

基于特征点矢量与纹理形变能量参数融合的人脸表情识别

易积政^① 毛 峡^{*①} Ishizuka Mitsuru^② 薛雨丽^①

^①(北京航空航天大学电子信息工程学院 北京 100191)

^②(Department of Information and Communication Engineering, University of Tokyo, Japan)

摘 要: 人脸表情识别是人机交互领域的研究热点和难点之一。为了有效去除由于个体差异而造成的表情特征差异,该文首先基于特征点矢量提出特征点距离比例系数;其后,又给出纹理形变能量参数的概念;最后,将二者融合用于人脸表情识别。所提方法在 Cohn-Kanade 数据库及 BHU 人脸表情数据库进行了测试,实验结果表明该方法较传统的方法在识别率上分别提高了 4.5%与 3.9%。

关键词: 人脸表情识别;特征点矢量;纹理形变能量参数;径向基神经网络

中图分类号: TP391.41

文献标识码: A

文章编号: 1009-5896(2013)10-2403-08

DOI: 10.3724/SP.J.1146.2012.01569

Facial Expression Recognition Based on Feature Point Vector and Texture Deformation Energy Parameters

Yi Ji-zheng^① Mao Xia^① Ishizuka Mitsuru^② Xue Yu-li^①

^①(School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

^②(Department of Information and Communication Engineering, University of Tokyo, Japan)

Abstract: Facial expression recognition is a popular and difficult research field in human-computer interaction. In order to remove effectively the differences in expression feature caused by individual differences, this paper firstly presents the feature point distance ratio coefficient based on feature point vector, and then gives the concept of texture deformation energy parameters. Finally, merges previously mentioned two parts to form a new expression feature for facial expression recognition. The proposed method is tested in the Cohn-Kanade database and the BHU facial expression database, and the experimental results show the recognition rates of the proposed method comparing with the existing ones increased by 4.5% and 3.9%.

Key words: Facial expression recognition; Feature point vector; Texture deformation energy parameters; Radial basis function neural network

1 引言

情感识别一直是人机交互领域中重要的研究方向之一。为了建立友好和谐的人机交互模式,大量研究者从语音、面部表情、文本等方面入手,期望以单模或者多模融合的方法实现更好的人机交互效果。面部情感识别是情感识别的重要研究方向,近些年涌现出大量优秀的研究成果^[1-11],有力地推动了人机交互研究的发展。文献[12,13]提出了高兴、悲伤、惊讶、生气、嫌恶和害怕 6 种基本表情,并提出了面部动作编码理论。然而,真正意义上依靠计算机的自动化人脸表情识别始于上世纪 90 年代,许多基于文献[12,13]所提理论的研究成果相继出现。当然,也有一些研究者^[14,15]提出 6 种表情以外的表情模式,极大地丰富了人脸情感识别研究的内

容。随着计算机技术的飞速发展,研究者们所提出的研究方向大致分为两个主要分支:一是基于整体外观的方法,二是基于几何特征的方法。

在基于整体外观的人脸表情识别研究领域,局部纹理信息对捕获人脸表情信息非常重要。要得到满足条件的纹理特征,首先,需要利用形状特征、局部纹理、光流场等信息来获得表情的原始特征。然而,这些原始特征都存在信息冗余和维数过高的问题,甚至夹杂了干扰数据,为后续分类研究带来困难。所以,需要对原始特征数据进行处理,如特征提取和分解、特征降维等,以降低维数,去除干扰数据,得到对分类更为有利的特征数据。常用的特征提取及降维的方法有主成分分析(PCA)、线性判别分析(LDA)、独立成分分析(ICA)等。PCA 提取了最具本质的特征,可以有效地消除干扰因素,同时降低维数,但未考虑不同类别数据之间的区分性。而 LDA 则通过最大化数据的类间离散度和最小化类内离散度来选择合适的投影方向,侧重于寻找

2012-12-04 收到, 2013-02-06 改回

国家科技部中日国际科技合作项目(2010DFA11990)和国家自然科学基金(61103097)资助课题

*通信作者: 毛峡 moukyoun@yahoo.com.cn

具有最大分辨力的方向。PCA 与 LDA 仅在线性流形中具有良好的提取人脸表情信息的作用。人脸图像在数据结构上通常都呈现非线性流形的特征。因此,关于数据降维的研究又深入到非线性流形降维领域,出现 ISOMetric MAPping(ISOMAP), Locally Linear Embedding(LLE), Stochastic Neighbor Embedding(SNE)等算法^[16-18]。

在基于几何特征的人脸表情识别方法能够得到比较满意的识别效果。Kotsia 等人^[19]运用形状和纹理信息进行人脸表情识别研究,取得了很好的效果。伊朗的 Mahdi Ilbeygi 等人^[20]提出了基于彩色图像的模糊人脸表情识别系统,并详细给出了脸部器官的特征提取与后续的模糊分类方法,达到了很好的识别效果。Shi 等人^[21]验证了与人脸几何信息相关的标记点在人脸识别与人脸表情识别中的效果,并提出标记点的几何坐标不能直接用来训练分类器,因为不同的人可能有不同的面貌特征。因此,相对距离成了衡量几何距离的主要手段。文献[22]利用主动外观模型提取人脸几何特征,进而进行人脸表情识别。Valstar 和 Pantic^[23]提出人脸几何特征提取与运动单元识别,实现了人脸表情预测。然而,对于一些不同的面部表情,其几何特征有可能是相似的,从而使 Valstar 和 Pantic 的研究变得困难。Song Ming-li 等人^[24]利用主动形状模型对经过图像处理单元(Graphic Processing Unit, GPU)预处理后的人脸图像进行特征点检测,并提出了将图像比例特征与面部动画参数相结合作为支持向量机的输入,最终实现人脸情感识别,得到了较好的识别效果。上述基于几何特征的人脸情感识别研究都存在同一个问题,即未考虑个体差异对人脸情感识别结果的影响。

本文综合利用人脸图像的结构特征与纹理信息,并将得到的人脸特征通过比例系数的形式标准化,以此减小由于个体差异造成的人脸表情特征的差异。实验结果表明,本文所提方法对人脸表情识别是鲁棒的。

2 表情识别系统结构

图 1 为本文提出的人脸表情识别系统整体框图。系统的基本思想是融合人脸图像的结构特征与纹理信息,然后利用径向基函数(RBF)神经网络进行人脸表情识别。用于人脸表情识别研究的特征参数主要分为两部分:特征点矢量和含有纹理信息的特征块。

在获取特征点矢量与特征块之前,本文利用主动外观模型(Active Appearance Model, AAM)^[25]对特征点进行定位。经典的人脸特征点定位算法有主动形状模型(Active Shape Model, ASM)^[26]与 AAM。AAM 与 ASM 的不同之处在于:AAM 不仅利用了人脸形状信息,而且利用了人脸纹理信息,将形状与纹理结合起来建立模型,对于人脸各个器官的特征点定位更加准确。AAM 的形状模型是由一组人脸图像上的三角网格和三角网格的顶点集 \mathbf{S} 构成,通常 \mathbf{S} 记为坐标点集合,如式(1)所示,式中 v 为坐标点个数。

$$\mathbf{S} = [x_1, y_1, x_2, y_2, \dots, x_v, y_v]^T \quad (1)$$

通过训练给定的顶点集得到 AAM 形状模型,对顶点集数据进行对齐、统计估计、PCA 等处理,最终得到形状模型,如式(2)所示。

$$\mathbf{S} = \mathbf{S}_0 + \sum_{i=1}^n \rho_i \mathbf{S}_i \quad (2)$$

式中 \mathbf{S}_0 为平均形状; $\{\mathbf{S}_i\} (i=1, 2, \dots, n)$ 为正交基底,对应于 n 个最大特征值的特征向量; $\{\rho_i\}$ 为一组形状参数。

AAM 纹理模型是建立在 \mathbf{S}_0 上的。同样对纹理数据进行对齐、统计估计、PCA 等处理,便可获得纹理模型 \mathbf{T} , 由式(3)表示为

$$\mathbf{T} = \mathbf{T}_0 + \sum_{i=1}^m \lambda_i \mathbf{T}_i \quad (3)$$

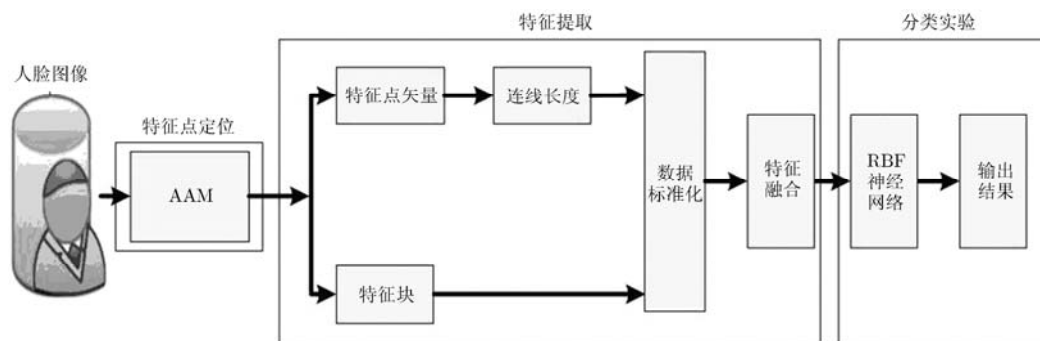


图 1 人脸表情识别系统框图

式中 T_0 为平均纹理； $\{T_i\}(i = 1, 2, \dots, m)$ 为正交基底，对应于 m 个最大特征值的特征向量； $\{\lambda_i\}$ 为一组纹理参数。

3 基于特征点矢量的特征提取

众所周知，个体之间存在很大差异，在脸型和器官形状上表现尤为突出。面部各器官的不同导致个体在表达同种感情时都具有很大的差异。如图 2 所示，人在做惊讶的表情时都趋向于睁大眼睛，但是，眼小的人，其上眼皮与下眼皮的距离，即睁眼的幅度，可能与眼大的人在平静状态下一样。在基于结构特征的人脸情感识别研究中，应考虑个体差异对识别结果的影响。



图 2 Cohn-Kanade 数据库惊讶表情实例

本文利用 AAM 算法对人脸共计 106 个特征点进行定位，最后选取的 26 个特征点如图 3 所示，其中“+”表示各特征点所在位置，箭头后的数字代表特征点对应的编号。特征点的对应说明如表 1 所示。

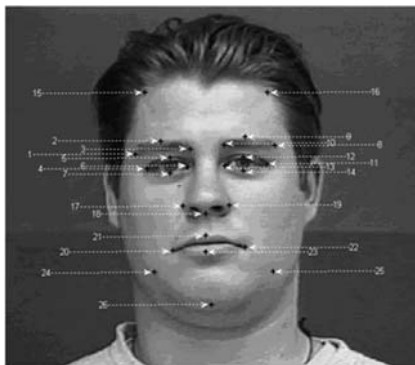


图 3 特征点位置示意图

将选取的特征点记为 $p_i (i = 1, 2, 3, \dots, 26)$ ，定义面部各特征点两两构成的矢量为特征点矢量。 p_{15} 相对于各点的特征点矢量如图 4 所示。基于特征点矢量，本文计算特征点之间的欧氏距离 $d_{ij}(i, j = 1, 2, \dots, 26, i \neq j)$ 如式(4)所示。

$$d_{ij} = \|p_i - p_j\| = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (4)$$

表 1 特征点编号及对应名称

编号	名称	编号	名称
1	右眉尾	14	左下眼睑
2	右眉中	15	右前额
3	右眉头	16	左前额
4	右外眼角	17	右鼻翼
5	右上眼睑	18	鼻小柱
6	右内眼角	19	左鼻翼
7	右下眼睑	20	右嘴角
8	右眉尾	21	唇珠
9	右眉中	22	左嘴角
10	右眉头	23	下唇尖
11	左外眼角	24	右腮
12	左上眼睑	25	左腮
13	左内眼角	26	下巴

对表演者的中性与其它表情图像分别进行特征点定位，分别记为： p_i^n, p_i^e ，其中图像大小均为 640×490 。表演者 1，表演者 2 中性与惊讶表情特征点定位如图 5 所示。

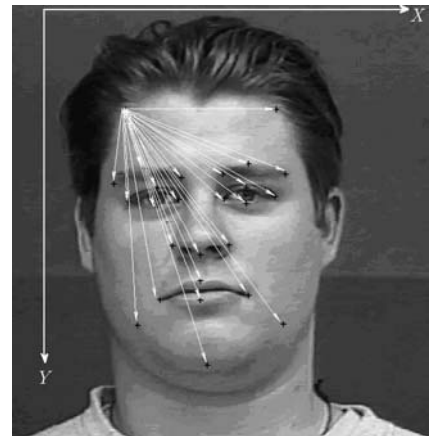


图 4 特征点 15 相对于各点的特征点矢量示意图

在得到表演者在中性及某一表情状态下的特征点位置后，计算特征点之间的距离 d_{ij}^n, d_{ij}^e ，然后将距离标准化，得到特征点距离比例系数 k_d 如式(5)所示。

$$k_d = d_{ij}^e / d_{ij}^n \quad (5)$$

表演者 1，表演者 2 惊讶与中性的特征点距离比例系数 k_d 分别如图 6 所示，由于共有 26 个特征点，根据 $C_{26}^2 = 325$ ，因此，图 6 横轴最大序号为 325。

从大量实验研究中发现，一些特征点之间的距离比例系数在所有的表情条件下变化并不明显，在直方图中表现为其值始终在 1 附近波动，本文将满



图 5 表演者 1, 表演者 2 中性与惊讶表情特征点定位

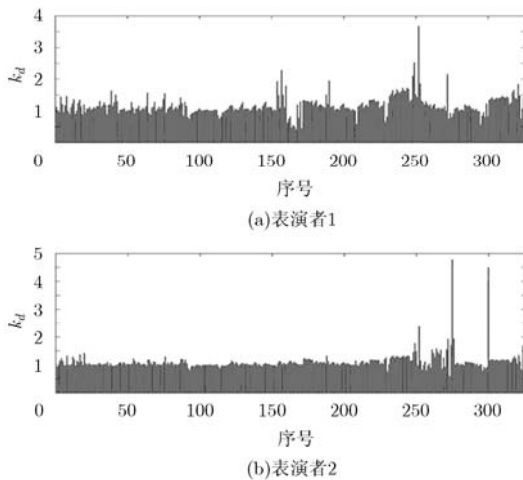


图 6 表演者 1, 表演者 2 惊讶与中性特征点距离比例系数

足条件 $|k_d - 1| \leq 0.1$ 的部分 k_d 去除, 其余定义为 k_{df} , 并作为最终参与后续训练学习的输入特征之一。

4 纹理形变能量参数

人在做某些表情时, 总会伴随着表情纹, 在面部图像上表现为纹理变化, 从而使原本平滑的皮肤表面变得复杂。如何利用表情纹理的变化进行人脸表情识别是学者们长期关注的研究热点。Song 等人^[24]选择一些出现表情纹理变化的特征块来检测皱纹, 提出皮肤形变参数的概念, 实现了较好的识别效果。Liu 等人^[27]指出表情比例图像能捕捉面部局部运动信息, 其中包括表情纹。考虑人在做表情时, 表情纹在额头与眉间最为明显, 因此, 本文选取额头及眉间两个特征块为讨论对象, 如图 7 所示。

两个特征块记为 $F_{15,16,9,2}$ 和 $F_{3,10,13,6}$, 其中下标代表特征点的序号。以特征块 $F_{3,10,13,6}$ 为例, 表演者做表情时, 相应的肌肉均在运动, 导致 $F_{3,10,13,6}$ 的纹理发生变化, 在图像上表现为灰度变化和出现明显边缘, 原先平滑的图像区域变得复杂, 从而导致图像能量增加, 这在图像频域中得到体现。本文提出纹理形变能量参数, 其计算过程如图 8 所示。

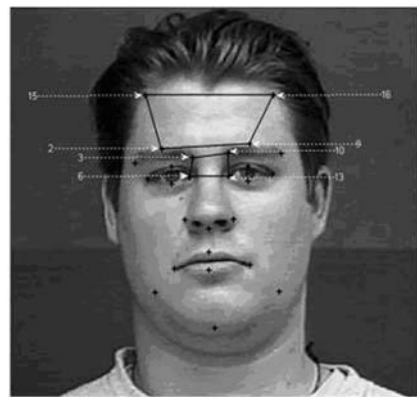


图 7 特征块示意图

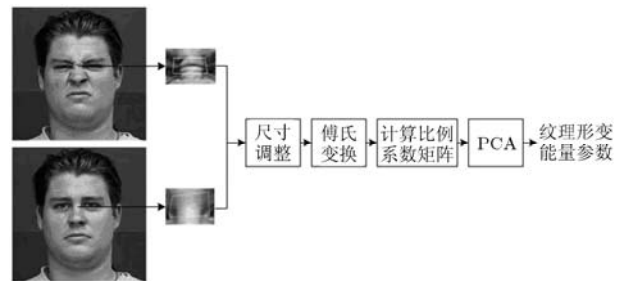


图 8 纹理形变能量参数获取流程

由于特征点的位置变化, 相应特征块的大小会发生变化。因此, 本文对选取的特征块通过中值插值的方法进行尺寸调整, 使之规格统一, 其中 $F_{15,16,9,2}$ 调整为 155×65 , $F_{3,10,13,6}$ 调整为 35×25 。考虑特征块内的纹理信息对特征点位置较为敏感, 对尺寸调整后的特征块进行傅氏变换, 如图 9 所示。

得到的表情频谱图与中性频谱图记为 $B1$ 与 $B2$, 大小为 $N \times M$ (可为 155×65 或者 35×25), 分别如式(6)与式(7)所示。

$$B1 = \begin{pmatrix} b1_{11} & \cdots & b1_{1N} \\ \vdots & \ddots & \vdots \\ b1_{M1} & \cdots & b1_{MN} \end{pmatrix} \quad (6)$$

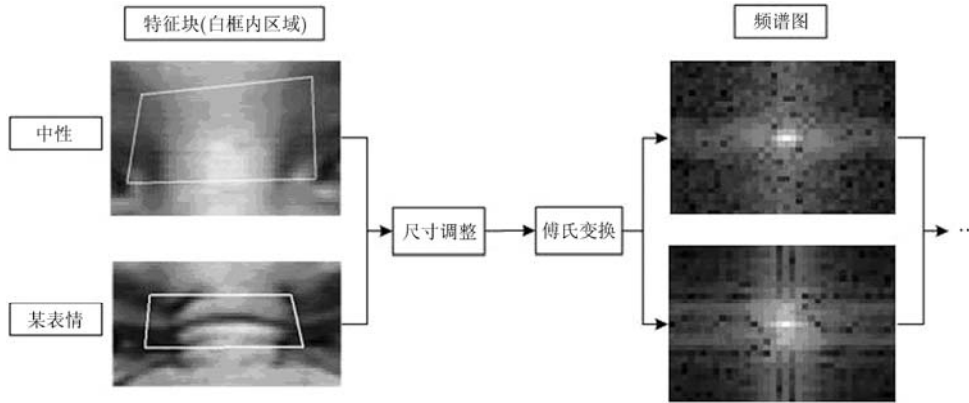


图 9 特征块傅氏变换

$$B2 = \begin{pmatrix} b2_{11} & \dots & b2_{1N} \\ \vdots & \ddots & \vdots \\ b2_{M1} & \dots & b2_{MN} \end{pmatrix} \quad (7)$$

由此，本文得到比例系数矩阵 B 如式(8)所示。

$$B = B1 / B2 = \begin{pmatrix} b1_{11}/b2_{11} & \dots & b1_{1N}/b2_{1N} \\ \vdots & \ddots & \vdots \\ b1_{M1}/b2_{M1} & \dots & b1_{MN}/b2_{MN} \end{pmatrix} = \begin{pmatrix} b_{11} & \dots & b_{1N} \\ \vdots & \ddots & \vdots \\ b_{M1} & \dots & b_{MN} \end{pmatrix} \quad (8)$$

最后将 B 进行 PCA，最终得到纹理形变能量参数，记为 k_{sf} 。

结合标准化后的特征点距离比例系数与纹理形变能量参数，最终的输入特征如式(9)所示。

$$k_f = k_{df} + k_{sf} \quad (9)$$

得到融合特征后，本文选择 RBF 神经网络进行表情训练与分类实验。人工神经网络(ANNs)是一种仿生神经网络，进行分布式并行信息处理的数学模型。ANNs 通过调整网络内部大量节点之间相互连接的关系实现信息处理的目的。RBF 神经网络是 ANN 重要的一支，优点如下：具有全局逼近的性质，且不存在局部极小问题；有较强的输入和输出映射能力，理论证明 RBF 神经网络在前向网络中拥有最优的完成映射功能；连接权值与输出表现为线性关系；分类能力强；收敛速度快。图 10 为 RBF 神经网络的结构图。

RBF 神经网络共分 3 层，即输入层、隐含层和输出层。每个隐含层神经元对应一个中心，RBF 取高斯函数。当输入特征向量，即表情矢量送入每个中心后，RBF 根据其到中心的欧氏距离做非线性变换，再通过连接权值 W 做线性运算进入输出层，得到最终结果。

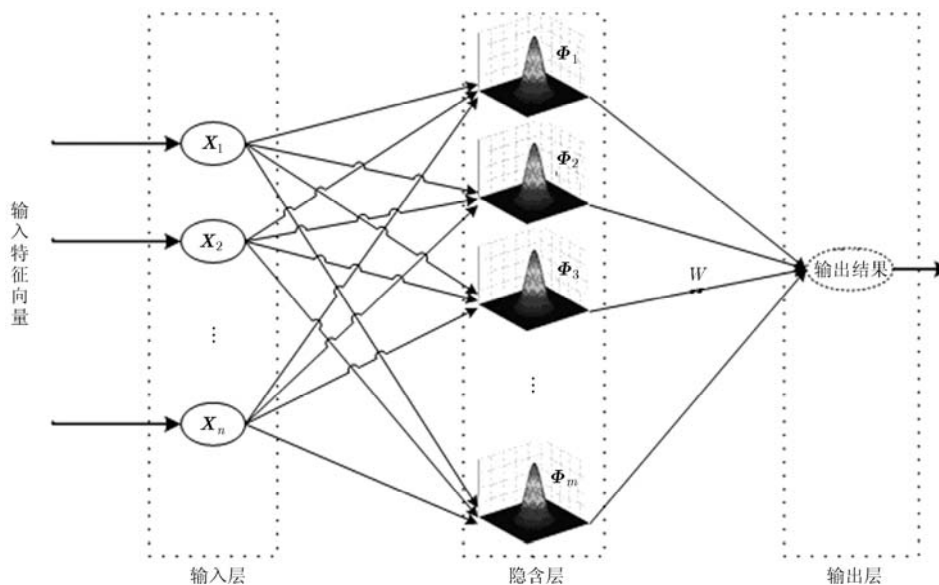


图 10 RBF 神经网络结构图

5 实验结果及分析

实验选择 Cohn-Kanade 数据库及 BHU(北京航空航天大学)人脸表情数据库。其中 Cohn-Kanade 数据库包括多于 100 位来自不同地区、不同肤色和不同年龄表演者的人脸表情序列。BHU 人脸表情数据库由 3 部分组成: 18 种单一表情、3 种混合表情和 4 种复杂表情构成, 32 位 21 至 25 岁的大学生(男性: 14 位, 女性: 18 位)参加了数据采集^[14]。部分 BHU 人脸表情数据库实例如图 11 所示。



图 11 BHU 人脸表情数据库实例图

本文选取 Cohn-Kanade 数据库中 100 位表演者的共计 520 位正面表情序列, 范围包括不同的肤色、光照和性别; 选取 BHU 人脸表情数据库中 6 种基本情感和中性表情, 由于每个表情序列都包括中性表情, 32 个对象每个表情序列各一组, 共计 192 组。实验设置 RBF 神经网络的输出为一个节点, 值可为 1, 2, 3, 4, 5, 6, 7, 相应的表示为高兴, 悲伤, 惊讶, 生气, 厌恶, 害怕, 中性 7 种表情。

对于 Cohn-Kanade 数据库, 实验设置训练样本为 60 位表演者的表情序列, 识别样本为 40 位表演者的表情序列, 实验结果如表 2 所示。从实验结果可以看出, 高兴、惊讶和中性 3 种表情的识别率排在前 3 位, 分别为: 94.8%, 92.7%和 91.3%, 而对另外 4 种情感的识别率都在 80%~90%之间。悲伤的识别率最低, 值为 80.2%。另外, 将生气理解为悲伤的误判率为 8.7%, 将厌恶理解为生气与将悲伤理解为厌恶的误判率均为 5.7%。本文方法的误判率还有待改进, 但总体上取得了良好的识别效果, 最终平均识别率为 87.0%。

对于 BHU 人脸表情数据库, 本文以 20 位表演者的表情序列为训练样本, 以 12 位表演者的表情序列为识别样本, 识别结果如表 3 所示。从表 3 可以看出, BHU 人脸表情数据库与 Cohn-Kanade 数据

表 2 Cohn-Kanade 数据库识别结果(%)

输入	输出						
	高兴	悲伤	惊讶	生气	厌恶	害怕	中性
高兴	94.8	0.2	0.7	0.8	0	3.5	0
悲伤	0.3	80.2	0.5	5.1	5.7	2.7	5.5
惊讶	1.3	0.3	92.7	0.5	0.7	4.5	0
生气	0	8.7	0.1	85.0	1.3	4.9	0
厌恶	0	4.7	0.8	5.7	83.7	4.3	0.8
害怕	0.2	3.4	3.8	4.2	4.1	81.2	3.1
中性	0	3.6	0.2	0.5	1.5	2.9	91.3

平均识别率: 87.0

表 3 BHU 人脸表情数据库识别结果(%)

输入	输出						
	高兴	悲伤	惊讶	生气	厌恶	害怕	中性
高兴	93.4	0.1	1.1	0.8	0	4.4	0.2
悲伤	0.5	79.8	0.5	6.2	5.7	2.2	5.1
惊讶	2.3	0.1	91.8	0.5	0.4	4.1	0.8
生气	0	10.3	0.5	81.2	0.9	5.7	1.4
厌恶	0	5.3	0.6	7.8	78.2	6.2	1.9
害怕	0.2	0.9	4.7	5.3	3.9	83.7	1.3
中性	0	4.6	0.3	0.4	1.3	1.7	91.7

平均识别率: 85.7

库的实验结果分布情况较为相似, 但总体识别效果比后者低, 最终平均识别率为 85.7%。

在均使用 RBF 神经网络的前提下, 实验使用不同的特征输入进行比较, 其结果如表 4 所示。从表 4 可以看出, 以 PCA, LDA, SNE 为基础的人脸表情识别效果均比本文所提方法的识别效果差, 识别率最高为 82.5%, 与本文所提方法相差 4.5%。表 4 也给出了单独使用特征点距离比例系数作为特征输入的识别率, 在两个数据库中分别为 83.6%与 82.5%, 这进一步验证了综合利用人脸的结构特征与纹理信息更有利于人脸表情识别研究。

表 4 不同特征经 RBF 神经网络分类器后的识别结果比较(%)

方法	Cohn-Kanade 数据库	BHU 人脸表情数据库
PCA	68.5	64.3
LDA	74.6	71.7
SNE	82.5	81.8
特征点距离比例系数	83.6	82.5
本文方法	87.0	85.7

本文也给出了在不同分类器条件下的识别结果, 如表 5 所示。所使用的 RBF 神经网络相比于 SVM, ADABOOST 更有优势, 识别效果更好, 这也表明 RBF 神经网络更能突显本文所提出的特征点矢量与纹理形变能量参数两个输入特征的优势, 实现了较好的识别效果。

表 5 本文所提特征在不同分类器条件下的识别结果比较(%)

分类器	Cohn-Kanade 数据库	BHU 人脸表情 数据库
SVM	77.3	74.1
ADABOOST	83.2	80.3
本文方法	87.0	85.7

本文算法考虑了个体差异对表情特征提取的影响。为分析算法复杂度, 所有实验均在 Inter(R) Core(TM)2 Duo 2.2 GHz 的 CPU, 2.00 GB 内存, Microsoft Windows XP 操作系统和 Matlab2011 的环境下进行。算法分为 3 部分, 即特征点定位、特征提取和识别分类。从样本输入到结果输出, 仅选取人脸关键部位的 26 个特征点参与运算, 且在求取特征点距离比例系数过程中较少涉及循环运算。另外, 所选取的特征块 $F_{15,16,9,2}$ 和 $F_{3,10,13,6}$ 尺寸分别为 155×65 和 35×25 , 计算量较小。因此本文算法的计算复杂度不大。

6 结束语

本文基于特征点矢量提出特征点距离比例系数, 并将之与纹理形变能量参数结合, 通过 RBF 神经网络进行人脸表情识别。该方法综合利用人脸图像的结构特征与纹理信息, 并在 Cohn-Kanade 数据库与 BHU 人脸表情数据库上进行了验证, 取得了较好的识别结果。如何更好地利用人脸表情序列的时空信息, 实现人脸身份信息与运动信息分离是下一步的研究重点。

参考文献

- [1] 谭华春, 章毓晋. 基于人脸相似度加权距离的非特定人表情识别[J]. 电子与信息学报, 2007, 29(2): 455-459.
Tan Hua-chun and Zhang Yu-jin. Person-independent facial expression recognition based on person-similarity weighted distance[J]. *Journal of Electronics & Information Technology*, 2007, 29(2): 455-459.
- [2] 王晓佩, 毛峽, Ishizuka Mitsuru. 基于非线性流形学习的人脸面部运动估计[J]. 电子与信息学报, 2011, 33(10): 2531-2535.
Wang Xiao-kan, Mao Xia, and Ishizuka M. Human face analysis with nonlinear manifold learning[J]. *Journal of Electronics & Information Technology*, 2011, 33(10): 2531-2535.
- [3] 廖海斌, 陈庆虎, 鄢煜尘. 基于因子分析的实用人脸识别研究[J]. 电子与信息学报, 2011, 33(7): 1611-1677.
Liao Hai-bin, Chen Qing-lu, and Yan Yu-chen. Practical face recognition via factor analysis[J]. *Journal of Electronics & Information Technology*, 2011, 33(7): 1611-1677.
- [4] 郑建炜, 王万良, 姚信威. 基于子块优化及全局整合的局部判别投影法[J]. 电子与信息学报, 2011, 33(9): 2175-2180.
Zheng Jian-wei, Wang Wan-liang, and Yao Xin-wei. Locally discriminant projection algorithm based on the block optimization and combination strategy[J]. *Journal of Electronics & Information Technology*, 2011, 33(9): 2175-2180.
- [5] Li Zheng and Mao Xia. EEMML: the emotional eye movement animation toolkit[J]. *Multimedia Tools and Applications*, 2012, 60(1): 181-201.
- [6] Moore S and Bowden R. Local binary patterns for multi-view facial expression recognition[J]. *Computer Vision and Image Understanding*, 2011, 115(4): 541-558.
- [7] Zhao Guo-ying, Huang Xiao-hua, Taini M, et al.. Facial expression recognition from near-infrared videos[J]. *Image Vision Computing*, 2011, 29(9): 607-619.
- [8] 张如艳, 王士同. 基于重尾噪声分布特性的多分类人脸识别方法[J]. 电子与信息学报, 2012, 34(3): 523-528.
Zhang Ru-yan and Wang Shi-tong. Multi-classification recognition method applied to facial image based on distribution characteristic of heavy-tailed noise[J]. *Journal of Electronics & Information Technology*, 2012, 34(3): 523-528.
- [9] Lajvardi S M and Wu H R. Facial expression recognition in perceptual color space[J]. *IEEE Transactions on Image Processing*, 2012, 21(8): 3721-3733.
- [10] Lajvardi S M and Hussain Z M. Automatic facial expression recognition: feature extraction and selection[J]. *Signal, Image Video Processing*, 2012, 6(1): 159-169.
- [11] Yi Ji and Khalid Idrissi. Automatic facial expression recognition based on spatiotemporal descriptors[J]. *Pattern Recognition Letters*, 2012, 33(10): 1373-1380.
- [12] Paul E and Friesen W V. Constant across cultures in the face and emotion[J]. *Personality and Social Psychology*, 1971, 17(2): 124-129.
- [13] Paul E and Friesen W V. Facial Action Coding System: Investigator's Guide[M]. Palo Alto, California: Consulting Psychologists Press, 1978, 156-163.
- [14] 薛雨丽, 毛峽, 张帆. BHU 人脸表情数据库的设计与实现[J]. 北京航空航天大学学报, 2007, 33(2): 224-228.
Xue Yu-li, Mao Xia, and Zhang Fan. Design and realization of BHU facial expression database[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2007, 33(2): 224-228.

- [15] Terence S, Simon B, and Maan B. The CMU Pose, illumination, and expression (PIE) database[C]. Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, Washington, 2002: 46–51.
- [16] Tenebaum J B, Silvam V D, and Langford J C. A global geometric framework for nonlinear dimensionality reduction[J]. *Science*, 2000, 290(5500): 2319–2323.
- [17] Roweis S T and Saul L K. Nonlinear dimensionality reduction by locally linear embedding[J]. *Science*, 2000, 290(5500): 2323–2326.
- [18] Hinton G and Roweis S. Stochastic neighbor embedding[C]. Advances in Neural Information Processing Systems, Vancouver, Canada, 2002: 833–840.
- [19] Kotsia I, Zafeiriou S, and Pitas I. Texture and shape information fusion for facial expression and facial action unit recognition[J]. *Pattern Recognition*, 2008, 41(3): 833–851.
- [20] Mahdi I and Hamed S H. A novel fuzzy facial expression recognition system based on facial feature extraction from color face images[J]. *Engineering Applications of Artificial Intelligence*, 2012, 25(1): 130–146.
- [21] Shi J, Samal A, and Marx D. How effective are landmarks and their geometry for face recognition? [J]. *Computer Vision and Image Understanding*, 2006, 102(2): 117–133.
- [22] Park S, Shin Jong-ju, and Kim Dai-jin. Facial expression analysis with facial expression deformation[C]. ICPR: 19th International Conference on Pattern Recognition, Tampa, 2008: 1–4.
- [23] Valstar M F and Pantic M. Biologically vs. logic inspired encoding of facial actions and emotions in video[C]. Proceedings of IEEE International Conference on Multimedia and Expo, Toronto, ON, Canada, 2006: 325–328.
- [24] Song Ming-li, Tao Da-cheng, Liu Zi-cheng, *et al.* Image ratio features for facial expression recognition application[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 2010, 40(3): 779–788.
- [25] Cootes T F, Taylor C J, Cooper D H, *et al.* Active shape models-their training and application[J]. *Computer Vision and Image Understanding*, 1995, 61(1): 38–59.
- [26] Cootes T, Edwards G J, and Taylor C J. Active appearance models[C]. 5th European Conference on Computer Vision Freiburg, Germany 1998, 2: 484–498.
- [27] Liu Zi-cheng, Shan Ying, and Zhang Zheng-you. Expressive expression mapping with ratio images[C]. Proceedings of ACM SIGGRAPH, Los Angeles, CA, 2001: 271–276.
- 易积政: 男, 1985 年生, 博士生, 研究方向为图像处理、模式识别、人工智能。
- 毛 峡: 女, 1952 年生, 博士, 教授, 博士生导师, 研究方向为模式识别、人工智能、情感计算。
- Ishizuka Mitsuru: 男, 1949 年生, 教授, 博士生导师, 研究方向为人机交互、情感计算、仿生代理等领域。