

基于特征融合和模糊核判别分析的 面部表情识别方法

周晓彦^{1),2)} 郑文明³⁾ 邹采荣^{1),4)} 赵力¹⁾

¹⁾ (东南大学信息科学与工程学院, 南京 210096) ²⁾ (南京信息工程大学电子与信息工程学院, 南京 210044)

³⁾ (东南大学学习科学研究中心, 南京 210096) ⁴⁾ (佛山科技学院, 佛山 528000)

摘要 提出了基于特征融合和模糊核判别分析(FKDA)的面部表情识别方法。首先,从每幅人脸图像中手工定位34个基准点,作为面部表情图像的几何特征,同时采用Gabor小波变换方法对每幅表情图像进行变换,并提取基准点处的Gabor小波系数值作为表情图像的Gabor特征;其次,利用典型相关分析技术对几何特征和Gabor特征进行特征融合,作为表情识别的输入特征;然后,利用模糊核判别分析方法进一步提取表情的鉴别特征;最后,采用最近邻分类器完成表情的分类识别。通过在JAFPE国际表情数据库和Ekman“面部表情图片”数据库上的实验,证实了所提方法的有效性。

关键词 模糊核判别分析 核判别分析 面部表情识别

中图法分类号 TP391.41 文献标识码 A 文章编号: 1006-8961(2009)08-1615-06

Facial Expression Recognition Based on Feature Fusion and Fuzzy Kernel Discriminant Analysis

ZHOU Xiao-yan^{1),2)}, ZHENG Wen-ming³⁾, ZOU Cai-rong^{1),4)}, ZHAO Li¹⁾

¹⁾ (School of Information Science and Engineering, Southeast University, Nanjing 210096)

²⁾ (School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044)

³⁾ (Research Center for Learning Science, Southeast University, Nanjing 210096) ⁴⁾ (Foshan University, Foshan 528000)

Abstract In this paper, a facial expression recognition method based on feature fusion and fuzzy kernel discriminant analysis (FKDA) is proposed. This method firstly locates 34 landmark points from each facial image as the Geometric features of the facial image. Then, these landmark points are converted into a labeled graph (LG) vector using the Gabor wavelet transformation method, and the LG vector are used as the Gabor feature vector of the facial image. Both Geometric feature and Gabor feature are further fused using the canonical correlation analysis (CCA) as the final input facial features for recognition. The FKDA method is finally used to further extract the discriminative expression features for classification and the nearest neighbor classifier is used to this goal. Experiments on both Japanese Female Facial Expression (JAFPE) database and the Ekman's 'Pictures of Facial Affect' database demonstrate the better performance of the proposed method.

Keywords fuzzy kernel discriminant analysis, kernel discriminant analysis, facial expression recognition

基金项目:国家自然科学基金项目(60503023,60872160);教育部新世纪优秀人才支持计划(NCET-050467);江苏省高校自然科学基金研究自筹经费项目(08KJD520009)

收稿日期:2008-09-30;**改回日期:**2008-12-12

第一作者简介:周晓彦(1980~),女,讲师。东南大学信息科学与工程学院信号与信息处理专业博士研究生。主要研究方向为模式识别、机器学习等。E-mail: xiaoyan_zhou@seu.edu.cn

1 引言

线性判别分析(LDA)方法是模式识别中非常重要的特征提取方法^[1],其基本思想是寻找某一投影矢量,使得样本经投影后具有最大的类间散射距离和最小的类内散射距离。LDA方法目前已在人脸识别、图像检索、字符识别等领域中得到成功的运用。但是,LDA方法在本质上仍是一种线性特征提取方法,对于非线性模式问题,其可能得不到较好的结果。为了克服LDA方法的不足,Mika等人借助核函数将LDA方法从线性空间推广到非线性空间中,提出了基于核函数的判别分析方法,简称核判别分析(KDA)方法^[2]。此后,Baudat等人^[3]进一步把LDA方法扩展到多类特征提取问题。

KDA方法克服了LDA方法对非线性特征提取问题的缺陷,已在人脸识别等领域中证明具有比LDA方法更好的识别性能。但是,同LDA方法一样,KDA方法在计算过程中,总是假设各个训练样本只唯一隶属于某一类的样本集。在实际问题中,这种假设在某些情况下并不合理。例如,对于面部表情图像模式,如果把每幅表情图像简单归属于6种基本面部表情(高兴、悲伤、惊奇、生气、厌恶和害怕)中的一种,并不能全面刻画表情图像的分布特性。事实上,把每个人的面部表情看成是6种基本表情的混合体将更真实地反映表情的分布特点。为了克服LDA方法的这种不足,Kwak等人借鉴了模糊集理论^[4],提出了模糊判别分析(FLDA)方法^[5],并成功地应用于人脸的识别中。Zheng^[6]和Wu^[7]独立地把FLDA方法推广到核函数空间中,提出了模糊核判别分析(FKDA)方法。FKDA方法通过引入样本的类隶属度信息来弥补了传统的核判别分析方法中每个样本唯一隶属于某一类别的不足。

面部表情识别(FER)是计算机视觉和模式识别领域中一个非常热门的研究课题。最近十年,针对这一问题已提出了多种方法,有关FER的一些研究概况,可参见文献[8]~[9]。一般来说,面部表情识别分为面部特征提取和面部表情分类两个部分。本文在面部表情特征提取方面遵循前期的研究工作^[10],首先从每幅面部图像上定位出34个基准点,作为表情图像的几何特征,同时采用Gabor小波变换方法提取基准点处的Gabor系数作为表情图像的

Gabor特征^[11]。其次,运用典型相关分析(CCA)技术对所提取的几何特征和Gabor特征进行特征融合,把融合后的特征作为表情识别的输入特征。然后再利用FKDA方法进一步提取用于分类的表情鉴别特征。最后利用最近邻分类器完成表情的分类识别。

2 模糊核判别分析

2.1 核判别分析

令 $\mathbf{X} = \{\mathbf{x}_i^j | j = 1, \dots, N_i; i = 1, \dots, c\}$ 是一个有 N 个元素的 n 维训练样本集, c 是所有类的类别数, N_i 表示第 i 类的样本数目,且 $N = N_1 + N_2 + \dots + N_c$ 。令 Φ 是一个非线性映射,将 $\mathbf{X}$ 从输入空间映射到某一高维的希尔伯特空间 F 中,即 $\Phi: \mathbf{X} \rightarrow F, \mathbf{x} \rightarrow \Phi(\mathbf{x})$ 。在特征空间 F 中,任意两点 $\Phi(\mathbf{x})$ 和 $\Phi(\mathbf{y})$ 的内积可由核函数 $k(\mathbf{x}, \mathbf{y})$ 的函数值算得,即

$$k(\mathbf{x}, \mathbf{y}) = \langle \Phi(\mathbf{x}), \Phi(\mathbf{y}) \rangle = (\Phi(\mathbf{y}))^T \Phi(\mathbf{x}) \quad (1)$$

令 $\mathbf{S}_B^\Phi$, $\mathbf{S}_W^\Phi$ 和 $\mathbf{S}_T^\Phi$ 分别表示 F 中样本的类间散射矩阵、类内散射矩阵和总体散射矩阵。则可得

$$\mathbf{S}_B^\Phi = \sum_{i=1}^c N_i (\mathbf{u}_i^\Phi - \mathbf{u}^\Phi) (\mathbf{u}_i^\Phi - \mathbf{u}^\Phi)^T \quad (2)$$

$$\mathbf{S}_W^\Phi = \sum_{i=1}^c \sum_{j=1}^{N_i} (\Phi(\mathbf{x}_i^j) - \mathbf{u}_i^\Phi) (\Phi(\mathbf{x}_i^j) - \mathbf{u}_i^\Phi)^T \quad (3)$$

$$\mathbf{S}_T^\Phi = \mathbf{S}_B^\Phi + \mathbf{S}_W^\Phi \quad (4)$$

式中, $\mathbf{x}^T$ 表示矢量 $\mathbf{x}$ 的转置, $\mathbf{x}_i^j$ 表示第 i 类的第 j 个样本, $\mathbf{u}_i^\Phi$ 表示第 i 类样本的均值,而 $\mathbf{u}^\Phi$ 表示所有样本的均值。

$$\mathbf{u}_i^\Phi = \frac{1}{N_i} \sum_{j=1}^{N_i} \Phi(\mathbf{x}_i^j), \mathbf{u}^\Phi = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} \Phi(\mathbf{x}_i^j) \quad (5)$$

于是,KDA的目标在于找到特征矢量以最大化Fisher判别准则:

$$J(\boldsymbol{\omega}) = \frac{\boldsymbol{\omega}^T \mathbf{S}_B^\Phi \boldsymbol{\omega}}{\boldsymbol{\omega}^T \mathbf{S}_T^\Phi \boldsymbol{\omega}} \quad (6)$$

从文献[3]可知,求解上述最佳判别矢量问题等价于求解如下的广义特征方程的特征值 λ 和相应的特征矢量 $\boldsymbol{\omega}$:

$$\mathbf{S}_B^\Phi \boldsymbol{\omega} = \lambda \mathbf{S}_T^\Phi \boldsymbol{\omega} \quad (7)$$

设 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ 是式(7)中得到的前 k ($k \leq c-1$) 个最大特征值,且 $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_k$ 是所对应的特征矢量,则 $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_k$ 是KDA方法的 k 个最佳投影矢量。

2.2 模糊核判别分析方法

KDA 方法在使用中要求每个训练样本唯一隶属于某一类别,该条件在某些情况下并不适合。FKDA 方法通过引入样本的类隶属度信息来弥补了传统的核判别分析方法中这一不足之处。

为简化 FKDA 方法的推导过程,首先用 $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ 表示所有的 N 个训练样本。假设 $u_{ij} (i = 1, \dots, c; j = 1, \dots, N)$ 是第 j 个样本隶属于第 i 类的隶属度。则样本的类均值可表示为

$$\bar{\mathbf{u}}_i^\Phi = \sum_{j=1}^N u_{ij} \Phi(\mathbf{x}_j) / \sum_{j=1}^N u_{ij} \quad (8)$$

此时,样本的类间散射矩阵,类内散射矩阵以及总体散射矩阵可分别表示为

$$\tilde{\mathbf{S}}_B^\Phi = \sum_{i=1}^c N_i (\bar{\mathbf{u}}_i^\Phi - \mathbf{u}^\Phi) (\bar{\mathbf{u}}_i^\Phi - \mathbf{u}^\Phi)^T \quad (9)$$

$$\tilde{\mathbf{S}}_W^\Phi = \sum_{i=1}^c \sum_{j=1}^{N_i} (\Phi(\mathbf{x}_i^j) - \bar{\mathbf{u}}_i^\Phi) (\Phi(\mathbf{x}_i^j) - \bar{\mathbf{u}}_i^\Phi)^T \quad (10)$$

$$\tilde{\mathbf{S}}_T^\Phi = \tilde{\mathbf{S}}_B^\Phi + \tilde{\mathbf{S}}_W^\Phi \quad (11)$$

此外,Fisher 判别准则可被改写为

$$\tilde{J}(\omega) = \frac{\omega^T \tilde{\mathbf{S}}_B^\Phi \omega}{\omega^T \tilde{\mathbf{S}}_T^\Phi \omega} \quad (12)$$

令 $\Phi(\mathbf{X}) = [\Phi(\mathbf{x}_1^1), \dots, \Phi(\mathbf{x}_1^{N_1}); \dots; \Phi(\mathbf{x}_c^1), \dots, \Phi(\mathbf{x}_c^{N_c})]$, 则式(9)和式(10)可表示为

$$\tilde{\mathbf{S}}_B^\Phi = \Phi(\mathbf{X}) (\mathbf{W}_u - \mathbf{M}) (\mathbf{W}_u - \mathbf{M})^T \Phi(\mathbf{X})^T \quad (13)$$

$$\tilde{\mathbf{S}}_W^\Phi = \Phi(\mathbf{X}) (\mathbf{I} - \mathbf{W}_u) (\mathbf{I} - \mathbf{W}_u) (\Phi(\mathbf{X}))^T \quad (14)$$

式中, $\mathbf{I}$ 是一个 $N \times N$ 阶的单位阵, $\mathbf{M} = (m_{ij})_{i=1, \dots, c; j=1, \dots, N}$ 是一个所有元素都等于 $1/N$ 的 $N \times N$ 阶的矩阵, $\mathbf{W}_u = (\mathbf{W}_l)_{l=1, \dots, c}$ 是一个 $N \times N$ 的分块矩阵,且 $\mathbf{W}_l$ 是一个每列的第 j 个元素都等于 $u_{ij} / \sum_{j=1}^N u_{ij}$ 的 $N \times N_l$ 阶的矩阵。

同文献[3]类似,若 ω 为 FKDA 的最佳判别矢量,则 ω 可由 $\Phi(\mathbf{x}_p^q) - \mathbf{u}^\Phi$ 线性组合而得,即存在系数矢量 α ,使得

$$\begin{aligned} \omega &= \sum_{p=1}^c \sum_{q=1}^{N_p} \alpha_{pq} (\Phi(\mathbf{x}_p^q) - \mathbf{u}^\Phi) \\ &= \Phi(\mathbf{X}) (\mathbf{I} - \mathbf{M}) \alpha \end{aligned} \quad (15)$$

由式(11)、式(13)~(15)可得:式(12)表示的 Fisher 判别准则可改写为

$$\tilde{J}(\alpha) = \frac{\alpha^T \mathbf{B} \alpha}{\alpha^T \mathbf{T} \alpha} \quad (16)$$

式中,

$$\begin{aligned} \mathbf{B} &= (\mathbf{I} - \mathbf{M}) \mathbf{K} (\mathbf{W}_u - \mathbf{M}) (\mathbf{W}_u - \mathbf{M}) \mathbf{K} (\mathbf{I} - \mathbf{M}) \\ \mathbf{T} &= (\mathbf{I} - \mathbf{M}) \mathbf{K} [(\mathbf{W}_u - \mathbf{M}) (\mathbf{W}_u - \mathbf{M}) + \\ &\quad (\mathbf{I} - \mathbf{W}_u) (\mathbf{I} - \mathbf{W}_u)] \mathbf{K} (\mathbf{I} - \mathbf{M}) \end{aligned}$$

其中, $\mathbf{K}$ 为核判别分析方法的 Gram 矩阵。

于是,求解 FKDA 的最佳判别矢量问题等价于求解使式(16)中的 Fisher 判别准则达到最大的判别矢量问题。这一优化求解问题可采用文献[3]中求解 KDA 方法最佳判别矢量的算法完成。假设 $\omega_i = \Phi(\mathbf{X}) (\mathbf{I} - \mathbf{M}) \alpha_i (i = 1, \dots, c-1)$ 是 FKDA 的前 $c-1$ 个最佳判别矢量,令 $\mathbf{W}_{\text{FKDA}} = [\omega_1, \omega_2, \dots, \omega_{c-1}] = \Phi(\mathbf{X}) [\alpha_1, \alpha_2, \dots, \alpha_{c-1}]$ 是 FKDA 的变换矩阵,则某一测试样本 $\Phi(\mathbf{x}_{\text{test}})$ 在 $\mathbf{W}_{\text{FKDA}}$ 下的投影可通过式(17)算得。

$$\begin{aligned} \mathbf{y}_{\text{test}} &= \mathbf{W}_{\text{FKDA}}^T \Phi(\mathbf{x}_{\text{test}}) \\ &= [\alpha_1, \alpha_2, \dots, \alpha_{c-1}]^T \mathbf{k} \end{aligned} \quad (17)$$

式中, $\mathbf{k} = [k(\mathbf{x}_1, \mathbf{x}_{\text{test}}), \dots, k(\mathbf{x}_{c-1}, \mathbf{x}_{\text{test}})]$ 。

假设 $\mathbf{y}_i^j (j = 1, \dots, N_i; i = 1, \dots, c)$ 分别是 $\Phi(\mathbf{x}_i^j) (j = 1, \dots, N_i; i = 1, \dots, c)$ 到 $\mathbf{W}_{\text{FKDA}}$ 的投影矢量,则对 $\Phi(\mathbf{x}_{\text{test}})$ 的分类问题可通过式(18)完成:

$$i_{\text{test}} = \arg \min_i \|\mathbf{y}_{\text{test}} - \mathbf{y}_i^j\| \quad (18)$$

3 基于特征融合和 FKDA 的表情识别

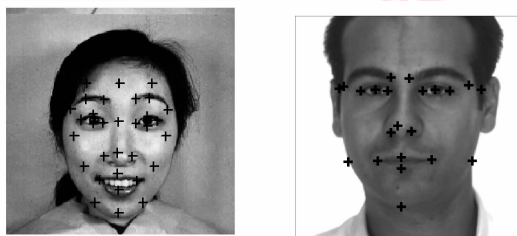
3.1 面部表情特征提取

本节主要研究面部图像的特征提取问题,包括图像的面部几何特征提取和 Gabor 小波特征提取两部分。其中,几何特征由脸部若干关键特征点的坐标值组成,刻画了面部肌肉的运动变化情况,可用于描述面部表情信息。而 Gabor 特征则是运用 Gabor 小波变换技术对面部图像进行小波变换后得到的特征,它既包含了人脸的局部特征,又包含了人脸图像的全局特征。

3.1.1 几何特征

为了提取脸部的几何特征,在每幅人脸图像中标定出若干个关键点位置,并将关键点的坐标值组成的矢量定义为几何特征矢量。面部关键点位置可有不同的定义方法:Zhang 等人^[12]提出了 34 个关键点的定义方法(见图 1(a)所示),这种定义方法已广泛用于 JAFFE (Japanese female facial expression)

面部表情数据库中^[10-11,13];另一种定义方法可见图 1(b),该方法在面部仅定义了 22 个关键特征点。文中采用第一种关键点的定义方式进行相关实验,关键点的位置采用人工标定。为了获得脸部关键点的自动标定,可结合主动形状模型(ASM)方法^[14]来实现。



(a) 34个基准点

(b) 22个基准点

图 1 基准点示例

Fig. 1 Example of the landmark points

3.1.2 Gabor 特征

Gabor 小波变换特征是人脸表情特征表示的一种重要方式。人脸图像的 Gabor 小波特征可由下述方法得到:首先用一组 Gabor 滤波器(6 个方向和 5 个尺度)对人脸图像进行卷积运算,再取卷积图像在关键点位置的幅值作为 Gabor 小波在该处的 Gabor 特征。典型的 Gabor 小波核函数定义如下:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right) \times \left[\exp(i k_{u,v} \cdot z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (19)$$

式中, u 和 v 分别表示 Gabor 函数的方向和尺度, $k_{u,v}$ 定义为: $k_{u,v} = k_v \exp(i\phi_u)$ 。本文采用与文献[11]相同的 Gabor 小波函数,即 $k_v = \pi/2^v$, ($v \in \{1, \dots, 5\}$), $\phi_u = \pi u/6$, ($u \in \{0, \dots, 5\}$), 且 $\sigma = \pi$ 。

3.2 基于 CCA 的表情特征融合方法

假设 f_x 和 f_y 分别代表面部表情的几何特征矢量和 Gabor 特征矢量,则典型相关分析(CCA)的目的就是寻找投影方向 ω_x 和 ω_y ,使得 $\omega_x^T f_x$ 和 $\omega_y^T f_y$ 相关系数 $\rho(f_x, f_y, \omega_x, \omega_y)$ 最大,其中

$$\begin{aligned} \rho(f_x, f_y, \omega_x, \omega_y) &= \frac{E\{\omega_x^T f_x f_y^T \omega_y\}}{\sqrt{E\{\omega_x^T f_x f_x^T \omega_x\}} \sqrt{E\{\omega_y^T f_y f_y^T \omega_y\}}} \\ &= \frac{\omega_x^T E\{f_x f_y^T\} \omega_y}{\sqrt{\omega_x^T E\{f_x f_x^T\} \omega_x} \sqrt{\omega_y^T E\{f_y f_y^T\} \omega_y}} \end{aligned} \quad (20)$$

假设 $\{f_x^i, f_y^i\}_{i=1, \dots, N}$ 为 $\{f_x, f_y\}$ 的 N 个观察值,则 CCA

的目标就等价于寻找 ω_x 和 ω_y ,使得式(21)最大。

$$\begin{aligned} \{\omega_x^*, \omega_y^*\} &= \arg \max_{\omega_x, \omega_y} \rho(\omega_x, \omega_y) \\ &= \arg \max_{\omega_x, \omega_y} \left\{ \frac{\omega_x^T X Y^T \omega_y}{\sqrt{\omega_x^T X X^T \omega_x} \sqrt{\omega_y^T Y Y^T \omega_y}} \right\} \end{aligned} \quad (21)$$

式中, $X = [f_x^1, f_x^2, \dots, f_x^N]$, $Y = [f_y^1, f_y^2, \dots, f_y^N]$ 。式(21)所示的优化问题可用式(22)的拉格朗日算子法求解。

$$\begin{aligned} L(\omega_x, \omega_y, \lambda, \mu) &= \omega_x^T X Y^T \omega_y - \lambda (\omega_x^T X X^T \omega_x - 1)/2 - \\ &\quad \mu \frac{(\omega_y^T Y Y^T \omega_y - 1)}{2} \end{aligned} \quad (22)$$

L 分别对 ω_x 和 ω_y 求偏导,并令其等于零,可得

$$\begin{aligned} \frac{\partial L}{\partial \omega_x} &= X Y^T \omega_y - \lambda X X^T \omega_x = 0 \\ \frac{\partial L}{\partial \omega_y} &= Y X^T \omega_x - \mu Y Y^T \omega_y = 0 \end{aligned} \quad (23)$$

由式(23)可得 $\mu = \lambda$, $\omega_y = \frac{(Y Y^T)^{-1} Y X^T}{\mu} \omega_x$, 且

$$X Y^T (Y Y^T)^{-1} Y X^T \omega_x = \lambda^2 X X^T \omega_x \quad (24)$$

求解式(24)所示的广义特征方程问题可通过对矩阵 $X X^T$ ^[15] 应用完全 Cholesky 分解方法,从而转换为一个对称的特征方程问题。

假设 (ω_x^i, ω_y^i) ($i = 1, 2, \dots, k$) 是 CCA 所求解的 k 组投影矢量。令 $W_x = [\omega_x^1, \omega_x^2, \dots, \omega_x^k]$ 及 $W_y = [\omega_y^1, \omega_y^2, \dots, \omega_y^k]$ 。则 f_x 和 f_y 到 W_x 和 W_y 上的投影值分别为 $p_x = W_x^T f_x$ 和 $p_y = W_y^T f_y$ 。于是,把 $p = p_x + p_y$ 作为 f_x 和 f_y 的融合特征。

3.3 基于 FKDA 的面部表情识别

假设 p_i ($i = 1, 2, \dots, N$) 是对应于 N 幅面部表情图像的特征矢量,且 u_{ij} ($i = 1, \dots, c; j = 1, \dots, N$) 是第 j 个样本隶属于第 i 类的隶属度,则 u_{ij} 可采用 Keller 等人^[16]提出的模糊 K 近邻(FK-NN)方法计算得到。下面是 FK-NN 算法计算 FKDA 类隶属度的伪代码:

算法:计算模糊核判别分析的隶属度

输入:(1)训练样本;(2)邻居的数目 k ;

输出:隶属度 μ_{ij} ($i = 1, \dots, c; j = 1, \dots, N$):

(1) 计算任意两个训练样本 $\Phi(p_p)$ 和 $\Phi(p_q)$ 之间的欧氏距离,用 d_{pq} 表示,其中

$$\begin{aligned} d_{pq}^2 &= \|\Phi(p_p) - \Phi(p_q)\|^2 \\ &= (\Phi(p_p) - \Phi(p_q))^T (\Phi(p_p) - \Phi(p_q)) \end{aligned}$$

$$\begin{aligned} &= (\Phi(\mathbf{p}_p))^T \Phi(\mathbf{p}_p) - (\Phi(\mathbf{p}_q))^T \Phi(\mathbf{p}_p) - \\ &\quad (\Phi(\mathbf{p}_p))^T \Phi(\mathbf{p}_q) + (\Phi(\mathbf{p}_q))^T \Phi(\mathbf{p}_q) \\ &= k(\mathbf{p}_p, \mathbf{p}_p) - 2k(\mathbf{p}_p, \mathbf{p}_q) + k(\mathbf{p}_q, \mathbf{p}_q) \end{aligned}$$

(2) 令 $\mathbf{D}$ 是一个 $N \times N$ 阶的矩阵,元素由 d_{pq} ($p, q = 1, \cdots, N; p \neq q$) 组成。设 $\mathbf{D}$ 的对角元素为无穷大;

(3) 对 $\mathbf{D}$ 的每列元素以升序进行排序,选择所有模式中位于最近邻的模式类别标签;

(4) 利用如下方法计算第 j 个样本属于第 i 类的隶属度:如果 i 等于第 j 个训练样本的类别标签,那么 $\mu_{ij} = 0.51 + 0.49 (n_{ij}/k)$; 否则 $\mu_{ij} = 0.49 (n_{ij}/k)$ 。其中 n_{ij} 代表与第 j 个样本相邻的,且属于第 i 类的样本数目。

4 实 验

分别采用国际著名的 JAFFE 面部表情数据库^[12]和 Ekman 的“面部表情图片”数据库^[17]来进行面部表情识别的实验。实验中分别选择单项式核和高斯核来计算 FKDA 方法中的 Gram 矩阵。两种核函数的定义如下:

单项式核函数: $k(\mathbf{x}, \mathbf{y}) = (\langle \mathbf{x}, \mathbf{y} \rangle)^d$,其中 d 是单项式核函数的度;

高斯核函数: $k(\mathbf{x}, \mathbf{y}) = \exp \{ \|\mathbf{x} - \mathbf{y}\|^2 / \sigma \}$,其中 σ 是高斯核函数的参数。

4.1 JAFFE 数据库上的实验

JAFFE 数据库由 10 位日本女性的 213 张人脸图像构成,这些人脸图像覆盖了全部的 7 种基本面部表情类型(即高兴、悲伤、惊奇、生气、厌恶、害怕和中性)。其中,每个人有 2 到 4 张每种表情的图片。JAFFE 数据库中每张图像是灰度级为 256、图像尺寸大小为 256×256 的灰度图像。实验采用“leave-one-class-out”交叉验证策略进行:将属于同一个人的图像作为测试数据,而剩余的部分则作为训练数据。实验重复进行直到所有人的 人脸图像都被用做测试数据。然后再将每次实验所得的识别率求平均作为最终的 表情识别率。为进行对比,也采用 KDA 方法完成了同样的实验。图 2 是识别率与高斯核函数参数 σ 之间的关系图,从中可以看出当 σ 取值在 1×10^5 附近时,识别率达到最高。表 1 是各种方法的识别率比较,从中可以看出,本文方法获得了最好的识别效果,识别率高达 82.51%。

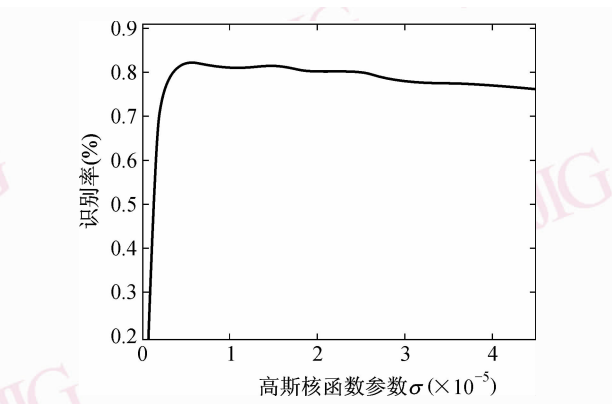


图 2 表情识别率与高斯核函数参数 σ 之间的关系图
Fig. 2 Recognition rate with the different choices of σ

表 1 采用“leave-one-class-out”交叉验证方法得到的面部表情识别率

方法	面部表情的平均识别率 (%)
FKDA + Gabor 特征 + 几何特征 (高斯核 $\sigma = 9 \times 10^4$)	82.51
FKDA + Gabor 特征 + 几何特征 (单项式核 $d = 3$)	80.87
FKDA + Gabor 特征 (高斯核 $\sigma = 2 \times 10^6$)	78.14
KDA + Gabor 特征 (高斯核 $\sigma = 2 \times 10^6$)	77.05
FKDA + Gabor 特征 (单项式核 $d = 3$)	76.50
KDA + Gabor 特征 (单项式核 $d = 3$)	68.85
KCCA + Gabor 特征 ^[10]	77.05
LDA + Gabor 特征 ^[11]	75.00

4.2 Ekman 面部表情数据库上的实验

Ekman 面部表情数据库由 6 位男性和 8 位女性的 110 张人脸图像构成。其中,每个人有 1 到 2 张每种表情的图片。原始图像尺寸大小为 768×512

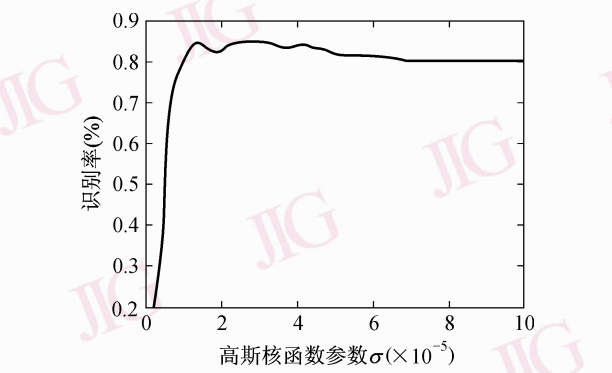


图 3 表情识别率与高斯核函数参数 σ 之间的关系图
Fig. 3 Recognition rate with the different choices of σ

像素,每个像素用 256 级的灰度值来表示。也采用“leave-one-class-out”交叉验证策略进行实验。图 3 是识别率与高斯核函数参数 σ 之间的关系图,从中可以看出当 σ 取值在 2×10^5 附近时,识别率达到最高。表 2 显示了各种方法的实验结果,从中可以看出,本文方法获得了最好的识别效果,识别率高达 84.38%。

表 2 采用“leave-one-class-out”交叉验证方法得到的面部表情识别率

Tab. 2 Comparison of Facial Expression Recognition using “leave-one-class-out” method

方法	面部表情的平均识别率 (%)
FKDA + Gabor 特征 + 几何特征 (高斯核 $\sigma = 2.2e5$)	84.38
FKDA + Gabor 特征 + 几何特征 (单项式核 $d = 2$)	80.21
FGDA + Gabor 特征 (高斯核 $\sigma = 7e6$)	82.29
GDA + Gabor 特征 (高斯核 $\sigma = 7e6$)	78.13
FGDA + Gabor 特征 (单项式核 $d = 2$)	79.17
GDA + Gabor 特征 (单项式核 $d = 2$)	76.04
KCCA + Gabor 特征 ^[10]	77.08
LDA + Gabor 特征 ^[11]	82.00

5 结 论

提出了一种基于表情特征融合和 FKDA 的面部表情识别方法。FKDA 方法通过引入样本的类隶属度信息弥补了传统的核判别分析(KDA)方法的不足,而基于 CCA 方法的表情特征融合方法通过 CCA 技术建立了表情的 Gabor 特征和几何特征之间的相关性,通过这种相关性所提取的表情特征可有效融合表情的全局和局部信息,更适合作为表情识别的特征。通过在两个著名的国际面部表情数据库上的实验,验证了本文方法的有效性。

参考文献 (References)

1 Fisher R A. The use of multiple measurements in taxonomic problems [J]. Annals of Eugenics, 1936, 7(2):179-188.
2 Mika S, Rätsch G, Weston J, et al. Fisher discriminant analysis with kernels[A]. In: Proceedings of IEEE International Workshop on Neural Networks for Signal Processing IX[C], Madison, WI, USA,

1999;41-48.
3 Baudat G, Anouar F. Generalized discriminant analysis using a kernel approach [J]. Neural Computation, 2000, 12 (10): 2385-2404.
4 Zadeh L A. Fuzzy sets[J]. Information and Control, 1965, 8(3): 338-353.
5 Kwak K C, Pedrycz W. Face recognition using a fuzzy fisherface classifier[J]. Pattern Recognition, 2005, 38(10):1717-1732.
6 Zheng Y J, Yang J, Yang J Y, et al. A reformative kernel discriminant algorithm and its application to face recognition [J]. Neurocomputing, 2006,69(13-15):1806-1810.
7 Wu Q, Zhou X. Zheng W. Facial expression recognition using fuzzy kernel discriminant analysis [A]. In: Proceedings of the 3rd International Conference on Fuzzy Systems and Knowledge Discovery [C], Xi'an, China, 2006, 4223:780-783.
8 Pantic M, Leon Rothkrantz J M. Automatic analysis of facial expressions: the state of the art [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000,22(12):1424-1445.
9 Fasel B, Luettin J. Automatic facial expression analysis: A survey [J]. Pattern Recognition, 2003,36(1):259-275.
10 Zheng W, Zhou X, Zou C, et al. Facial expression recognition using kernel canonical correlation analysis(KCCA)[J]. IEEE Transactions on Neural Networks, 2006,17(1):233-238.
11 Lyons M, Budynek J, Akamatsu S. Automatic classification of single facial images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(12):1357-1362.
12 Zhang Z, Lyons M, Schuster M, et al. Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptrons [A]. In: Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition [C], Nara, Japan, 1998:454-459.
13 Guo G, Dyer C R. Learning from examples in the small sample cases: face expression recognition [J]. IEEE Transactions on Systems, Man and Cybernetics-Part B: Cybernetics, 1995, 35(3): 477-488.
14 Lanitis A, Taylor C J, Cootes T F. Automatic interpretation and coding of face images using flexible models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19 (7): 743-756.
15 Hardoon D R, Szedmark S, Shawe-Taylor J. Caonical correlation analysis: an overview with application to learning methods [J]. Neural Computation, 2004, 16(12):2639-2664.
16 Keller J M, Gray M R, Givens J A. A fuzzy K-Nearest neighbor algorithm[J]. IEEE Transactions on Systems Man and Cybernetics, 1985,15(4):580-585.
17 Ekman P, Friesen W V. Pictures of Facial affect[M]. Palo Alto, CA, USA: Consulting Psychologists Press, 1976.