

中图法分类号: TP391 文献标志码: A 文章编号: 1006-8961(2011)03-0420-07

论文索引信息: 胡步发, 黄银成, 陈炳兴. 基于层次分析法语义知识的人脸表情识别新方法[J]. 中国图象图形学报, 2011, 16(3): 420-426

基于层次分析法语义知识的人脸表情识别新方法

胡步发, 黄银成, 陈炳兴

(福州大学机械工程及自动化学院, 福州 350108)

摘要: 在目前的人脸表情识别系统中, 人脸表情的机器识别和人类感知之间存在着本质的差异, 造成人脸表情识别率不高。为了减小人脸图像底层视觉特征与高层语义之间的语义鸿沟, 提出一种基于层次分析法(AHP)语义知识的人脸表情识别新方法。该方法首先采用层次分析法对训练集中人脸图像进行高层语义描述, 建立语义特征向量, 在底层视觉特征提取阶段, 提出一种二阶PCA(principal component analysis)方法来提取人脸图像的纹理特征; 在识别阶段, 仅利用输入人脸图像的底层视觉特征, 采用K-NN(k-nearest neighbor)算法并结合学习阶段建立的语义特征向量, 进行人脸表情分类识别。提出的人脸表情识别方法结合底层视觉特征和高层语义知识, 减小了人脸图像底层视觉特征与高层语义之间的语义鸿沟。在JAFPE人脸表情数据库中进行实验, 获得了93.92%的平均识别率。理论分析和实验结果表明, 与其他的人脸表情识别方法相比, 该方法具有更好的识别效果。

关键词: 人脸表情识别; 层次分析法; 底层视觉特征; 高层语义

A novel facial expression recognition method based on semantic knowledge of analytical hierarchy process

Hu Bufa, Huang Yincheng, Chen Bingxing

(College of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108 China)

Abstract: At present, there are intrinsic differences between machine recognition of facial expression and human perception in the facial expression recognition system, which affect the precision of facial expression recognition. In order to reduce the semantic gap between the low-level visual features of face images and high-level semantic, a novel facial expression recognition method based on semantic knowledge of analytical hierarchy process (AHP) is presented. The analytical hierarchy process method is adopted to describe the high-level semantic of face images of the training set, which further used to establish semantic features. In the stage of low-level visual features extraction, the 2nd-order principal component analysis method is proposed to extract the texture features of face images. In the recognition stage, only low-level visual features of the input face image is used, and k-nearest neighbor method combined with semantic features in the study stage is used to classify the facial expressions. The proposed method combines the low-level visual features with high-level semantic features, reducing the semantic gap between them. The experiments are conducted on Japanese Female Facial Expression (JAFPE) database and the overall recognition rate of 93.92% is achieved. Theoretical analysis and experimental results both show that the proposed method has higher recognition rate.

Keywords: facial expression recognition; analytical hierarchy process; low-level visual feature; high-level semantic

0 引言

人脸表情包含丰富的个人情感和精神状态方面

的信息, 在人机交互、机器视觉、人脸识别、行为科学和精神分析等方面有着广泛应用, 成为了研究热点^[1-4]。近年来, 研究者已经提出诸多用于人脸表情识别的方法, Abu 等人对人脸的眉毛、眼睛和嘴巴

收稿日期: 2009-08-10; 修回日期: 2009-11-19

第一作者简介: 胡步发(1963—), 男, 副教授。1991年获中国人民解放军国防科学技术大学通信与电子系统专业工学博士学位, 主要研究方向为计算机视觉、模式识别。E-mail: hubufa@21cn.com。

3 个表情特征区域进行特征点标定^[5],并通过计算 11 个不同的特征点与刚性模板特征点的位移,以此作为特征向量,利用 SVM(support vector machine)分类器,分别在 JAFFE(Japanese female facial expression)人脸表情库和 Cohn-Kanade 人脸表情库中,对 6 种基本人脸表情进行实验,分别获得 89.44% 和 84.86% 的平均识别率。Zhang 等人采用 MPEG-4 标准的人脸动画参数(FAPs)来提取人脸表情的底层视觉特征^[6],并在人脸表情识别中得到广泛运用。Aleksic 等人采用 FAPs 提取人脸表情特征,根据多重流隐马尔可夫模型实现表情分类,获得了 93.66% 的平均识别率^[7]。

从以上的研究分析可以看出,目前大部分的人脸表情识别方法都是依照这样一个过程:首先提取静态或序列人脸图像的底层视觉特征,然后采用这些特征作为分类器的输入,最后识别出表情类别。这些人脸表情识别方法在某种意义上获得了一定的成功,并给后来的研究者提供了良好的基础和有益启发,但是底层视觉特征和高层语义之间的语义鸿沟对于构造人脸表情识别系统仍然是一个巨大的挑战,表情识别准确率和算法鲁棒性还有待进一步提高和增强。此外,人脸表情识别也面临着一些不确定性,比如表情的细微变化复杂、表情的文化差异和性别差异、人脸的边缘遮挡等因素存在,增加了人脸表情识别的难度和不确定性。同时,由于底层视觉特征受图像光照、年龄等变化条件影响较大,因此在识别过程中引入语义知识对于提高人脸表情识别率是非常有必要的。Cheng 等人^[8]提出了采用层次分析法来构造人脸图像的语义知识,利用特征点标定方法在人脸脸上定位特征点,然后选择眉毛、眼睛和嘴巴区域的共 14 个特征点来表示各特征子区域的特征参数,并计算这些特征点的相对位置距离,但是该方法的训练集图像特征点手工标记工作费时而且繁琐,提取的底层特征精度不高,造成表情识别率不高。

针对这一问题,提出一种基于层次分析法语义知识的人脸表情识别新方法。通常情况下,底层视觉特征忽略了人脸图像的先验感知内容,因此仅采用人脸图像的底层视觉特征来表征人脸表情是不充分的。为了减小底层视觉特征与高层语义之间的语义鸿沟,采用层次分析法对训练集中人脸图像进行高层语义描述,建立语义特征向量,在底层视觉特征提取阶段,提出一种二阶 PCA 方法来提取人脸图像

的纹理特征。在识别阶段,仅利用输入人脸图像的底层视觉特征,采用 K-NN 算法并结合学习阶段建立的语义特征向量,进行表情分类识别。实验结果表明,与其他的人脸表情识别方法相比,提出的方法具有更好识别效果。

1 基于语义知识的人脸表情识别方法

提出的人脸表情识别方法分为两个阶段。在学习训练阶段,通过层次分析法得到训练样本的语义特征向量,并采用二阶 PCA 方法来提取人脸图像的纹理特征。最后采用 K-NN 算法并结合学习阶段建立的语义特征向量,进行表情分类识别。流程图如图 1 所示。

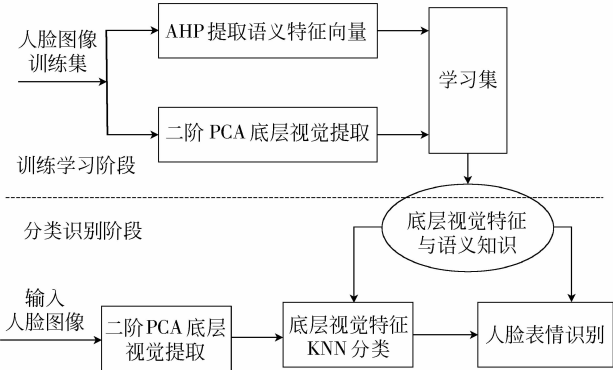


图 1 本文方法的流程图
Fig. 1 The flow chart of the proposed method

1.1 基于层次分析法的人脸表情语义特征提取

层次分析法是美国著名运筹学家 Satty 在 20 世纪 70 年代提出的,是一种定性与定量相结合的方法。层次分析法主要适用于多准则、多目标的复杂问题的决策分析。人脸表情识别问题也是属于多准则问题,如图 2 的两幅人脸图像都可以归类为高兴



图 2 底层视觉特征差别较大的两幅高兴表情图像
Fig. 2 Two face images of “happiness” emotion with different low-level visual features

表情,但是它们的外在表现(如嘴唇)存在较大的差异,因此在人脸表情识别方法中通过层次分析法引入语义知识就显得非常必要。

层次分析法的一个优点是可以通过专家组的群组决定,每个专家可以通过自己来判断层次分析结构中的语义描述指标相对重要性,然后取所有专家的相对重要性的均值作为最后的判断数据,可以有效避免人为主观误差。

1.1.1 人脸表情的高层语义层次结构模型

人脸表情的高层语义层次结构模型分为 3 个层次,如图 3 所示。在第 1 层次包含七种表情类别,然后每种表情类别又在第 2 层次分为 3 个子类别,并在第 3 层次根据 MPEG-4 的人脸动画参数 FAPs 对上一级的子类别分别进行特征描述^[6]。从第 3 层次经过第 2 层次,到第 1 层次的路径就形成了一条语义特征描述,而多个语义特征描述就可以表征一种表情类别。很显然,要描述不同的表情类别,对于各个节点路径的语义特征描述的权重是不一样的。

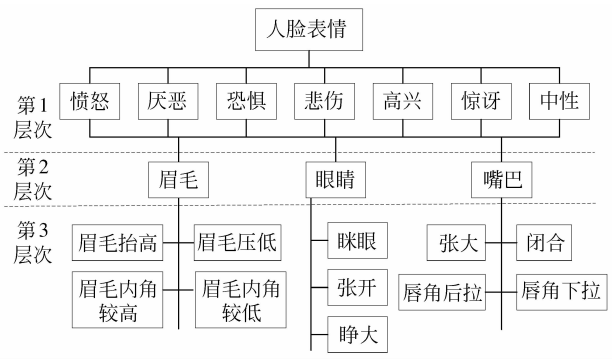


图 3 人脸表情的高层语义层次结构模型

Fig. 3 The high-level semantic hierarchy model of the facial expression

1.1.2 人脸表情的语义特征向量

假设语义层次结构的路径节点编码为 1 ~ N, 给定一幅人脸图像 q , 则该图像的语义知识可以用一组语义特征向量 I 来表示为^[8]

$$I = (s_1, s_2, \dots, s_N), \sum_{p=1}^N s_p = 1 \quad (1)$$

式中, s_p 表示第 p 个路径节点编码的权重。

显然,给每个路径节点编码分配权重是一个费时复杂的过程,而且难以实现。本文的语义特征向量的权重分配则采用层次分析法,层次分析法的优点是仅需要用户对若干对的语义描述指标的相对重要性进行判断,然后根据算法就可以自动计算得到

权重,而权重的顺序组合就构成人脸表情的语义特征向量。

在构建层次结构模型后,上下层次元素间的隶属关系就被确定了。下一步就是要确定各层次元素的权重。层次分析法利用决策者给出判断矩阵的方法导出权重。设某一层元素 Ω 所支配的下一层次元素为 $U_1, U_2, \dots, U_n$ 。针对该层元素 Ω 比较两个元素 U_i 和 U_j 哪个更重要,重要程度如何,并按按照表 1 定义的比例标度对重要性程度赋值,形成判断矩阵 $A = (a_{ij})_{n \times n}$,其中 a_{ij} 为元素 U_i 和 U_j 相对重要性比例标度。

表 1 1~9 标度的定义

Tab. 1 The definition of scale 1 ~ 9

比例标度	定义
1	2 个元素相比,具有相同的重要性
3	2 个元素相比,前者比后者稍重要
5	2 个元素相比,前者比后者明显重要
7	2 个元素相比,前者比后者强烈重要
9	2 个元素相比,前者比后者极端重要
2,4,6,8	表示上述相邻判断的中间值

例如图 4(a)为一幅人脸图像,这幅图像在层次结构模型中的第 1 层次可分为 3 个表情类别:高兴(H)、惊讶(S)和中性(N)。图 4(b)为对应的判断矩阵 A_1 ,可以这样计算得来

$$A_1 = \begin{bmatrix} \frac{w_H}{w_H} & \frac{w_H}{w_S} & \frac{w_H}{w_N} \\ \frac{w_S}{w_H} & \frac{w_S}{w_S} & \frac{w_S}{w_N} \\ \frac{w_N}{w_H} & \frac{w_N}{w_S} & \frac{w_N}{w_N} \end{bmatrix} \quad (2)$$

式中, w_H 、 w_S 、 w_N 分别是对应于高兴(H)、惊讶(S)和中性(N)3 种表情类别的相对重要性比例标度,根据表 1 的定义,分别对 3 种表情类别的相对重要性比例标度进行赋值(即与各种表情类别的相似性进行赋值),得到 $w_H = 8$ 、 $w_S = 4$ 和 $w_N = 1$,对于 $w_H = 8$ 的赋值表示图 4(a)的人脸图像与高兴表情类别的相对重要性位于强烈重要与极端重要之间(即相似度非常高),同理, $w_S = 4$ 的赋值表示图 4(a)的人脸图像与惊讶表情类别的相对重要性位于稍重要与明显重要之间(即相似度较高)、 $w_N = 1$ 的赋值表示图 4(a)的人脸图像与中性表情类别的相对重要性同等重要(即相似度低)。而对于另外 4 种表情的语义描述在语义特征向量 I 可以用 0 表示,则第 1 层次

的语义描述就可由 A_1 得到,因此构建的语义特征向量能够很好的表征人脸图像的先验感知语义知识。



(a) 人脸图像

	高兴	惊讶	中性
高兴	1	4	8
惊讶	1/4	1	2
中性	1/8	1/2	1

(b) 对应的第1层次语义描述判断矩阵

图4 人脸图像范例

Fig. 4 One face image example

不失一般性,令 l, m 和 n 分别是对应于第1、第2和第3层次的语义描述元素数目,式(3)为判断矩阵 A_1 ,其中 $w_i^{(1)}$ 为第1层次第 i 个语义描述的相对重要性比例标度。

$$A_1 = \begin{bmatrix} \frac{w_1^{(1)}}{w_1^{(1)}} & \frac{w_1^{(1)}}{w_2^{(1)}} & \cdots & \frac{w_1^{(1)}}{w_l^{(1)}} \\ \frac{w_2^{(1)}}{w_1^{(1)}} & \frac{w_2^{(1)}}{w_2^{(1)}} & \cdots & \frac{w_2^{(1)}}{w_l^{(1)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_l^{(1)}}{w_1^{(1)}} & \frac{w_l^{(1)}}{w_2^{(1)}} & \cdots & \frac{w_l^{(1)}}{w_l^{(1)}} \end{bmatrix} \quad (3)$$

对于第1层次的判断矩阵 A_1 的每行,计算行权重得到行权重矩阵 A'_1 ,式(4)为行权重矩阵

$$A'_1 = \begin{bmatrix} v_1^{(1)} \\ v_2^{(1)} \\ \vdots \\ v_l^{(1)} \end{bmatrix} \quad (4)$$

行权重 $v_i^{(1)}$ 的计算公式如下式所示

$$v_i^{(1)} = \sqrt[l]{\frac{w_i^{(1)}}{w_1^{(1)}} \times \frac{w_i^{(1)}}{w_2^{(1)}} \times \cdots \times \frac{w_i^{(1)}}{w_l^{(1)}}} \quad i = 1, 2, \dots, l \quad (5)$$

然后对第1层次的行权重 $v_i^{(1)}$ 进行归一化,得到相对权重为

$$u_i^{(1)} = v_i^{(1)} / \sum_{j=1}^l v_j^{(1)} \quad i = 1, 2, \dots, l, \quad (6)$$

同样可以计算得到,第2层次的相对权重 $u_{i,j}^{(2)}$, $j = 1, 2, \dots, m$ 和第3层次的相对权重 $u_{i,j,k}^{(3)}$, $k = 1, 2, \dots, n$,最后可以得到一幅图像的语义特征向量 $I = (s_1, s_2, \dots, s_N)$, N 为层次结构的路径节点编码总数,其中 s_p 为

$$s_{p(i,j,k)} = u_i^{(1)} \times u_{i,j}^{(2)} \times u_{i,j,k}^{(3)} \quad (7)$$

1.2 基于二阶 PCA 的人脸表情底层视觉特征提取

PCA 的目的是从一组零均值的随机变量中找出一个线性无关的组合,使得只需要通过较少的变量就可以描述原始变量尽可能多的信息。

1.2.1 PCA 基本理论

设 N 维人脸图像样本集 $X = \{x_1, x_2, \dots, x_M\}$, M 为样本数。根据 PCA 理论, X 经变换矩阵 U 后降为 N_1 维特征向量集 $Y = \{y_1, y_2, \dots, y_M\}$, 其中:

$$y_i = U^T(x_i - m) \quad i = 1, 2, \dots, M \quad (8)$$

$$m = \frac{1}{M} \sum_{i=1}^M x_i \quad (9)$$

$$U = (u_1, u_2, \dots, u_{N_1}) \quad N_1 \leq N \quad (10)$$

式中,向量 u_j 表示样本协方差矩阵 S_i 的第 j 大特征值 λ_j 所对应的特征向量。

$$S_i = \frac{1}{M} \sum_{i=1}^M (x_i - m)(x_i - m)^T \quad (11)$$

S_i, u_j, λ_j 满足

$$S_i u_j = \lambda_j u_j \quad (12)$$

式中, m 为 X 的平均向量。 Y 为方差最大的 N_1 个分量,且 $x_j = Uy_j + m$,在所有正交重构中,具有最小的均方误差。

1.2.2 二阶 PCA 特征提取方法

在一些情况下,如在光照变化情况,PCA 方法并不能很有效地表达人脸图像,针对这一问题, Kima 等人提出采用二阶 PCA 方法^[9],该方法不仅使用了原始图像的 PCA 特征,还使用了残余图像特征脸,能够更有效地表达出人脸图像特征。因此,提出一种基于二阶 PCA 的人脸底层视觉特征提取方法。

二阶 PCA 就是提取残余图像 $X' = \{x'_1, x'_2, \dots, x'_M\}$ 的 PCA 特征,其中, $x'_i = x - \hat{x}_i$, $x_i, \hat{x}_i$ 分别是原始人脸图像和重构后的人脸图像, $\hat{x}_i = Uy_i + m$ 。 N 维的残余人脸图像 $X' = \{x'_1, x'_2, \dots, x'_M\}$ 通过转换矩阵 V 降为 N_2 维的残余人脸图像,即

$$y'_i = V^T(x'_i - m') \quad (13)$$

式中, m' 是 X' 的平均向量, $m' = \frac{1}{M} \sum_{i=1}^M x'_i$, $V = \{v_1, v_2, \dots, v_{N_2}\}$ 。向量 v_j 表示样本协方差矩阵 S'_i 的第 j 大特征值 λ'_j 所对应的特征向量。

同时考虑变换矩阵 U 和 V ,则人脸图像样本 x_i 可以通过二阶 PCA 重构人脸图像 $\tilde{x}_i$ 近似表示

$$x_i \cong \tilde{x}_i = Uy_i + m + Vy'_i \quad (14)$$

式中,

$$y_i = U^T(x_i - m) \quad (15)$$

$$y'_i = V^T(x_i - Uy_i - m) \quad (16)$$

通过人脸图像 PCA 特征和残余人脸图像 PCA 特征获得人脸图像的重构过程,如图 5 所示。

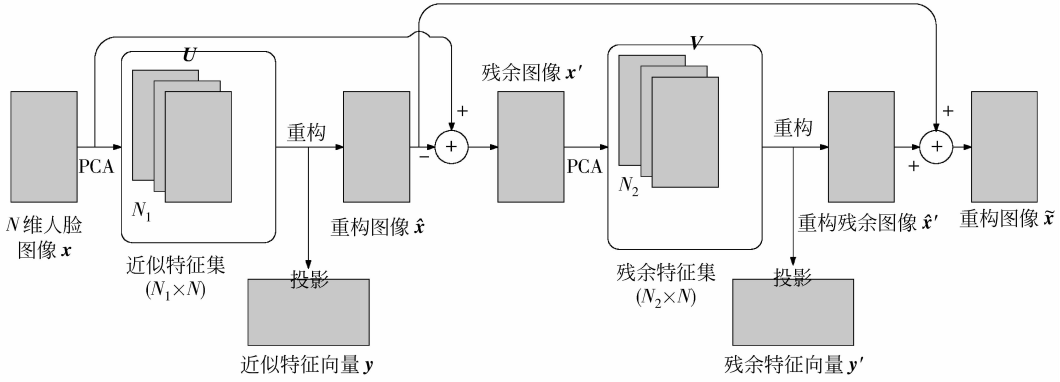


图 5 二阶 PCA 人脸图像重构流程图

Fig. 5 The flow chart of the face image reconstruction using the 2nd-order PCA

1.3 基于底层视觉特征的 K-NN 算法

训练集中的人脸图像经过语义特征提取和底层视觉特征提取后组成学习集,如果简单地只将一张最佳匹配人脸图像作为最终的匹配结果,则容易导致单个人脸图像的偶然性所引起的误判。为了提高分类识别的可靠性,减少偶然性误差引起的误判,对最近邻法(NN)进行改进,提出 K-NN 算法来查找人脸图像。

传统上的 NN 方法的判决规则是只查找底层视觉特征距离最小的人脸图像,这样容易引起误判,本文的 K-NN 算法用来查找学习集中的 k 个底层视觉特征距离最小的人脸图像。而要想找出 k 个最近邻匹配人脸图像,必须要有一个衡量标准,根据该标准找出最近邻匹配。采用底层视觉特征欧式距离作为最佳匹配度量,度量函数定义如下式所示

$$D(q, c) = \sum_{j=1}^m (q_j - c_j)^2 \quad (17)$$

式中, q_j 为被测人脸图像 q 底层视觉特征对应的第 j 个单个特征, c_j 为学习集人脸图像 c 底层视觉特征对应的第 j 个单个特征, m 为特征数。

被测人脸图像与训练集中所有人脸图像进行匹配,通过式(17)来度量匹配结果,然后按度量距离的大小对人脸图像进行排列,选择距离最小的 k 个最近邻样本作为匹配结果。

2 人脸表情分类识别

查找出训练集中的 k 个底层视觉特征距离最小

的人脸图像后,然后根据这 k 个最近邻人脸图像的语义特征向量,判断输入人脸图像的表情类别。

根据图 3 的人脸图像表情高层语义层次结构模型,一幅人脸图像的语义特征向量包括 7 个(愤怒、厌恶、恐惧、悲伤、高兴、惊讶和中性)子向量组成,每个子向量 11 维,可得到一幅人脸图像的表情语义特征向量 $I = (s_1, s_2, \dots, s_{77})$ 。根据其语义特征向量 I ,可得到一幅人脸图像隶属于各种表情类别的概率:

$$\begin{aligned} P_{AN} &= \sum_{i=1}^{11} s_i, P_{DI} = \sum_{i=12}^{22} s_i, P_{FE} = \sum_{i=23}^{33} s_i, P_{SA} = \sum_{i=34}^{44} s_i \\ P_{HA} &= \sum_{i=45}^{55} s_i, P_{SU} = \sum_{i=56}^{66} s_i, P_{NE} = \sum_{i=67}^{77} s_i \end{aligned} \quad (18)$$

然后取所有表情类别概率的最大值所属的类别判为输入人脸图像的表情类别,即

$$P_j = \max_{1 \leq i \leq 7} P_i$$

(1, 2, ..., 7 分别为愤怒(AN)、厌恶(DI)、恐惧(FE)、悲伤(SA)、高兴(HA)、惊讶(SU)、中性(NE))则将该输入人脸图像表情判为第 j 类表情。

本文方法可以归纳如下:

1) 根据二阶 PCA 方法提取输入人脸图像 q 的底层视觉特征 F_q 。

2) 利用得到的底层视觉特征 F_q , 根据 K-NN 算法来查找训练集中的 k 个底层视觉特征距离最小的人脸图像,并组成候选集 C 。

3) 在候选集 C 中,计算组合语义特征向量

$$\bar{I}_q = \sum_{j=1}^{|C|} a_j I_j \quad (19)$$

式中, I_j 为候选集 C 中第 j 幅图人脸图像 Q_j 的语义特征向量, a_j 为相对应于 I_j 的权重, a_j 的值可以通过下式计算得到

$$a_j = \left(1 - \frac{D(q, Q_j)}{D_{\max}}\right) / \sum_{i=1, 2, \dots, |C|} \left(1 - \frac{D(q, Q_j)}{D_{\max}}\right)$$

(20)

式中, $D_{\max} = \max [D(q, Q_j), j = 1, 2, \dots, |C|]^{[8]}$ 。

4) 根据组合语义特征向量 I_q , 采用式(18) 计算所有表情类别的概率值。

5) 将所有表情类别概率的最大值所属的类别判为输入人脸图像的表情类别。

显然, 本文方法的识别率受 K-NN 算法中近邻数目 k 的影响, 根据实验结果分析, 取近邻数目为 $k=4$ 的时候就能够取得较高的表情识别率和较好的系统鲁棒性。

3 人脸表情识别实验及分析

为了验证提出方法的有效性, 在 JAFFE 人脸表情数据库中 进行实验。JAFFE 人脸表情库由 10 个人, 共 213 幅图像组成, 每人都包含愤怒、厌恶、恐惧、高兴、悲伤、惊讶和中性 7 种表情, 每人每种表情 2~4 幅, 每种表情共有 27 幅。实验的测试集和训练集采用不同的人脸图像集, 对每种表情的 27 幅人脸表情图像, 其中 9 幅图像进行训练, 其余 18 幅图像用来识别, 交叉实验 3 次。

由于本文方法的识别率受 K-NN 算法中近邻数目 k 的影响, 因此实验首先对不同的近邻数目 k 对于识别率的影响进行实验比较, 实验结果如表 2 所示。由表 2 可知, 取近邻数目 $k=4$ 的时候就能够取得较高的表情识别率, 而取近邻数目 $k=5$ 的时候表情识别率急剧下降, 这是由于在 $k=5$ 时增加了大量的误查找的人脸图像, 使得人脸表情的正确识别率急剧减少。

表 2 对应于不同的近邻数目 k 的实验结果

Tab.2 The experimental results corresponding to different nearest neighbor number k

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
识别率/%	90.27	92.61	93.92	89.27	86.39

本文方法采用层次分析法构建人脸表情的语义特征向量, 在提高人脸表情识别率方面有着重要的作用, 此实验分为两类(近邻数目 $k=4$), 第 1 类实

验为直接采用 K-NN 算法对人脸表情进行分类, 而不利用人脸表情的语义知识, 表情识别率如表 3 所示; 第 2 类实验为本文提出的人脸表情识别方法, 表情识别率如表 4 所示。从表 3 和表 4 可以看出, 采用本文方法后, 表情识别率从 88.36% 提高到 93.92%; 同时还可以观察到惊讶表情的识别率最高, 主要原因是惊讶表情变化幅度较大, 识别率高。

表 3 第 1 类实验结果 ($k=4$)

Tab.3 The first experimental results ($k=4$)

	愤怒	厌恶	恐惧	悲伤	高兴	惊讶	中性
愤怒	47	2	4	0	0	1	2
厌恶	3	44	3	2	0	0	2
恐惧	0	1	42	1	0	0	0
悲伤	2	5	0	51	0	0	0
高兴	0	0	0	0	49	2	0
惊讶	2	0	3	0	2	51	0
中性	0	2	2	0	3	0	50
识别率/%	87.04	81.48	77.78	94.44	90.74	94.44	92.59

表 4 第 2 类实验结果 ($k=4$)

Tab.4 The second experimental results ($k=4$)

	愤怒	厌恶	恐惧	悲伤	高兴	惊讶	中性
愤怒	49	2	1	0	0	0	0
厌恶	1	47	2	1	0	0	0
恐惧	0	0	48	0	0	0	1
悲伤	2	3	0	52	0	0	0
高兴	0	0	0	1	52	0	0
惊讶	1	0	2	0	2	54	0
中性	1	2	1	0	0	0	53
识别率/%	90.74	87.04	88.89	96.30	96.30	100	98.15

另外, 还将本文实验结果与其他表情识别方法的实验结果进行比较, 表 5 列出了一些采用不同的表情识别方法和表情库所得到的表情识别率。表 6 列出了本文方法和文献[10]中不同的分辨率(像素×像素)下得到的平均表情识别率。

表 5 表情识别率比较结果 1

Tab.5 The contrast results one of expression recognition rate

文献	方法	表情库	平均识别率/%
文献[5]	AUs + SVM	JAFFE	89.44
文献[5]	AUs + SVM	Cohn-Kanade	84.86
文献[7]	FAPs + HMMs	Cohn-Kanade	93.66
文献[8]	FAPs + AHP	JAFFE	85.2
本文	本文方法	JAFFE	93.92

由表5可知, 与引言中总结的4种表情识别方

表 6 表情识别率比较结果 2

Tab.6 The contrast results two of expression recognition rate

文献	方法	分辨率	平均识别率/%
文献[10]	ALBP + Tsallis + NLDAI	16 × 16	84.62
文献[10]	ALBP + Tsallis + NLDAI	32 × 32	88.82
文献[10]	ALBP + Tsallis + NLDAI	48 × 48	90.54
文献[10]	ALBP + Tsallis + NLDAI	64 × 64	94.59
本文	本文方法	67 × 55	93.92

法的平均识别率(89.44%、84.86%、93.66%和85.2%)相比,本文提出的人脸表情识别方法的平均识别率更高。这主要是得益于二阶PCA方法具有良好的人脸图像表征性能,以及采用层次分析法构造人脸表情的语义特征向量,结合了底层视觉特征和高层语义知识,减少了人脸图像底层视觉特征与高层语义的语义鸿沟,获得较高的表情识别率和较好的系统鲁棒性。

由表6可知,在文献[10]中,表情识别率随着分辨率的不同而改变,且呈现出分辨率越大,平均识别率越高的趋势,文献[10]中在分辨率(像素×像素)为64×64情况下得到的平均识别率达到94.59%,要比本文中使用的分辨率(像素×像素)为67×55下得到的识别率要高,这是由于图像分辨率越大,人脸图像越清晰,细节特征就越丰富,表情识别率也相应的提高了;同时,文献[10]中组合了多种特征提取方法,ALBP、Tsallis、NLDAI分别用来提取人脸图像的灰度梯度特征、低中频的纹理特征、全局外观特征,因此得到较高的表情识别率。

4 结 论

提出了一种基于层次分析法语义知识的人脸表情识别新方法,该方法有以下两个特点:1)采用层次分析法建立语义特征向量,层次分析法适用于多准则、多目标的复杂问题的决策分析;2)采用二阶PCA方法来提取人脸图像底层视觉特征,二阶PCA方法具有良好的人脸图像表征性能。因此,该方法能够很好的对底层视觉特征和高层语义知识进行了结合,减少了人脸图像底层视觉特征与高层语义的语义鸿沟,取得了较高的人脸表情识别率。实验结

果表明,该方法在表情识别率和系统鲁棒性方面都具有优势。

参考文献 (References)

[1] Kim M, Lee H S, Jeong W P, et al. Determining color and blinking to support facial expression of a robot for conveying emotional intensity[C]// Proceedings of the 17th International Symposium on Robot and Human Interactive Communication. New Jersey: IEEE Computer Society, 2008: 219-224.

[2] Lee K K, Xu Y. Real-time estimation of facial expression intensity[C]// Proceedings of the 2003 IEEE International Conference on Robotics and Automation. New York: IEEE Inc. , 2003: 2567-2572.

[3] Shishir Bashyal, Ganesh K Venayagamoorthy. Recognition of facial expressions using Gabor wavelets and learning vector quantization [J]. Engineering Applications of Artificial Intelligence, 2008,21 (7): 1056-1064.

[4] Geetha A, Ramalingam V, Palanivel S, et al. Facial expression recognition - a real time approach[J]. Expert Systems with Applications, 2009,36 (1): 303-308.

[5] Abu Sayeed Md Sohail, Prabir Bhattacharya. Classifying facial expressions using point-based analytic face model and Support Vector Machines [C]//Proceedings of the 2007 IEEE International Conference on Systems Man and Cybernetics. New York: IEEE Inc. , 2007:1008-1013.

[6] Zhang Yongmian, Ji Qiang, Zhu Zhiwei. Facial expression analysis and synthesis with MPEG-4 facial animation parameters[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008,18(10):1383-1395.

[7] Aleksic P S, Katsaggelos A K. Automatic facial expression recognition using facial animation parameters and multistream HMMs[J]. IEEE Transactions on Information Forensics and Security, 2006,1(1):3-11.

[8] Cheng Shyichyi, Chen Mingyao, Chou T C, et al. Semantic-based facial expression recognition using analytical hierarchy process[J]. Expert Systems with Applications, 2007,33 (1): 86-95.

[9] Kima Hyunchul, Daijin Kima, Sung Yang Banga, et al. Face recognition using the second-order mixture-of-eigenfaces method[J]. Pattern Recognition, 2004,37(2):337-349.

[10] Shu Liao, Wei Fan, Albert C S, et al. Facial expression recognition based on local binary patterns and coarse-to-fine classification[C]// Proceedings of the 2006 IEEE International Conference on Image Processing. New Jersey: IEEE Computer Society, 2006:665-668.