者简介: 张 伟(1972—), 男, 博士研究生, 主研方向: 模式识别, 计算视觉; 黄 炜, 讲师、硕士; 夏利民、罗大庸, 教授、博士生导师

收稿日期: 2012-10-29 **修回日期:** 2012-12-24 **E-mail:** 360706711@qq.com

面部表情的低维特征, 然后通过 SVM 实现痛苦表情的分类。但是, LPP 在本质上是无监督的线性学习方法, 并受到人脸图像的光照、姿态等因素的干扰, 而 SVM 在本质上是二值分类方法, 无法实现对痛苦表情多样性的精确表述。

针对基础 LPP、SVM 方法在痛苦表情识别中的缺点, 本文采用 SLPP 与多核支持向量机(Multiple Kernel Support Vector Machines, MKSVM)相结合的方法, 首先通过 SLPP 方法获取痛苦表情的低维特征^[3], 然后通过 MKSVM 实现痛苦表情的分类识别^[4-5]。

2 基于SLPP的痛苦表情特征提取

2.1 保局投影方法

保局投影是近年来发展起来的基于流形学习的降维方法^[6], 其基本思想是通过投影矩阵 A , 将高维空间欧式空间 R^n 中的数据集 $X = [x_1, x_2, \dots, x_n]$ 投影到低维特征空间 R^d ($d \ll n$) 中, $Y = AX = [y_1, y_2, \dots, y_N]$, 为了在 R^d 中保持流形的局部结构, 需求解如下的极小化目标函数:

$$\min \sum_{i,j=1}^n \|y_i - y_j\|^2 W(i,j) \quad (1)$$

其中, $y_i = A^T x_i$; $W_{ij} = \begin{cases} 1 & x_i \in N_K(x_j) \\ 0 & \text{otherwise} \end{cases}$; $N_K(x_j)$ 为 x_j

的 K 近邻集合。式(1)的极小化求解可转化为如下广义特征值求解问题:

$$XLX^T a = \lambda XDX^T a \quad (2)$$

其中, $D_{ij} = \sum_j W(i,j)$ 为对角阵, 且 $L = D - W$ 为 Laplace 阵。给定约束条件 $A^T XDX^T A = 1$, 求解式(2), 得到排序后的前 d 个广义特征值 $\lambda_0, \lambda_1, \dots, \lambda_{d-1}$, 其对应的 d 个特征向量 $a_0, a_1, \dots, a_{d-1}$ 构成了投影矩阵 A , 即 $x_i \rightarrow A^T x_i$, $A = [a_0, a_1, \dots, a_{d-1}]$, 其中, y_i 是 d 维向量, A 是 $N \times d$ 矩阵。

采用 LPP 方法对痛苦表情数据进行降维的缺点是没有充分利用训练样本的类内几何结构及训练点的先验类别信息, 对光照、姿态等噪声干扰较为敏感, 可能产生类内差异过大等问题, 并最终超过类间差异从而导致识别错误。

2.2 痛苦表情的特征提取

LPP 的本质是通过拉普拉斯特征映射的线性逼近而实现的一种无监督线性降维方法^[6], 针对采用 LPP 实现高维非线性表情数据降维过程中所存在的问题, 在其极小化目标函数构造过程中引入了先验类标签信息, 使得 LPP 方法有监督化。对于表情图像集 $X = [(S_1, C_1), (S_2, C_2), \dots, (S_n, C_n)]$, 其中, $S_i = [x_1, x_2, \dots, x_m]$ 是具有 N 维特征的原始脸部图像, 而 C_i 是类别标签。参考文献[7]所给出的方法, 通过非线性函数 ϕ 将 X 映射到另一高维特征空间 $F: \phi(X) = [\phi(x_1), \phi(x_2), \dots, \phi(x_n)]$ 中, 并在特征空间 F 中通过求解样本权重距离和极小化目标函数的方式, 寻找

可以保持数据 $\phi(x)$ 类内几何结构的投影变换 P , 其样本权重距离和极小化目标函数定义如下:

$$\min_P \sum_{i,j=1}^m \|z_i - z_j\|^2 W(i,j) \quad (3)$$

$$W_{ij} = \begin{cases} \phi(x_i) \cdot \phi(x_j) & x_i, x_j \text{ belong to the same label} \\ 0 & \text{otherwise} \end{cases}$$

$$z_i = P^T \phi(x_i)$$

对于每个权重矩阵 W 都可以看作是一对带有类标签信息样本的相似性度量。定义目标函数:

$$\sum_{i,j=1}^n \|z_i - z_j\|^2 W(i,j) = 2P^T \phi(X)(D - W)\phi(X^T)P \quad (4)$$

其中, $D_{ii} = \sum_j W(i,j)$ 为对角阵, 因为线性变换 P 位于向量组 $\phi(x_1), \phi(x_2), \dots, \phi(x_n)$ 所张成的空间, 即 $Span\{\phi(x_1), \phi(x_2), \dots, \phi(x_n)\}$ 上, 且存在系数向量 $a = [a_1, a_2, \dots, a_n]$ 使得:

$$P = \sum_{i=1}^n a_i \phi(x_i) = \phi(X)a \quad (5)$$

将式(5)代入式(4)中得到:

$$\sum_{i,j=1}^n \|z_i - z_j\|^2 W(i,j) = 2a^T K(D - W)Ka \quad (6)$$

其中, 矩阵 $K(i,j) = \phi(x_i) \cdot \phi(x_j)$ 。因此, 给定约束条件 $a^T KDKa = 1$, 则极小化问题的求解可转化为广义特征值的求解问题, 即:

$$K(D - W)Ka = \lambda KDXa \quad (7)$$

投影变换 P 可以通过求解式(7)和式(5)获得, 因此, 痛苦表情的特征 $Z = (z_1, z_2, \dots, z_m)$ 由式(8)决定:

$$Z = P^T \phi(x) \quad (8)$$

3 痛苦表情识别

3.1 支持向量机

SVM^[8]是建立在统计学习理论和结构风险最小原理基础上的分类方法, 并根据有限的样本信息在模型复杂性和学习能力之间给出了最佳折衷, 在小样本、非线性及高维模式识别过程中表现出许多特有的优势。对于给定的标签样本数据集 $D = \{x_i, y_i\}_{i=1}^m$, 其中, $x_i \in X$ 为输入向量, 且 $y_i \in \{-1, 1\}$ 。从输入空间 X 中导入特征映射 ϕ 到再生核 Hilbert 空间 H , 则 H 中线性分类器的形式如下:

$$f(x) = w^T \phi(x) + b \quad (9)$$

其参数 (w, b) 由下式的解决定:

$$\min_{w,b,\xi} \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^n \xi_i, \quad \text{s.t. } \forall i, y_i \{w^T \phi(x_i) + b\} \geq 1 - \xi_i; \xi_i \geq 0 \quad (10)$$

其中, $\|\cdot\|_2$ 表示 l_2 范数; $C > 0$ 为正则化常量。将式(10)转化

为双优化问题:

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j k(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq C, \forall i; \sum_{i=1}^n y_i \alpha_i = 0 \end{aligned} \quad (11)$$

其中,核函数 $k(\mathbf{x}, \mathbf{x}_i) = \langle \phi(\mathbf{x}), \phi(\mathbf{x}_i) \rangle$ 。至此,通过式(11)可计算得出优化参数,而最终的决策函数由下式决定:

$$f(\mathbf{x}) = \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}) + b \quad (12)$$

单核函数 SVM 适用于二值分类应用,对于痛苦表情的多样性,存在分类表述模糊、分类效率低下等问题。

3.2 基于 MKLMSVM 的痛苦表情分类

针对痛苦表情的多样性,采用单核函数 $k(\mathbf{x}_i, \mathbf{x})$ 已无法实现清晰的分类,因此,采用多核学习的方法,通过联合优化分类器系数以及多核权重系数,可有效提高目标识别的性能^[9-10]。

令 $\mathbf{K}_1, \mathbf{K}_2, \dots, \mathbf{K}_m$ 为从不同的特征值获得的形如 $\mathbf{K}_t = [k_t(\mathbf{x}_i, \mathbf{x}_j)]_{i,j=1,2,\dots,n}, 1 \leq t \leq m$ 的 m 个核矩阵,多核学习构架通过学习多核线性混合的方式扩展了 SVM 的一般形式:

$$\mathbf{K}_{\beta} = \sum_{i=1}^m \beta_i \mathbf{K}_i \quad (13)$$

因此,式(9)变成如下形式:

$$f(\mathbf{x}) = \sum_{i=1}^n \beta_i \mathbf{w}_i^T \phi_i(\mathbf{x}) + b \quad (14)$$

通过同时优化参数 $\mathbf{w}, b, \xi, \beta$, 对应优化问题可最大化算法的泛化性能:

$$\begin{aligned} \min_{\beta, \mathbf{w}, b, \xi} \quad & \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & \forall i, y_i \{ \langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b \} \geq 1 - \xi_i \\ & \xi_i \geq 0, \beta_i \geq 0, \|\beta\| = 1 \end{aligned} \quad (15)$$

其中, $\phi(\mathbf{x}_i) = (\sqrt{\beta_1} \phi_1(\mathbf{x}_i), \sqrt{\beta_2} \phi_2(\mathbf{x}_i), \dots, \sqrt{\beta_m} \phi_m(\mathbf{x}_i))^T$, $\mathbf{w} = (\sqrt{\beta_1} \mathbf{w}_1, \sqrt{\beta_2} \mathbf{w}_2, \dots, \sqrt{\beta_m} \mathbf{w}_m)^T$ 。上述优化问题可转化为半无限规划问题,并可通过标准方法求解^[11]。当 $m=1$ 时,上述 MKLMSVM 方法转化为一般 SVM 方法。

在痛苦表情识别过程中,通过 SLPP 方法获取的痛苦表情特征 $\mathbf{z} = (z_1, z_2, \dots, z_m)$ 作为 MKLMSVM 方法的输入,而其输出 $y_i \in \{-1, 1\}$, 当 $y_i = 1$ 时意味着痛苦表情,否则反之。

4 实验结果与分析

为了验证 SLPP 与 MKLMSVM 方法相结合识别痛苦表情的性能,建立了包含痛苦、正常 2 种表情的表情数据库,

在该数据库中,摄制并存储了 15 位男性和 15 位女性的痛苦和正常表情图像,每位对象的正常及痛苦表情图像各 5 幅,其中的痛苦表情可能包括困倦、疼痛等表情。表情图像的摄制在光照可控的实验室环境下完成,同时为了更方便的完成识别实验,对表情数据库中的表情图像进行了相关预处理:(1)通过 Photoshop 裁减了脸部位图;(2)进行了调整及伸缩等预处理,以确保同类面部图像中的两眼之间具有相同的距离,并在同一坐标系中拥有相近的坐标。以上预处理的目的是为了更方便脸部跟踪、AAM 等算法的实现。预处理后某女性对象的正常和痛苦表情如图 1 所示。

图 1 表情数据库中某对象的正常及痛苦表情

实验过程中所需要的 AAM、Contourlet 变换、LPP、SVM、SLPP、MKLMSVM 等相关算法全部在 64 位的 Windows 7 环境下采用 Matlab 7.11 设计实现,并通过配置为 Intel(R) Core(TM)i7、主频 1.6 GHz、内存 4 GB、硬盘容量 500 GB 的 ASUS N61J 笔记本电脑完成相关计算。

为了验证本文方法的识别性能,与文献[1-2]中所给出的方法进行了对比实验。对于每种方法的每次实验过程,都首先从 2 种表情中各自随机抽取 50 幅图像对模型进行训练,而每种表情各自剩余的 100 幅图像用于痛苦表情的识别实验。对于每种方法都执行上述实验过程 5 次,并采用平均识别率作为最终的检测准确率,实验结果如表 1 所示。

表 1 不同方法实验结果对比

方法	错误率	准确率
LPP+SVM ^[1]	0.237 5	0.762 5
AAM+SVM ^[2]	0.177 3	0.822 7
SLPP+MKMSVM	0.114 4	0.885 6

从表1可知, 本文所给出方法的识别率明显高于文献[1]及文献[2]中所给出的方法, 其根本原因在于对于痛苦表情识别的2个处理阶段: 首先通过有监督的LPP方法有效地提取了痛苦表情特征, 其低维特征充分保持了高维流形的几何结构; 而后通过MKLSVM提高了分类效率并使得分类更加明晰, 其分类效果显著高于单核SVM方法。

虽然采用预处理措施简化了面部跟踪过程, 但是本文所提及算法需要较大的计算量, 单幅面部位图的SLPP与MKLSVM计算过程耗时约3s左右。

5 结束语

针对采用LPP和SVM方法进行痛苦表情识别过程中类内局部结构被忽略及表情多样性支持较差等问题, 本文通过有监督保局投影方法获取痛苦表情的低维流形特征, 并以此为基础通过多核支持向量机实现了痛苦表情的最终识别, 主要工作如下: (1)采用SLPP方法, 避免了LPP降维过程中对流形类内结构的忽略, 在低维投影空间有效保持了高维流形的几何结构, 提高了高维痛苦表情特征提取的有效性; (2)采用MKLSVM对痛苦表情的低维特征进行分类, 可有效应对痛苦表情的多样性, 提高了分类的效率; (3)采用SLPP与MKLSVM算法实现痛苦表情分类识别, 存在实时性较差等问题, 在实际应用中除对计算过程进一步优化外, 可在硬件系统设计上采用数字信号处理器^[12]完成相关计算以改善系统的实时性能。

总之, 本文结合SLPP与MKLSVM的优势, 可有效提高痛苦表情识别的准确性, 对于相关的机器学习过程具有一定的借鉴意义。

参考文献

- [1] 高现文, 付炜, 祝鹏. 基于Contourlet变换与LPP的表情识别[J]. 计算机工程, 2012, 38(6): 184-186.
- [2] Ashraf A B, Lucey S, Cohn J F, et al. The Painful Face-pain Expression Recognition Using Active Appearance Models[J]. Image and Vision Computing, 2009, 27(12): 1788-1796.
- [3] Xiao Yongliang, Xia Limin. Shot Boundary Detection Based on Supervised Locality Preserving Projections and KNN-SVM Classifier[C]//Proceedings of the 2nd International Asia Conference on Informatics in Control, Automation and Robotics. Wuhan, China: [s. n.], 2010.
- [4] Xiao Yongliang, Xia Limin, Zhang Wei. Face Recognition with Supervised Spectral Regression and Multiple Kernel SVM[C]//Proceedings of the 2nd IEEE International Conference on Advanced Computer Control. Shenyang, China: [s. n.], 2010.
- [5] 付燕, 鲜艳明. 基于多特征和改进SVM集成的图像分类[J]. 计算机工程, 2011, 37(21): 196-198.
- [6] He Xiaofei, Niyogi P. Locality Preserving Projections[C]//Proceedings of Conference on Advances in Neural Information Processing Systems. Vancouver, Canada: IEEE Press, 2003.
- [7] Yu Xuelian, Wang Xuegang, Liu Benyong. Supervised Kernel Neighborhood Preserving Projections for Radar Target Recognition[J]. Signal Processing, 2008, 88(9): 2335-2339.
- [8] Schölkopf B, Smola A J. Learning with Kernels: Support Vector Machines Regularization, Optimization and Beyond[M]. [S. l.]: MIT Press, 2002.
- [9] Vedaldi A, Gulshan V, Varma M, et al. Multiple Kernels for Object Detection[C]//Proceedings of International Conference on Computer Vision. Kyoto, Japan: IEEE Press, 2009.
- [10] Fu Siyao, Yang Guosheng, Hou Zengguang, et al. Multiple Kernel Learning from Sets of Partially Matching Image Features[C]//Proceedings of the 19th International Conference on Pattern Recognition. Tampa, USA: IEEE Press, 2008.
- [11] Felzenszwalb P F, McAllister D, Ramanan D. A Discriminatively Trained, Multiscale, Deformable Part Model[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Alaska, USA: IEEE Press, 2008.
- [12] 赵振华, 郑红. 嵌入式可重构的多DSP图像并行处理系统[J]. 单片机与嵌入式系统应用, 2009, 9(2): 12-15.

编辑 任吉慧

(上接第195页)

- [8] 陈伏兵, 陈秀宏, 张生亮, 等. 基于模块2DPCA的人脸识别方法[J]. 中国图象图形学报, 2006, 11(4): 580-485.
- [9] 陈才扣, 黄璞. 基于分块独立分量分析的人脸识别方法[J]. 中国图象图形学报, 2009, 14(9): 1837-1842.
- [10] 陈伏兵, 谢永华, 严云洋, 等. 分块PCA鉴别特征抽取能力的分析研究[J]. 计算机科学, 2006, 33(3): 155-159.
- [11] 黄晓华, 王春茂, 郑文明. 基于彩色人脸图像的信息融合与识别方法[J]. 中国图象图形学报, 2010, 15(3): 422-428.
- [12] 罗婵娟, 朱嘉钢, 陆晓. 小空间占用的快速彩色图像特征抽取方法[J]. 计算机应用, 2012, 32(10): 2827-2831.
- [13] 胡旭迈. 基于彩色图像的人脸识别研究[D]. 西安: 西北工业大学, 2005.

编辑 索书志