

基于偏最小二乘回归的人脸身份和表情同步识别方法

周晓彦^{1),2)} 郑文明³⁾ 赵力¹⁾ 邹采荣^{1),4)}

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 将偏最小二乘回归方法用于人脸身份和表情的同步识别。首先,对每幅人脸图像进行脸部特征提取以及相应的语义特征定义。在脸部特征提取方面,从每幅图像中标定出若干脸部关键点位置,并提取图像在该关键点处的 Gabor 小波系数 (Gabor 特征) 以及关键点的坐标值 (几何特征),作为该图像的输入特征。语义特征则定义为该人脸图像所属的表情类别信息以及所对应的人脸身份信息。其次,利用核主成分分析 (KPCA) 方法对脸部 Gabor 特征和几何特征进行融合,使得输入特征具有更好的识别特性;最后,运用偏最小二乘回归 (PLSR) 方法建立脸部特征和语义特征之间的关系模型,并运用此模型对某一测试人脸图像进行表情和身份的同步识别。通过在 JAFFE 国际表情数据库和 AR 人脸数据库上的对比实验,证实了所提方法的有效性。

关键词 偏最小二乘回归 人脸识别 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)05-801-08

Simultaneous Recognition of Face and Facial Expression Via Partial Least Squares Regression

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

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, we utilize the partial least squares regression (PLSR) method to solve the simultaneous recognition problem on face identity and face expression. Firstly, face features and the corresponding semantic features are extracted for each face image as the input features, where the face features are defined as the Gabor wavelet coefficients defined on several land marks of each face image; the geometric features are defined as the coordinates of the landmark points and the semantic features are defined as the facial expression category index and face identity index of each face image. The kernel principal component analysis (KPCA) method is then applied to deal with the feature fusion task of both Gabor features and geometric features. Finally, the PLSR method is used to model the correlation between the input facial feature vectors and the semantic vectors. Based on this model, both face identity and facial expression category of any test facial image can be predicted. Experiments on both JAFFE facial expression database and AR face database show the effectiveness of the proposed method.

Keywords partial least squares regression, face recognition, 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基础研究自筹经费项目(08KJD520009)

收稿日期:2008-07-14;**改回日期:**2008-09-05

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

人脸面部图像中包含着丰富的信息,通过面部图像不仅可以识别人的身份,而且可以识别人的面部表情。目前,面部表情识别^[1]和身份识别^[2]已成为计算机视觉和模式识别领域中的两大热点研究问题。面部表情识别研究的主要任务是从人脸图像中提取能反映情绪类别的主要特征,并在此基础上进行表情分类。根据 Ekman 等人的研究工作^[3],面部表情被分为六种基本类型(高兴、悲伤、惊奇、生气、厌恶和害怕)。一些研究学者也将中性表情作为第七种基本情绪^[4]。目前,大部分传统的面部表情识别方法是面部图像划分为七种基本表情类型中的一种。表情识别的另一种常用方法是参照 Ekman 等人提出的面部运动编码系统(FACS),通过识别 FACS 中定义的运动单元(AUs)来达到识别表情的目的^[5]。与面部表情识别方法类似,人脸识别的目标是将未知身份的人脸图像匹配到某一已知类别的人脸图像,从而达到人脸身份识别的目的^[6]。

由于识别目标的不同,人脸识别和表情识别往往被作为两个独立的任务来完成。然而,在某些场合下,进行人脸身份和表情的同步识别具有一定的特殊意义。例如,在基于面部表情的儿童情绪类型鉴别研究中,不仅需要识别儿童的面部表情,而且还要了解不同儿童个体间的情绪差异。因此,如果能同时识别儿童的身份和表情类型,将能更好地完成鉴别的任务。此外,心理学的研究也表明:人脸的身份识别和面部表情识别存在着交互作用^[7]。针对这一问题,汪亚珉和傅小兰提出了人脸身份识别和面部表情识别的多级整合模型。鉴于这一原因,本文借鉴了多级整合模型的思想,将偏最小二乘回归(PLSR)^[8]方法用于人脸身份和面部表情的同步识别。该方法首先从人脸图像中提取若干反映身份和表情的特征,并建立与之相应的语义特征矢量。在特征提取方面,由于人脸图像的 Gabor 小波表示在保留空间关系的同时还描述了图像的空间频率结构信息,能较好地刻画图像的局部特征,并且已在人脸识别和表情识别中分别得到了成功的应用^[4,9]。因此,本文采用 Gabor 小波方法来提取人脸的图像特征。除此之外,在脸部特征提取方面,为了更充分利用图像所包含的信息,还在每幅图像中标定出若干脸部关键点位置,并提取图像在该关键点处的坐标

值作为图像的几何特征,然后把几何特征和 Gabor 特征一起作为人脸身份和表情同步识别的输入特征。为了获得更好的识别性能,应用核主成分分析(KPCA)^[10]方法对 Gabor 特征和几何特征进行特征融合,使得输入特征具有更好的识别特性。最后,在此基础上利用 PLSR 方法建立脸部特征矢量和语义特征矢量之间的关系式,再根据该关系式预测某一待测人脸图像的身份和表情的类别。

2 相关工作回顾

2.1 偏最小二乘回归

PLSR 算法最早由 Wold^[8]于 1966 年提出,现已在经济学、化学、心理学等研究领域得到了广泛应用。最近,Baek 和 Kim^[11]运用 PLSR 方法进行人脸识别,获得了比主成分分析(PCA)方法^[12]更好的识别效果。此外,最近的研究也显示:在需要从一组大量的独立变量集预测另一组相关变量集的情况下,PLSR 方法具有显著的效果^[13]。

假设 $\{\mathbf{x}_i\}_{i=1}^N \in \mathbf{R}^p$ 表示一组 p 维的输入变量,而 $\{\mathbf{y}_i\}_{i=1}^N \in \mathbf{R}^q$ 则表示相应的 q 维的输出响应。PLSR 的目标在于建立输入变量和相应的输出响应之间的关系模型。目前为止,已出现了许多实现 PLSR 方法的算法,本文采用由 Wold^[14]提出的经典形式,它基于非线性迭代偏最小二乘(NIPALS)算法。Rosipal 等给出了该算法的详细实现流程,并称之为 NIPALS-PLS 算法^[15]。假设 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$ 和 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]^T$, 其中 $\mathbf{x}^T$ 表示 $\mathbf{x}$ 的转置,则 NIPALS-PLS 算法可按如下步骤完成^[15],其中, $\mathbf{u}_{(r)}$ 和 $\mathbf{t}_{(r)}$ 分别为 PLSR 中关于 $\mathbf{Y}_{(r)}$ 和 $\mathbf{X}_{(r)}$ 的隐变量。

(1) $\mathbf{X}_{(1)} = \mathbf{X}$, $\mathbf{Y}_{(1)} = \mathbf{Y}$;

(2) 对 $r = 1, \dots, m$ 重复

① 随机地初始化 $\mathbf{u}_{(r)}$;

② $\mathbf{w}_{(r)} = \mathbf{X}_{(r)}^T \mathbf{u}_{(r)}$;

③ $\mathbf{t}_{(r)} = \mathbf{X}_{(r)} \mathbf{w}_{(r)}$, $\mathbf{t}_{(r)} = \mathbf{t}_{(r)} / \|\mathbf{t}_{(r)}\|$;

④ $\mathbf{c}_{(r)} = \mathbf{Y}_{(r)}^T \mathbf{t}_{(r)}$;

⑤ $\mathbf{u}_{(r)} = \mathbf{Y}_{(r)} \mathbf{c}_{(r)}$, $\mathbf{u}_{(r)} \leftarrow \mathbf{u}_{(r)} / \|\mathbf{u}_{(r)}\|$;

⑥ 重复步骤①~⑤,直到收敛为止;

⑦ 对矩阵 $\mathbf{X}_{(r)}$ 和 $\mathbf{Y}_{(r)}$ 进行如下 rank-one 更新

i. $\mathbf{X}_{(r)} \leftarrow \mathbf{X}_{(r)} - \mathbf{t}_{(r)} \mathbf{t}_{(r)}^T \mathbf{X}_{(r)}$;

ii. $\mathbf{Y}_{(r)} \leftarrow \mathbf{Y}_{(r)} - \mathbf{t}_{(r)} \mathbf{t}_{(r)}^T \mathbf{Y}_{(r)}$;

(3) 回归方程可写成

$$\hat{\mathbf{Y}} = \mathbf{X}^T \mathbf{U} (\mathbf{T}^T \mathbf{X} \mathbf{X}^T \mathbf{U})^{-1} \mathbf{T}^T \mathbf{Y}$$

其中, $\mathbf{U} = [\mathbf{u}_{(1)}, \mathbf{u}_{(2)}, \dots, \mathbf{u}_{(m)}]$, $\mathbf{T} = [\mathbf{t}_{(1)}, \mathbf{t}_{(2)}, \dots, \mathbf{t}_{(m)}]$ 。

令 $\mathbf{x}_{\text{test}}$ 表示一个测试输入数据, 那么相应的响应 $\hat{\mathbf{y}}_{\text{test}}$ 可通过式(1)算得

$$\hat{\mathbf{y}}_{\text{test}} = \mathbf{x}_{\text{test}}^T \mathbf{U} (\mathbf{T}^T \mathbf{X} \mathbf{X}^T \mathbf{U})^{-1} \mathbf{T}^T \mathbf{Y} \quad (1)$$

2.2 核主成分分析

假设 Φ 是输入空间到再生核希尔伯特空间 (RKHS) $\mathbf{F}$ 中的某一非线性映射, 即 $\Phi: \mathbf{R}^p \rightarrow \mathbf{F}$, $\mathbf{x}_i \rightarrow \Phi(\mathbf{x}_i)$ 。令 $\mathbf{K} = \{k_{ij}\}$ 表示 $N \times N$ 阶的 Gram 阵, 其元素定义为 $k_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$, 其中 $k(\mathbf{x}_i, \mathbf{x}_j)$ 为某一核函数, 该函数可表示成 $\Phi(\mathbf{x}_i)$ 和 $\Phi(\mathbf{x}_j)$ 的内积值, 即

$$k(\mathbf{x}_i, \mathbf{x}_j) = (\Phi(\mathbf{x}_i))^T \Phi(\mathbf{x}_j) \quad (2)$$

令 $\Phi(\mathbf{X}) = [\Phi(\mathbf{x}_1), \Phi(\mathbf{x}_2), \dots, \Phi(\mathbf{x}_N)]$, 则求解 KPCA 问题等价于求解对角化(式(3))的协方差矩阵问题。

$$\mathbf{C}^\Phi = \frac{1}{N} \sum_{i=1}^N (\Phi(\mathbf{x}_i) - \mathbf{u}^\Phi)(\Phi(\mathbf{x}_i) - \mathbf{u}^\Phi)^T \quad (3)$$

式中, $\mathbf{u}^\Phi = \sum_{i=1}^N \Phi(\mathbf{x}_i) / N$ 表示经 Φ 映射后的输入数据点的均值。求解 $\mathbf{C}^\Phi$ 对角化问题的标准方法是求解特征方程 $N\lambda \mathbf{v} = \mathbf{C}^\Phi \mathbf{v}$ 的特征矢量 $\mathbf{v}$ 和相应的特征值 $\lambda \geq 0$ 的问题。由文献[10]可知, 特征矢量 $\mathbf{v}$ 可写成如下形式:

$$\mathbf{v} = \sum_{i=1}^N \alpha_i (\Phi(\mathbf{x}_i) - \mathbf{u}^\Phi) = \Phi(\mathbf{X}) (\mathbf{I} - \mathbf{M}) \boldsymbol{\alpha} \quad (4)$$

式中, $\boldsymbol{\alpha} = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_N]^T$, $\mathbf{I}$ 是 $N \times N$ 阶的单位阵, $\mathbf{M}$ 是元素为 $1/N$ 的 $N \times N$ 阶矩阵。将式(4)代入式(3), 再联合式(2)可得

$$N\lambda \mathbf{K}_n \boldsymbol{\alpha} = \mathbf{K}_n \mathbf{K}_n \boldsymbol{\alpha} \Rightarrow N\lambda \boldsymbol{\alpha} = \mathbf{K}_n \boldsymbol{\alpha} \quad (5)$$

式中, $\mathbf{K}_n = (\mathbf{I} - \mathbf{M}) \mathbf{K} (\mathbf{I} - \mathbf{M})$ 。测试样本 $\Phi(\mathbf{t})$ 到 $\mathbf{v}$ 上的投影可由下式算得: $\mathbf{v}^T \Phi(\mathbf{t}) = \boldsymbol{\alpha}^T \mathbf{K}_{\text{test}}$, 其中 $\mathbf{K}_{\text{test}} = [\kappa_{\text{test}}^1 \ \kappa_{\text{test}}^2 \ \dots \ \kappa_{\text{test}}^N]^T$, $\kappa_{\text{test}}^i = k(\mathbf{x}_i, \mathbf{t})$ 。

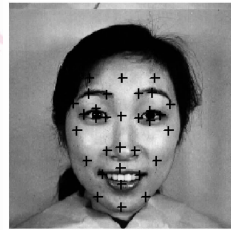
3 面部图像的特征提取

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、Gabor 小波特征提取以及语义特征提取等 3 个部分。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

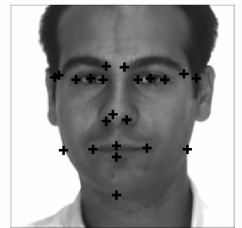
像进行小波变换后得到的特征, 它既包含了人脸的局部特征, 又包含了人脸图像的全局特征。人脸图像的语义特征则由该人脸图像所属的表情类别信息以及所对应的人脸身份信息组成。在以上三种特征中, 几何特征和 Gabor 特征反映了人脸图像的外在信息, 用于人脸识别和表情识别的输入特征。语义特征反映了图像类别信息, 是进行人脸身份和表情识别的依据。

3.1 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^[16]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 面部表情数据库中^[4,17-18]; 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。在本文中, 以上两种定义方式均被采用。在本文中, 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^[19]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

3.2 Gabor 特征

典型的 Gabor 小波核函数定义如下:

$$\psi_{u,v}(z) = \frac{\|\mathbf{k}_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|\mathbf{k}_{u,v}\|^2 \|\mathbf{z}\|^2}{2\sigma^2}\right) \left[\exp(i\mathbf{k}_{u,v} \cdot \mathbf{z}) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (6)$$

式中, u 和 v 分别表示 Gabor 函数的方向和尺度, $\mathbf{k}_{u,v}$ 定义为: $\mathbf{k}_{u,v} = \mathbf{k}_v \exp(i\phi_u)$ 。与文献[4]所采用的 Gabor 小波相同, 本文所定义的 Gabor 小波的方向和尺度分别为 6 个方向和 5 个尺度, 即: $\mathbf{k}_v = \pi/2^\circ$, ($v \in \{1, \dots, 5\}$), $\phi_u = \pi u/6$, ($u \in \{0, \dots, 5\}$), 且 $\sigma = \pi$ 。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。图 2 显示了 Gabor 小波变换的示意图,其中符号 $\otimes$ 表示卷积运算,右边显示的图像是由人脸图像和 Gabor 小波滤波器组经过卷积运算后得到的幅值图像。

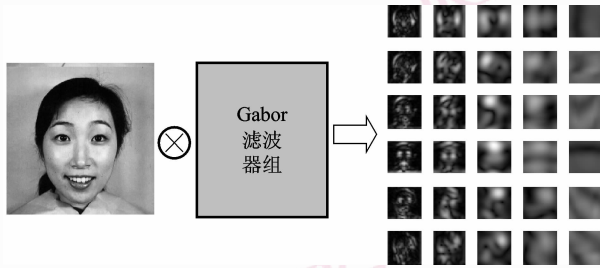


图 2 Gabor 小波变换示意图

Fig. 2 Example of Gabor wavelet transformation

3.3 人脸图像的语义特征

人脸图像的语义特征定义为图像所属的 7 种基本表情(包括中性情绪)类别信息以及所对应的人脸身份信息组成的二值特征矢量。假设 c 表示人脸的类别个数, y_i 表示与人脸图像 x_i 相对应的语义特征矢量,并且假设每幅人脸图像属于 7 种基本表情中的一种,则 y_i 是一个 $c+7$ 维的矢量,它的元素取 0 或 1。更具体地说,若 x_i 属于第 k 个人的某一幅图像,且 x_i 属于 7 种基本情绪中的第 j 种基本表情,其中 $1 \leq k \leq c$ 。则 y_i 可定义为

$$y_i = e_j + e_{7+k} \quad (7)$$

式中, e_j 是一个第 j 个元素等于 1 的 $c+7$ 维的单位矢量。

4 人脸身份和表情同步识别方法

本节将详细介绍基于 PLSR 的人脸身份和表情的同步识别方法。图 3 显示了人脸身份和表情同步识别方法的系统框架图。

由图 3 可以看出,整个识别程序可分为三部分内容:(1)面部特征提取部分;(2)基于 KPCA 的特征融合部分;(3)基于 PLSR 的人脸身份和表情的同步识别部分。下面详细介绍基于 KPCA 的特征融合技术和基于 PLSR 的识别方法。

假设 x_i^1 和 x_i^2 ($i=1, \dots, N$) 分别表示从 N 个训练样本所提取的 N 个几何特征矢量及所对应的 Gabor 特征矢量。设 Φ_1 和 Φ_2 为两个非线性映射函

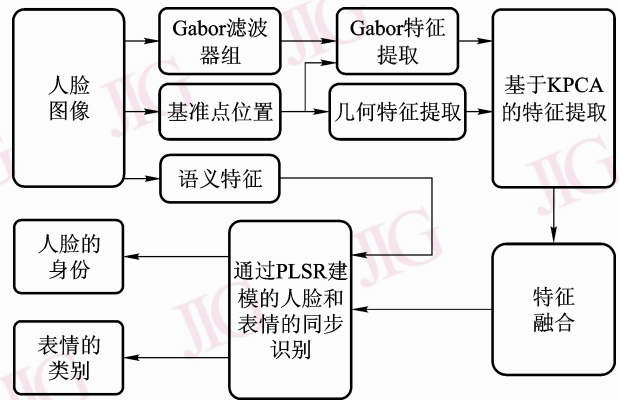


图 3 人脸身份和表情同步识别的系统框架

Fig. 3 System framework of simultaneous recognition of face and facial expression

数,分别将 x_i^1 和 x_i^2 ($i=1, \dots, N$) 从输入的 n_1 和 n_2 维的空间映射到高维的再生核希尔伯特空间 F_1 和 F_2 中,即

$$\Phi_1: R^{n_1} \rightarrow F_1, x_i^1 \rightarrow \Phi_1(x_i^1)$$

$$\Phi_2: R^{n_2} \rightarrow F_2, x_i^2 \rightarrow \Phi_2(x_i^2)$$

令 K^1 和 K^2 分别代表数据集 $\{\Phi_1(x_i^1)\}$ 和 $\{\Phi_2(x_i^2)\}$ 所对应的 Gram 矩阵。令 $K_n^1 = (I - M)K^1(I - M)$, $K_n^2 = (I - M)K^2(I - M)$ 。其中, M 为一个元素为 $1/N$ 的 $N \times N$ 阶矩阵。假设 α_i^1 ($i=1, \dots, m_1$) 是矩阵 K_n^1 前 m_1 个最大特征值所对应的特征矢量,而 α_i^2 ($i=1, \dots, m_2$) 是 K_n^2 前 m_2 个最大特征值所对应的特征矢量,令

$$v_i^1 = \Phi(X^1)(I - M)\alpha_i^1 \quad i=1, \dots, m_1 \quad (8)$$

$$v_i^2 = \Phi(X^2)(I - M)\alpha_i^2 \quad i=1, \dots, m_2 \quad (9)$$

式中, $\Phi(X^1) = [\Phi(x_1^1) \quad \Phi(x_2^1) \quad \dots \quad \Phi(x_N^1)]$, $\Phi(X^2) = [\Phi(x_1^2) \quad \Phi(x_2^2) \quad \dots \quad \Phi(x_N^2)]$ 。则由文献[15]可知, v_i^1 ($i=1, \dots, m_1$) 是 K^1 前 m_1 个最大特征值所对应的特征矢量, v_i^2 ($i=1, \dots, m_2$) 是 K^2 前 m_2 个最大特征值所对应的特征矢量。

令 $V_1 = [v_1^1 \quad v_2^1 \quad \dots \quad v_{m_1}^1]$ 和 $V_2 = [v_1^2 \quad v_2^2 \quad \dots \quad v_{m_2}^2]$ 。假设 u_1^Φ 和 u_2^Φ 分别代表集合 $\{\Phi(x_i^1), i=1, \dots, N\}$ 和 $\{\Phi(x_i^2), i=1, \dots, N\}$ 的均值,则 $\Phi(x_i^1) - u_1^\Phi$ 和 $\Phi(x_i^2) - u_2^\Phi$ 到 V_1 和 V_2 上的投影可分别计算为

$$z_i^1 = V_1^T (\Phi(x_i^1) - u_1^\Phi) = A_1^T (K_n^1)_i$$

$$z_i^2 = V_2^T (\Phi(x_i^2) - u_2^\Phi) = A_2^T (K_n^2)_i \quad (10)$$

式中, $A_1 = [\alpha_1^1 \quad \alpha_2^1 \quad \dots \quad \alpha_{m_1}^1]$, $A_2 = [\alpha_1^2 \quad \alpha_2^2 \quad \dots \quad \alpha_{m_2}^2]$,

$(\mathbf{K}_n^1)_i$ 和 $(\mathbf{K}_n^2)_i$ 分别代表 $\mathbf{K}_n^1$ 和 $\mathbf{K}_n^2$ 的第 i 列。对特征矢量 $\mathbf{z}_i^1$ 和 $\mathbf{z}_i^2$ 进行规范化处理使得 $\|\mathbf{z}_i^1\| = 1$ 和 $\|\mathbf{z}_i^2\| = 1$ 。令 $\mathbf{z}_i = [(\mathbf{z}_i^1)^T, (\mathbf{z}_i^2)^T]^T$, 则特征矢量 $\mathbf{z}_i$ 为进行人脸身份和表情识别用的输入特征矢量。应用 PLSR 方法建立面部特征矢量 $\mathbf{z}_i$ 和相应语义矢量 $\mathbf{y}_i$ 之间的回归关系式为

$$\hat{\mathbf{Y}} = \mathbf{Z}^T \mathbf{U} (\mathbf{T}^T \mathbf{Z} \mathbf{Z}^T \mathbf{U})^{-1} \mathbf{T}^T \mathbf{Y} \quad (11)$$

式中, $\mathbf{Z} = [\mathbf{z}_1 \ \mathbf{z}_2 \ \cdots \ \mathbf{z}_N]$, $\mathbf{Y} = [\mathbf{y}_1 - \bar{\mathbf{y}} \ \mathbf{y}_2 - \bar{\mathbf{y}} \ \cdots \ \mathbf{y}_N - \bar{\mathbf{y}}]$, $\bar{\mathbf{y}}$ 是 $\mathbf{y}_i (i=1, \dots, N)$ 的均值, $\mathbf{U}$ 和 $\mathbf{T}$ 通过 NIPALS-PLS 算法求得。

假设 $\mathbf{x}_{\text{test}}$ 为一待测人脸图像, 则其所对应的语义特征矢量 $\hat{\mathbf{y}}_{\text{test}}$ 可通过式(12)求得

$$\hat{\mathbf{y}}_{\text{test}} = \mathbf{z}_{\text{test}}^T \mathbf{U} (\mathbf{T}^T \mathbf{Z} \mathbf{Z}^T \mathbf{U})^{-1} \mathbf{T}^T \mathbf{Y} + \bar{\mathbf{y}} \quad (12)$$

式中, $\mathbf{z}_{\text{test}} = [(\mathbf{z}_{\text{test}}^1)^T, (\mathbf{z}_{\text{test}}^2)^T]^T$, $\mathbf{z}_{\text{test}}^i (i=1, 2)$ 是测试图像 $\mathbf{x}_{\text{test}}$ 所提取的规范化几何特征矢量和 Gabor 特征矢量。假设 $(\hat{\mathbf{y}}_{\text{test}})_i$ 表示 $\hat{\mathbf{y}}_{\text{test}}$ 的第 i 个元素, 则 $\hat{\mathbf{y}}_{\text{test}}$ 的前 7 个元素组成的 7 维矢量 $\mathbf{y}_{\text{test}}^{\text{exp}} = [(\hat{\mathbf{y}}_{\text{test}})_1 \cdots (\hat{\mathbf{y}}_{\text{test}})_7]^T$ 代表该测试图像对 7 种基本表情的类属情况, 而 $\hat{\mathbf{y}}_{\text{test}}$ 剩下的元素组成的矢量 $\mathbf{y}_{\text{test}}^{\text{face}} = [(\hat{\mathbf{y}}_{\text{test}})_8 \cdots (\hat{\mathbf{y}}_{\text{test}})_{7+c}]^T$ 则表示该人脸图像对每个人的类属情况。假设 i^{exp} 表示最后分类得出的表情类别, 而 i^{face} 则表示人脸的身份类别, 则 i^{exp} 和 i^{face} 可分别由下式得到:

$$i^{\text{exp}} = \arg \max_i (\mathbf{y}_{\text{test}}^{\text{exp}})_i, i^{\text{face}} = \arg \max_i (\mathbf{y}_{\text{test}}^{\text{face}})_i \quad (13)$$

基于 KPCA 的特征融合和识别算法流程如下:

- (1) 输入数据 $\mathbf{x}_i^1$ 和 $\mathbf{x}_i^2 (i=1, \dots, N)$;
- (2) 运用 KPCA 算法求解数据集 $\mathbf{x}_i^1 (i=1, \dots, N)$ 的投影矢量 $\mathbf{v}_i^1 (i=1, \dots, m_1)$;
- (3) 运用 KPCA 算法求解数据集 $\mathbf{x}_i^2 (i=1, \dots, N)$ 的投影矢量 $\mathbf{v}_i^2 (i=1, \dots, m_2)$;

(4) 令 $\mathbf{V}_1 = [\mathbf{v}_1^1 \mathbf{v}_2^1 \cdots \mathbf{v}_{m_1}^1]$ 和 $\mathbf{V}_2 = [\mathbf{v}_1^2 \mathbf{v}_2^2 \cdots \mathbf{v}_{m_2}^2]$, 运用式(10)计算 $\Phi(\mathbf{x}_i^1) - \mathbf{u}_1^\Phi$ 和 $\Phi(\mathbf{x}_i^2) - \mathbf{u}_2^\Phi$ 到 $\mathbf{V}_1$ 和 $\mathbf{V}_2$ 上的投影值 $\mathbf{z}_i^1$ 和 $\mathbf{z}_i^2 (i=1, \dots, N)$, 并分别进行归一化处理;

(5) 令 $\mathbf{z}_i = [(\mathbf{z}_i^1)^T, (\mathbf{z}_i^2)^T]^T$, $\mathbf{Z} = [\mathbf{z}_1 \mathbf{z}_2 \cdots \mathbf{z}_N]$;

(6) 假设 $\mathbf{x}_{\text{test}}$ 为一待测人脸图像, 利用式(12)计算 $\mathbf{x}_{\text{test}}$ 所对应的语义特征矢量 $\hat{\mathbf{y}}_{\text{test}}$;

(7) 由式(13)得出身份和表情的类别。

5 实验

采用国际著名的 JAFFE 人脸数据库^[4,16] 以及 AR 人脸数据库^[20] 来进行人脸身份和表情同步识别的实验。考虑到几何特征和 Gabor 特征位于不同的度量空间, 因此实验中选择规范化的单项式核函数来计算 KPCA 方法中的 Gram 矩阵, 以消除不同度量空间中的特征差异。规范化的单项式核函数的定义如下^[6]:

$$k(\mathbf{x}, \mathbf{x}') = \frac{(\mathbf{x}^T \mathbf{x}')^d}{\sqrt{(\mathbf{x}^T \mathbf{x})^d} \sqrt{((\mathbf{x}')^T \mathbf{x}')^d}} \quad (14)$$

式中, d 是核函数的度。为了进一步验证所提方法的有效性, 采用了目前在人脸识别和表情识别中常用的其他方法进行了对比实验。

5.1 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(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。图 4 显示了数据库中的部分图像。

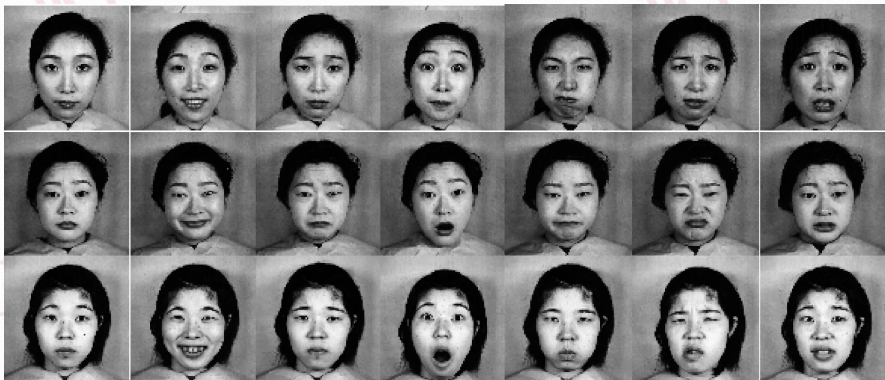


图4 JAFFE 数据库示例

Fig. 4 Examples of the JAFFE facial expression database

在进行面部特征提取时,一方面从每幅人脸图像中人工定位 34 个关键点(参见图 1(a)),并用这 34 个关键点的坐标值组成人脸的几何特征矢量。另一方面,对每幅人脸图像进行 Gabor 变换,并在 Gabor 变换后的图像中提取每个面部关键点处的幅值组成 Gabor 特征矢量。此外,为每幅图像建立相应的语义特征矢量。实验采用交叉验证策略进行,在交叉验证实验中,先将整个数据集随机地分为大致相同的 10 个子集,然后选取 9 个子集作为训练数据,剩下的一个作为测试数据。整个实验重复 10 次,直到每个子集都被用作测试数据。最后将 10 次实验得到的识别率求平均作为最后的平均识别率。本实验中采用度为 6 的单项式核函数。

图 5 中显示了交叉验证中每次实验的识别结果,而表 1 显示了在 JAFFE 数据库上采用交叉验证法时各种面部表情识别方法的平均识别率。从图 5 中可以看出人脸识别的平均识别率高达 100%。另外,从表 1 中可以看出本文方法获得了最好的表情识别效果,平均识别率高达 98.08%。

为了进一步测试本文方法在面部表情识别中的性能,采用如下新的交叉验证策略进行了同样的实验: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图像都被用作测试数据。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。实验中,单项式核函数的度取为 3。表 2 是各种方法的识别率比较,从表 2 中可以看出:本文方法获得了最好的识别效果,识别率高达 85.45%。

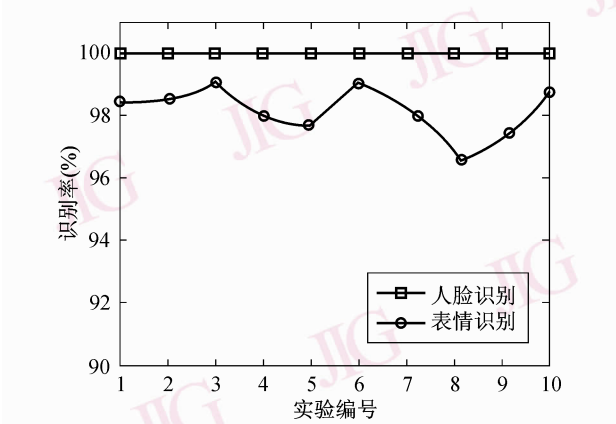


图 5 JAFFE 数据库上的人脸和表情识别率
Fig. 5 Face recognition rate and facial expression rate on JAFFE database

表 1 采用交叉验证方法得到的面部表情识别率
Tab. 1 Comparison of facial expression recognition using ten-fold cross validation

方法	面部表情的平均识别率 (%)
PLSR + Gabor + Geometry	98.08
NN + Gabor + Geometry ^[16]	92.30
LDA + Gabor ^[4]	92.00
Bayes + Gabor ^[17]	71.00
AdaBoost + Gabor ^[17]	71.90
Linear SVM + Gabor ^[17]	92.40
Nonlinear SVM + Gabor ^[17]	91.90
FLSP + Gabor ^[17]	91.00

表 2 采用新的交叉验证法得到的面部表情识别率
Tab. 2 Comparison of facial expression recognition using “leave-one-class-out” cross validation

方法	面部表情的平均识别率 (%)
PLSR + Gabor + Geometry	85.45
LDA + Gabor ^[4]	75.00
KCCA + Gabor ^[18]	77.05
K-NN + LBP ^[21]	77.00

5.2 AR 人脸数据库上的实验

AR 人脸数据库^[20]由 126 个人,超过 3 000 张的人脸图像组成。每个人的图像分两次拍摄,时间间隔为两个星期,每次拍摄的图像为 13 张。原始图像尺寸大小为 768×576 像素,每个像素用 256 级的灰度值来表示。图 6 显示了数据库中一个人的所有 26 张图像,其中图像标号 1 到 13 是第一次拍摄的,而标号 14 到 26 是第二次拍摄的。本实验从数据库中所有的 126 个人中选取 69 个人进行实验,并且对于每个人的所有图像中,仅选取编号为 1、2、3 和 5 的图像用于实验,分别对应于中性、微笑、生气以及光源为左的图像。在提取图像的几何特征和 Gabor 特征中,采用图 1(b)所定义的 22 个关键特征点作为实验采用的脸部关键点。

实验采用如下交叉验证策略进行:首先选取 1 幅图像作为测试数据,而剩余的图像作为训练数据。重复此过程直至所有的图像都被用作测试数据。计算正确分类的图像数目以获得平均识别率。实验获得了 98.55% 的人脸识别率和 81.16% 的面部表情识别率,其中单项式核函数的度设为 6。此外,为了说明本文所采用的脸部特征以及所提方法的有效性,还采用 PCA 方法^[12]、直接线性判别分析(D-

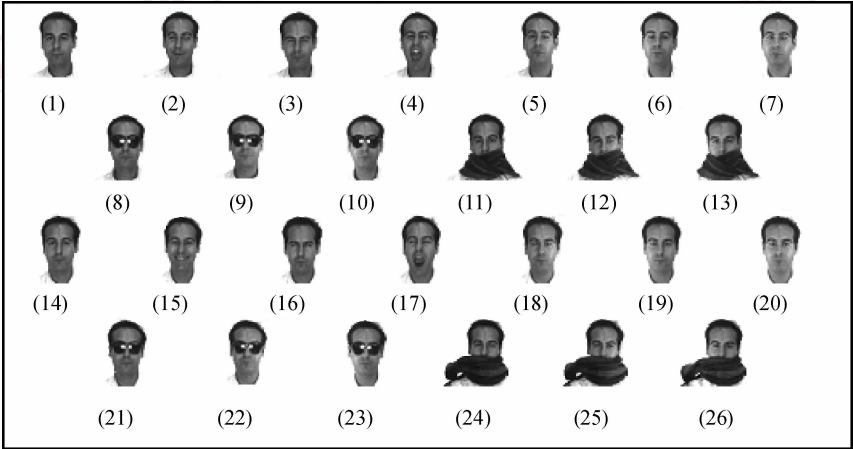


图 6 AR 人脸数据库示例
Fig. 6 Examples of the AR face database

LDA)方法^[22]、PCA + LDA 方法^[23]和零空间方法^[24]进行了人脸识别和表情识别的对比实验,其中每种方法所采用的输入数据为整幅人脸图像的灰度值。图 7 显示了各种方法的对比实验结果。

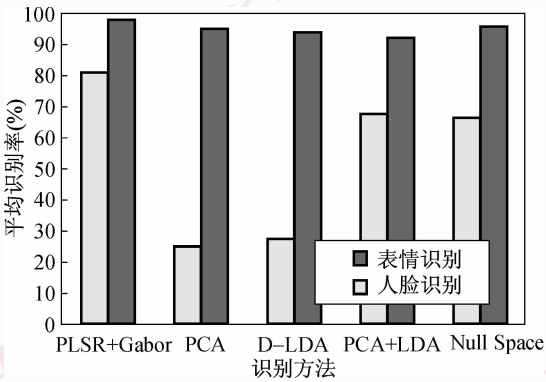


图 7 各种识别方法识别率的比较
Fig. 7 Comparison of the recognition rates among several recognition methods

从图 7 可以看出:无论是在人脸识别还是在表情识别方面,基于 PLSR 的方法均获得了最好的识别结果。此外也说明:采用图像灰度值作为输入数据并不能很好地进行面部表情识别,其原因在于:整个人脸灰度图像主要刻画人脸的全局信息,而难以刻画图像的局部信息。而运用 Gabor 特征和几何特征相混合的组合特征不仅可以刻画图像的全局信息,同时也刻画了图像的局部信息,因此在人脸身份识别和表情识别中均得到了较好的结果。

6 结 论

提出了一种新的基于 PLSR 的人脸身份和表情同步识别的方法。该方法充分利用 PLSR 技术的优点来建立面部图像特征和相应语义特征之间的关系式。基于两个著名的人脸图像数据库的实验表明:基于 PLSR 的识别方法不仅获得了较好的人脸识别效果,同时也获得了较好的表情识别效果。此外,实验结果也表明:同传统的灰度值特征相比,联合 Gabor 特征和几何特征在进行人脸身份和表情的统一识别方面具有较好的效果。虽然本文的重点在于人脸身份和表情的识别问题上,但所提出的识别方法也可用于其他多个目标同步识别问题(例如性别和身份的同步识别)。

参考文献 (References)

1 Fasel B, Luettin J. Automatic facial expression analysis: A survey [J]. Pattern Recognition, 2003, 36(1): 259-275.
2 Tolba A S, El-Baz A H, El-Harby A A. Face recognition: A literature review [J]. Journal of Signal Processing, 2005, 2(2): 88-103.
3 Ekman P, Friesen W V. Pictures of Facial Affect [M]. Palo Alto, CA, USA: Consulting Psychologists Press, 1976.
4 Lyons M L, Budynek J, Akamatsu S. Automatic classification of single facial images [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(12):1357-1362.
5 Tian Y L, Kanade T, Cohn J F. Facial expression analysis [A]. In: Handbook of Face Recognition [C], New York, USA: Springer Press, 2005:247-275.
6 Zhao W, Chellappa R, Phillips P J. Face recognition: A literature

- survey [J]. ACM Computing Surveys, 2003, **35**(4):399-458.
- 7 Wang Ya-min, Fu Xiao-lan. Recognizing facial expression and facial identity: parallel processing or interative processing [J]. Advances in Psychological Science, 2005, **13**(4):497-516. [汪亚珉, 傅小兰, 面部表情识别与面孔身份识别的独立加工与交换作用机制[J]. 心理科学进展, 2005, **13**(4): 497-516.]
- 8 Wold H. Estimation of principal components and related models by iterative least squares [A]. In: Multivariate Analysis [C], New York, USA: Academic Press, 1966:391-420.
- 9 Wiskott L, Fellous J. M, Krueger N, *et al.* Face recognition by elastic bunch graph matching [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, **19**(7):775-779.
- 10 Schölkopf B, Smola A, Müller K R. Nonlinear component analysis as a kernel eigenvalue problem [J]. Neural Computation, 1998, **10**(5): 1299-1319.
- 11 Baek J, Kim M. Face recognition using partial least squares components [J]. Pattern Recognition, 2004, **37**(6): 1303-1306.
- 12 Turk M, Pentland A. Eigenfaces for recognition [J]. Journal of Cognitive Neuroscience, 1991, **3**(1): 71-86.
- 13 Abdi H. Partial least squares (PLS) regression [A]. In: Encyclopedia for Research Methods for The Social Sciences [C], Thousand Oaks, CA, USA: Sage Publications, 2003:792-795.
- 14 Wold H. Soft Modeling by Latent Variables: The nonlinear iterative partial least squares approach [A]. In: Gani J, Ed., Perspectives in Probability and Statistics, Papers in Honour of M. S. Bartlett [C], London, UK: Academic Press, 1975: 520-540.
- 15 Rosipal R, Trejo L J. Kernel partial least squares regression in reproducing kernel hilbert space [J]. Journal of Machine Learning Research, 2001, **2**(6):97-123.
- 16 Zhang Z, Lyons M, Schuster M, *et al.* Comparison between geometry-based Gabor wavelets-based facial expression recognition using multi-layer perceptron [A]. In: Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition [C], Nara, Japan, 1998:454-459.
- 17 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.
- 18 Zheng W, Zhou X, Zou C, *at al.* Facial expression recognition using kernel canonical correlation analysis (KCCA) [J]. IEEE Transactions on Neural Networks, 2006, **17**(1):233-238.
- 19 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.
- 20 Martinez A M, Benavente R. The AR Face Database [R]. CVC Technical Report # 24, West Lafayette, Indiana, USA: Purdue University, 1998.
- 21 Feng X, Hadid A, Peitikäinen M. A coarse-to-fine classification scheme for facial expression recognition [A]. In: Campilho A, Kamel M, Eds. Proceedings of International Conference on Image Analysis and Recognition, Lecture Notes in Computer Science [C], Berlin, Germany: Springer Press 2004, **3212**: 668-675.
- 22 Yu H, Yang J. A direct LDA algorithm for high-dimensional data-with application to face recognition [J]. Pattern Recognition, 2001, **34**(10):2067-2070.
- 23 Belhumeur P N, Hespanha J P, Kriegman D J. Eigenfaces vs Fisherfaces: recognition using class specific linear projection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, **19**(7):711-720.
- 24 Chen L F, Mark Liao H Y, Ko M T, *et al.* A new LDA-based face recognition system which can solve the small sample size problem [J]. Pattern Recognition, 2000, **33**(10):1713-1726.