

基于概率协作表示的多表情序列融合识别

王秀友^{1,2*}, 范建中¹, 刘华明¹, 徐冬青¹, 刘争艳¹

¹ 阜阳师范学院计算机与信息工程学院, 安徽 阜阳 236037;

² 安徽大学计算机科学与技术学院, 安徽 合肥 230601

摘要 传统表情识别往往是基于单一图像进行特征提取、训练及识别, 缺乏在动态时间上的细微表情变化提取。利用时间前后的人脸表情变化信息, 提出了一种基于概率协作表示的多视频序列融合表情识别方法。先采用主动外观模型(AAM)定位出人脸表情的 68 个特征点, 利用提出的融合策略将相邻 3 帧表情图像的 AAM 特征进行融合, 最后利用概率协作表示的分类优势进行识别。实验结果表明, 在 CK+ 表情数据库上, 所提出的方法能够把握表情的时间变化信息, 相比于近几年的表情识别算法, 在识别率上取得了较好的效果。

关键词 机器视觉; 人脸表情识别; 主动外观模型; 多序列融合; 概率协作表示

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.131004

Multi-Expression Sequence Fusion Recognition Based on Probabilistic Cooperative Representation

Wang Xiuyou^{1,2*}, Fan Jianzhong¹, Liu Huaming¹, Xu Dongqing¹, Liu Zhengyan¹

¹ School of Computer and Information Engineering, Fuyang Normal University, Fuyang, Anhui 236037, China;

² School of Computer Science and Technology, Anhui University, Hefei, Anhui 230601, China

Abstract Traditional facial expression recognition often uses a single image to extract features, train, and recognize; however, subtle changes in dynamic facial expressions are not recognized. This study proposes a multi-expression sequence fusion recognition method based on probabilistic cooperative representation using the changes in facial expression before and after time. First, 68 feature points of facial expression are located using an active appearance model (AAM). Then, the AAM features of three adjacent facial expressions are combined using the proposed method. Finally, the classification advantages of probabilistic cooperative representation are used for recognition. Experimental results indicate that the proposed method can grasp the temporal change information of expression on the CK+ expression database. Moreover, this method can achieve higher recognition rates compared with traditional expression recognition algorithms.

Key words machine vision; facial expression recognition; active appearance model; multi-sequence fusion; probabilistic cooperative representation

OCIS codes 150.0155; 100.5010; 070.5010

1 引言

表情是人们进行沟通交流、表达意图与情感的重要途径, 是计算机理解人类的根基^[1-2]。在人工智能领域, 人脸表情识别作为人机交互的一种重要技术, 成为

广大学者研究的热点, 涉及图像处理、计算机视觉、模式识别等众多交叉学科, 有着广泛的应用前景^[3-5]。

通常表情识别以静态图像为研究对象, 是动态序列表情识别的基础, 由图像预处理、特征提取与分类识别 3 个部分组成。特征提取有主动外观模型

收稿日期: 2018-12-21; 修回日期: 2019-01-20; 录用日期: 2019-01-24

基金项目: 安徽省高校优秀青年骨干人才项目(gxfx2017072)、安徽省自然科学基金(1708085MF155)、阜阳市政府-阜阳师范学院横向合作科研项目(XDHX2016020, XDHX201710)、安徽省教育厅自然科学研究重点项目(KJ2018A0345)、阜阳师范学院青年人才基金重点项目(rcxm201706)

* E-mail: wangxiuyou@163.com

(AAM)^[6]、可变形部件模型(DPM)^[7]等基于几何的方法,有 Gabor 变换^[8]、时空局部三值方向角模式^[9]等基于像素、频域的提取方法。在分类识别方面,K 近邻^[10]、支持向量机^[11]、隐马尔可夫模型^[12]等浅层机器学习算法和卷积神经网络(CNN)^[13]、深度置信网络(DBN)^[14]、深度自编码器^[15]等深度学习的应用,都使得识别效果显著增强。然而,上述关于表情识别的研究存在的不足是仅以单一的静态图像作为输入来提取图像的空间信息,忽略了面部表情是一个连续变化的过程,其时域信息更能有效表达人脸表情。

在关于利用时序的表情识别中,文献[13]中利用卷积神经网络,搭建了人脸表情图像的多帧融合,但该融合只是针对输出结果,属于决策层的融合,在融合之前就已损失序列之间的相关信息;另外,其输入使用人脸整体区域,引入了大量与表情无关的特征,最终识别率仅达到 92.88%。文献[16]中虽然考虑到表情的动态时间问题,但仅利用滑动窗口动态时间规整算法选出表情明显的图像,最后依然只将单一图像输入到卷积神经网络中。文献[1]中将韦伯局部特征扩展到时空域,并融合分块光流直方图特征,达到了融合视频表情序列的时间信息,识别结果更为精确;然而,该方法在识别中采用支持向量机(SVM),对于帧与帧之间像素点的细微变化,SVM 的分类能力不具有优势。文献[17]中采用时空纹理映射的方法,较好地反映了表情在时间变化下的动态信息,但最后依然使用 SVM 进行分类。

针对上述问题,基于几何方法能够较好地描述表情变化的局部信息,且获取的数据量小,其中 AAM 将纹理信息与形状信息结合起来,其定位出的关键点可以表征人脸局部特征,在表情识别中经常使用^[8,18]。近年来稀疏表示分类(SRC)方法在图像分类^[19]、人脸识别^[20]等方面得到了广泛应用,而协作表示^[21](CR)在 SRC 中因具有低计算复杂度而逐渐受到关注。Cai 等^[22]在协作表示分类的基础上提出了概率协作表示的方法(ProCRC),其有明确的概率解释,最大化样本属于每一类的概率,更好地处理了识别对象的细微变化,在人脸识别上取得突出的效果。

综上所述,本文考虑到表情序列帧之间的相关性,沿着时间轴将表情图片堆叠成 3 维时空立体图,随后利用 AAM 在人脸定位出的特征点作为特征,提出融合策略,将时间轴上相邻的 3 帧图像进行融合,最后将 ProCRC 算法应用于表情识别中。实验

结果表明,该方法将表情序列的细微变化融合在一起,相比于单一图像的认识,不仅操作简单,而且识别率得到进一步提高。

2 AAM 模型

不同于人脸识别,表情识别主要是识别人脸用五官表达出来的感情,而与面部肤色、五官大小等用于人脸识别的特征无关,若在特征提取中加入这些干扰因素,势必会有一定的影响。

主动外观模型是一种应用非常广泛的统计模型,通常用来描述人脸表情以及特征提取,其特点是集人脸的形状信息与纹理信息于一体,能够获得面部各个关键点的精确位置。

AAM 算法包括形状模型与纹理模型的建立以及模型匹配。

建立形状模型时主要应用训练集已标记的位置信息,将所有标记点坐标按固定顺序排列起来,即

$$\mathbf{S} = (x_1, y_1, x_2, y_2, \dots, x_n, y_n)^T, \quad (1)$$

式中: $x_1 \sim x_n$ 为人脸关键点的横坐标; $y_1 \sim y_n$ 为人脸关键点的纵坐标; n 为关键点数; T 表示矩阵的转置。

AAM 允许形状包含线性变化,所以可将 \mathbf{S} 表示为

$$\mathbf{S} = \mathbf{S}_0 + \sum_{i=1}^n p_i \mathbf{S}_i = \mathbf{S}_0 + \mathbf{Q}_s \mathbf{P}, \quad (2)$$

式中: $p_i (i=1, 2, \dots, n)$ 是形状参数; \mathbf{S}_0 是平均形状向量; \mathbf{Q}_s 是由 n 个形状向量 $\mathbf{S}_i (i=1, 2, \dots, n)$ 构成的特征矩阵; \mathbf{P} 是形状参数向量。

建立纹理模型是将所有人脸变形到平均形状并采样,得到不同图像中采样出来的像素意义对应的纹理样本,任意纹理图像 $\mathbf{A}(x)$ 可估计为

$$\mathbf{A}(x) = \mathbf{A}_0(x) + \sum_{i'=1}^m \lambda_{i'} \mathbf{A}_{i'}(x) = \mathbf{A}_0 + \mathbf{Q}_g \mathbf{A}, \quad \forall x \in \mathbf{S}_0, \quad (3)$$

式中: $\mathbf{A}(x)$ 是在平均形状上的图像纹理向量; x 是像素点; $\mathbf{A}_0(x)$ 是平均纹理; $\lambda_{i'}$ 是纹理参数; \mathbf{A} 是纹理参数向量; \mathbf{Q}_g 是由 m 个纹理向量 $\mathbf{A}_{i'}(x) (i'=1, 2, \dots, m)$ 构成的特征矩阵。

AAM 匹配过程就是最小化待测图像和模型之间的差的过程,这一优化过程主要通过改变模型参数实现,即

$$\sum_{x \in \mathbf{S}_0} \left\{ \mathbf{A}_0(x) + \sum_{i'=1}^m \lambda_{i'} \mathbf{A}_{i'}(x) - \mathbf{I}[\mathbf{W}(x; \mathbf{P})] \right\}^2, \quad (4)$$

式中: $\mathbf{W}(x; \mathbf{P})$ 表示形状向量中的像素点和平均形状中的像素点 x 相对应, 并由形状参数 \mathbf{P} 控制; $\mathbf{I}[\mathbf{W}(x; \mathbf{P})]$ 是该像素点上的纹理向量。

文献[6, 10]中对特征点提取的个数有 28, 48, 51, 68 等。为了在表征表情部分上能够较为充分地覆盖多个关键点, 本文选择 68 个特征点。利用 AAM 算法对某人脸图像特征点进行提取, 其结果如图 1 所示, 可以看出该 68 个特征点能够较好地描述出人脸形状, 且能有效地表示出人脸的轮廓和五官, 包括脸颊、嘴唇、眼睛、眉毛、鼻子这几个主要部分, 而当表情发生变化时, 喜、怒、哀、乐等表情信息也正是由这几个部分共同体现, 因此通过 AAM 模型的特征点提取, 保留了主要的表情变化部分, 同时也除去了一些对表情识别无用的干扰因素。

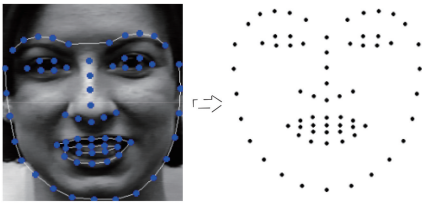


图 1 对人脸特征点的提取

Fig. 1 Extraction of facial feature points

3 表情序列融合策略

在表情识别中一般仅使用单一人脸图像进行特征提取, 而人脸表情通常是一个动态的连续变化过程, 因此基于该思想, 在静态图像的基础上, 沿着时间轴加入了该帧图像前一帧与后一帧图像的动态

变化信息, 提出了两种融合相邻帧表情序列的方法。

在进行融合前, 需要将表情图像的 AAM 特征点进行序列化, 如图 2 所示, 将表情特征分为 8 部分 (依次为左边脸颊、右边脸颊、左边眉毛、右边眉毛、鼻子、左眼睛、右眼睛、嘴巴) 并按顺序合并为一个完整表情序列。由于每个特征点 f_i 都具有横坐标 x_i 与纵坐标 y_i , 在序列化时, 将每个点的横坐标排成一行, 随后将纵坐标接在后面排成一行, 形成该表情图像 AAM 的一个序列, 即 $\mathbf{F}_j = (x_{f_1}^{(j)}, \dots, x_{f_{68}}^{(j)}, y_{f_1}^{(j)}, \dots, y_{f_{68}}^{(j)})^T$, 其中: \mathbf{F}_j 表示第 j 幅图像, $x_i^{(j)}$ 和 $y_i^{(j)}$ 分别表示第 j 幅图像的第 i 个特征点的横坐标和纵坐标。

图 3 为相邻帧图像融合过程。若相邻 3 帧表情图像的特征点序列分别为 $\mathbf{F}_1 = (x_{f_1}^{(1)}, \dots, x_{f_{68}}^{(1)}, y_{f_1}^{(1)}, \dots, y_{f_{68}}^{(1)})^T$, $\mathbf{F}_2 = (x_{f_1}^{(2)}, \dots, x_{f_{68}}^{(2)}, y_{f_1}^{(2)}, \dots, y_{f_{68}}^{(2)})^T$, $\mathbf{F}_3 = (x_{f_1}^{(3)}, \dots, x_{f_{68}}^{(3)}, y_{f_1}^{(3)}, \dots, y_{f_{68}}^{(3)})^T$, 融合策略为将 3 个特征序列向量依次进行拼接, 融合公式为

$$\mathbf{c}(\mathbf{F}_1, \mathbf{F}_2, \mathbf{F}_3) = (x_{f_1}^{(1)}, \dots, x_{f_{68}}^{(1)}, y_{f_1}^{(1)}, \dots, y_{f_{68}}^{(1)}, x_{f_1}^{(2)}, \dots, x_{f_{68}}^{(2)}, y_{f_1}^{(2)}, \dots, y_{f_{68}}^{(2)}, x_{f_1}^{(3)}, \dots, x_{f_{68}}^{(3)}, y_{f_1}^{(3)}, \dots, y_{f_{68}}^{(3)})^T. \quad (5)$$

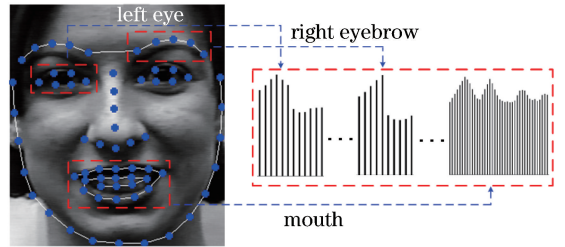


图 2 AAM 特征序列化示意图

Fig. 2 Diagram of AAM feature serialization

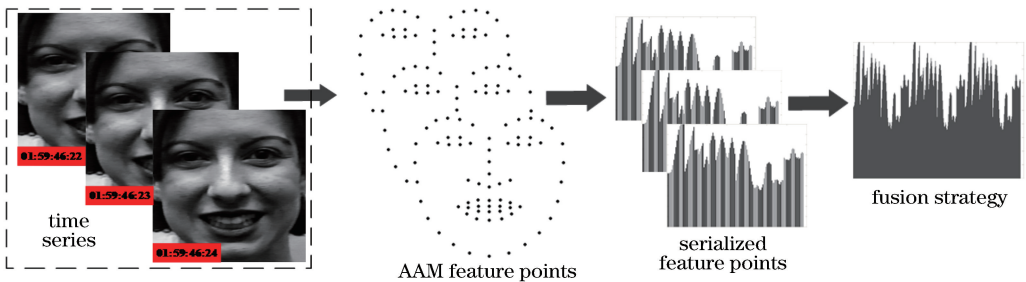


图 3 相邻帧图像融合过程

Fig. 3 Image fusion process of adjacent frames

4 概率协作表示

概率协作表示的基本原理可用图 4 进行解释。假设有 C 类样本集 $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_C\}$, 其中 \mathbf{X}_C 是 C 类样本构成的矩阵, 每一列为一个样本向量,

且 l_x 为样本集中类别的标签组, 另外设 S 为跨越 \mathbf{X} 中所有类的线性协作子空间。若数据点 x 在该子空间中, 则可以用 \mathbf{X} 线性表示为 $x = \mathbf{X}\alpha$, 其中, α 定义为表示向量。由于 \mathbf{X} 包含所有类的样本, 则协作子空间 S 比单独每类 \mathbf{X}_C 的子空间大得多, 因此尽

管数据点 x 落在 S 中,但是具体会被列为标签组 l_x 中哪一类也会随着向量 α 的组合不同而变化。图 4 为一组来自不同个体的人脸表情图像,可以看出,由于 x_1 的系数 α_1 比 x_2 的系数 α_2 小,即 x_1 更接近子空间的中心区域并且 x_2 更靠近边缘区域,因此 x_1 有更大的可能性被列为 l_x 组中。其相关公式与计算可参考文献[22]。

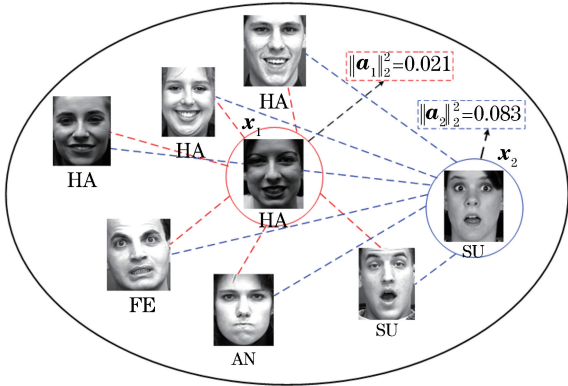


图 4 概率协作表示子空间示意图

Fig. 4 Probabilistic cooperative representation subspace diagram

5 实验结果及分析

5.1 实验数据与训练测试集预处理

为验证方法的有效性,实验采用 CK + (extended Cohn-Kanade dataset) 数据库。该数据库为当前评估面部表情识别应用最为广泛的数据库,共包含 123 名实验对象以及 593 个表情序列,每个序列至少有 10 帧图像组成,组成了人脸表情变化由平静到高峰到回落的 3 个过程。实验选取了经常使用的 7 种基本的人脸表情:生气、厌恶、恐惧、高

兴、悲伤、惊讶、中性,如图 5 所示。

选取已标注的 225 个视频序列,其中生气、厌恶、恐惧、高兴、悲伤、惊讶、中性的对象数量分别为 24, 32, 21, 50, 23, 51, 24, 其中,除中性外每个序列包含表情开始到高峰的 8 帧图像。由于要选择连续的 3 帧图像进行融合,因此对每一个序列按照如图 6 所示的方法进行处理,这样每个视频序列中 8 帧图像最终得到 6 组数据(图中①~⑥表示分组序号),则一共得到 1206 组数据组成训练集。而对于测试集,要识别的是单个图像,为保证维度一致,将单个表情图像首先复制为 3 份,然后再按照融合策略进行融合,则一共有 1608 组数据组成测试集。

实验验证中,采用文献[23]中的样本选取方法,即“与人有关”和“与人无关”,其中“与人有关”的选取方法将每个对象的每类表情一部分用于训练,一部分用于测试;“与人无关”的选取方法将一部分对象所有表情用于训练,其余未参与训练的对象用于测试。因此,实验中“与人有关”的选取方法是随机抽取训练集中的 18 个对象,随后抽取每个对象的每种表情中①②③组融合数据,该对象对应的测试集中,将剩下 3 帧未参与训练的图像用于测试;“与人无关”的选取方法是在每种表情中随机抽取训练集中的 18 个对象,剩下未参与训练的对象作为测试集。测试中,分别进行 6 类(除去中性表情)和 7 类表情的 10 次随机抽取识别实验。

依托 MATLAB 2014a 平台对提出的融合方法进行实验对比验证,实验环境:CPU 为 Inter(R) Core(TM) i7-8750H, 8G 内存, Win10, 64 位操作系统。

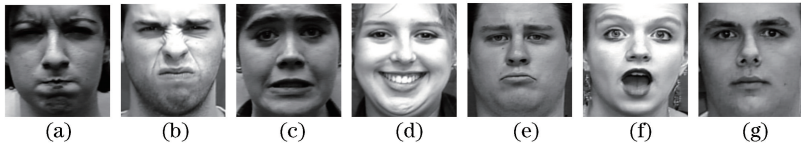


图 5 CK+人脸表情示例。(a)生气;(b)厌恶;(c)恐惧;(d)高兴;(e)悲伤;(f)惊讶;(g)中性

Fig. 5 CK+ facial expression examples. (a) Anger; (b) disgust; (c) fear; (d) happy; (e) sad; (f) surprise; (g) neutral

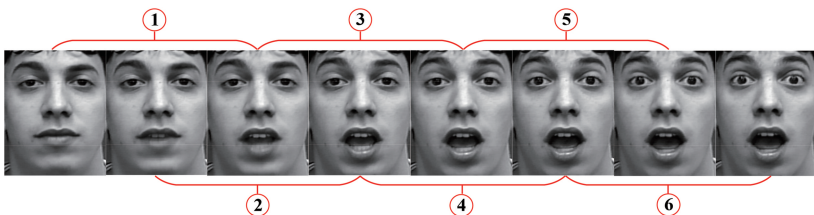


图 6 测试数据预处理

Fig. 6 Test data preprocessing

5.2 参数选择

由文献[22]中(12)式可知,在求解 α 过程中需要调节 λ, γ 两个参数,在此以某一固定的训练测试集进行融合策略的实验,得到两参数下不同值的识别率直方图,如图7所示,当 $\lambda=0.1, \gamma=0.01$ 时,得到的识别效果最佳。

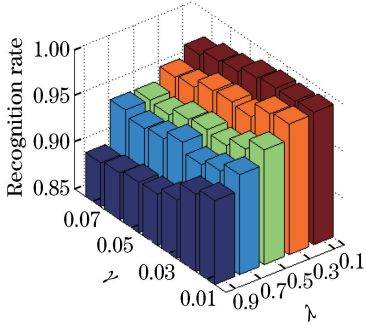


图7 不同参数下的识别率

5.3 实验结果分析

为验证采用该融合方式的可行性,实验使用Ori-AAM以及Ori-pic与融合策略进行对比,其中Ori-AAM表示使用单一的图像AAM特征点, Ori-pic表示使用数据库原始图像。将通过Adaboost人脸检测算法得到的人脸所在区域的大小均归一化为 $32 \text{ pixel} \times 32 \text{ pixel}$ 。对上述得到的数据通过ProCRC进行10次随机抽取实验,最终得到识别结果。表1为在“与人无关”样本选取方法下识别6类表情的测试结果,表2为各种状态下对比结果的汇合。从表1和表2可知,由于Ori-pic引入了与表情无关的大量信息,识别率较低;相比于Ori-AAM,本文算法中表情序列之间的融合使其在6类表情识别中“与人有关”和“与人无关”两种情况下的识别率分别提高了6.5%、17.2%,在7类表情识别中“与人有关”和“与人无关”两种情况下的识别率分别提高了9.5%、14.7%,故该融合方式是可行有效的。

Fig. 7 Recognition rate under different parameters

表1 融合方法识别率对比

Table 1 Comparison of recognition rates of fusion methods

Method	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8	Test 9	Test 10	Average
Ours	91.1	88.4	86.7	87.7	88.0	87.1	91.2	86.5	88.1	90.3	88.5
Ori-AAM	67.7	71.2	73.9	71.5	71.9	70.1	68.1	73.3	72.9	72.3	71.3
Ori-pic	59.4	66.7	61.5	60.2	63.9	64.9	61.1	65.8	61.8	62.9	62.8

表2 各状态下的识别率

Table 2 Recognition rate in each state

Method	6-class		7-class	
	Related to people	Not related to people	Related to people	Not related to people
Ours	99.4	88.5	99.0	83.2
Ori-AAM	92.9	71.3	89.5	68.5
Ori-pic	90.2	62.8	88.0	54.3

为验证整体方法的有效性,在“与人有关”的实验中,将本文结果与文献[1,3,13,17,24,25]的识别结果进行对比,如表3所示。可以看出文献[1,17,25]中虽然都考虑到了表情变化的时间因素,但使用的SVM算法在一定程度上使识别性能受限。文献[13]中将多个CNN的识别结果进行融合,在一定程度上利用了表情序列的信息,但其最终属于决策层的融合,而相比于本文的数据层融合,该方法在融合之前就已损失序列之间的相关信息,另外,其使用人脸整体区域作为输入,引入了大量与表情无关的特征。文献[24]在利用CNN提取到表情的共同特征之后,进一步细化提取表情的特定特征,故最后的识别率较高。文献[3]与本文都使用了协作表示的方法进行分类,识别率相差无几,说明基于协作表示的分类方法具有良好的分类效果。本文方法基本上

高于对比的文献(与文献[3,24]持平),表明了本文整体方法的有效性。在“与人无关”的实验中,将本文结果与文献[23,26-28]的识别结果进行对比,如表4所示。可以看出,本文方法能够把握住表情的时间变化信息,即使在识别未训练的人脸表情时,相比于其他文献,依然能取得较好的识别效果。另一方面,其他文献仅选择表情强度最大的几帧进行训练,而本文的融合能够扩展到表情初始部分,扩展了训练样本。

6 结论

提出了利用AAM特征点将多表情序列进行融合,并采用概率协作表示进行识别的方法,相比于其他文献方法,具有以下优点:1)使用AAM特征点,能够较好地描述表达表情器官的结构轮廓,排除了

表3 与人有关情况下不同方法识别率

Table 3 Recognition rates of different methods

related to person		%	
Document	Method	6-class	7-class
Ref. [1]	STWLD+BHOF+SVM	91.6	—
Ref. [3]	DDL+CRC	99.5	98.9
Ref. [13]	5-frame CNN	92.9	—
Ref. [17]	STTM+SVM	97.7	—
Ref. [24]	ETEL_J	—	99.0
Ref. [25]	GLDPE	97.1	—
Ours	3-frame AAM+ProCRC	99.4	99.0

表4 与人无关情况下不同文献识别率

Table 4 Recognition rates of different documents

unrelated to person			
Document	Method	6-class	7-class
Ref. [23]	CS-LSBP+HOAG+ CCA+SVM	—	81.4
Ref. [26]	SLFDA	78.7	—
Ref. [27]	AdaBoost+DBN	87.4	—
Ref. [28]	WPHOG_TOP+SVM	—	84.9
Ours	3-frame AAM+ProCRC	88.5	83.2

人脸中不属于表情部分的干扰因素;2) 融合相邻表情图像的时间变化信息,避免单一图像的局限性,获取了一定的动态特征;3) 利用 ProCRC 分类方法的优越性,大幅提高融合表情图像的识别性能。本文从特征提取到分类,整体方法简单高效,易于实现,但未考虑表情中存在遮挡、头部姿态变化大等因素,因此下一步工作将围绕该方面展开。

参 考 文 献

- [1] Wang X H, Xia C, Hu M, *et al.* Facial expression recognition based on the fusion of spatio-temporal features in video sequences[J]. *Journal of Electronics & Information Technology*, 2018, 40(3): 626-632.
王晓华, 夏晨, 胡敏, 等. 融合时空特征的视频序列表情识别[J]. *电子与信息学报*, 2018, 40(3): 626-632.
- [2] He Z C, Zhao L Z, Chen C. Convolution neural network with multi-resolution feature fusion for facial expression recognition[J]. *Laser & Optoelectronics Progress*, 2018, 55(7): 071503.
何志超, 赵龙章, 陈闯. 用于人脸表情识别的多分辨率特征融合卷积神经网络[J]. *激光与光电子学进展*, 2018, 55(7): 071503.
- [3] Moeini A, Faez K, Moeini H, *et al.* Facial expression recognition using dual dictionary learning [J]. *Journal of Visual Communication and Image*

Representation, 2017, 45: 20-33.

- [4] Silva Cruz E A, Jung C R, Esparza Franco C H. Facial expression recognition using temporal POEM features[J]. *Pattern Recognition Letters*, 2018, 114: 13-21.
- [5] Ahmed F, Paul P P, Gavrilova M L. Evolutionary fusion of local texture patterns for facial expression recognition[C]//2015 IEEE International Conference on Image Processing (ICIP), September 27-30, 2015, Quebec City, QC, Canada. New York: IEEE, 2015: 1031-1035.
- [6] Zhou H L, Lam K M, He X J. Shape-appearance-correlated active appearance model [J]. *Pattern Recognition*, 2016, 56: 88-99.
- [7] Zhu X X, Ramanan D. Face detection, pose estimation, and landmark localization in the wild[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), June 16-21, 2012, Providence, RI, USA. New York: IEEE, 2012: 2879-2886.
- [8] Zhou J, Zhang S E, Mei H Y, *et al.* A method of facial expression recognition based on Gabor and NMF[J]. *Pattern Recognition and Image Analysis*, 2016, 26(1): 119-124.
- [9] Jiang M X, Hu M, Wang X H, *et al.* Dual-modal emotion recognition based on facial expression and body posture in video sequences [J]. *Laser & Optoelectronics Progress*, 2018, 55(7): 071004.
姜明星, 胡敏, 王晓华, 等. 视频序列中表情和姿态的双模态情感识别[J]. *激光与光电子学进展*, 2018, 55(7): 071004.
- [10] Xu L F, Wang J Y, Cui J N, *et al.* Dynamic expression recognition based on dynamic time warping and active appearance model[J]. *Journal of Electronics & Information Technology*, 2018, 40(2): 338-345.
许良凤, 王家勇, 崔婧楠, 等. 基于动态时间规整和主动外观模型的动态表情识别[J]. *电子与信息学报*, 2018, 40(2): 338-345.
- [11] Farajzadeh N, Hashemzadeh M. Exemplar-based facial expression recognition [J]. *Information Sciences*, 2018, 460/461: 318-330.
- [12] Siddiqi M H, Ali R, Khan A M, *et al.* Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields[J]. *IEEE Transactions on Image Processing*, 2015, 24(4): 1386-1398.
- [13] Zhang J G, Fang Y, Yuan H, *et al.* Multiple

- convolutional neural networks for facial expression sequence recognition [J]. Journal of Xidian University, 2018, 45(1): 150-155.
- 张金刚, 方圆, 袁豪, 等. 一种识别表情序列的卷积神经网络[J]. 西安电子科技大学学报, 2018, 45(1): 150-155.
- [14] Wang L L, Liu J H, Fu X M. Facial expression recognition based on fusion of local features and deep belief network [J]. Laser & Optoelectronics Progress, 2018, 55(1): 011002.
- 王琳琳, 刘敬浩, 付晓梅. 融合局部特征与深度置信网络的人脸表情识别[J]. 激光与光电子学进展, 2018, 55(1): 011002.
- [15] Zeng N Y, Zhang H, Song B Y, *et al.* Facial expression recognition via learning deep sparse autoencoders[J]. Neurocomputing, 2018, 273: 643-649.
- [16] Hu M, Zhang K K, Wang X H, *et al.* Video facial expression recognition combined with sliding window dynamic time warping and CNN[J]. Journal of Image and Graphics, 2018, 23(8): 1144-1153.
- 胡敏, 张柯柯, 王晓华, 等. 结合滑动窗口动态时间规整和CNN的视频人脸表情识别[J]. 中国图象图形学报, 2018, 23(8): 1144-1153.
- [17] Kamarol S K A, Parkkinen J, Jaward M H, *et al.* Spatiotemporal feature extraction for facial expression recognition[J]. IET Image Processing, 2016, 10(7): 534-541.
- [18] Cheon Y, Kim D. Natural facial expression recognition using differential-AAM and manifold learning[J]. Pattern Recognition, 2009, 42(7): 1340-1350.
- [19] Yang M, Zhang L, Feng X C, *et al.* Sparse representation based fisher discrimination dictionary learning for image classification [J]. International Journal of Computer Vision, 2014, 109(3): 209-232.
- [20] Lai J, Jiang X D. Modular weighted global sparse representation for robust face recognition[J]. IEEE Signal Processing Letters, 2012, 19(9): 571-574.
- [21] Wei D M, Zhou W D. Face recognition using collaborative representation with neighbors [J]. Journal of Xidian University, 2015, 42(3): 115-121.
- 魏冬梅, 周卫东. 近邻样本协作表示的人脸识别算法[J]. 西安电子科技大学学报, 2015, 42(3): 115-121.
- [22] Cai S J, Zhang L, Zuo W M, *et al.* A probabilistic collaborative representation based approach for pattern classification[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 2950-2959.
- [23] Hu M, Teng W D, Wang X H, *et al.* Facial expression recognition based on local texture and shape features [J]. Journal of Electronics & Information Technology, 2018, 40(6): 1338-1344.
- 胡敏, 滕文娣, 王晓华, 等. 融合局部纹理和形状特征的人脸表情识别[J]. 电子与信息学报, 2018, 40(6): 1338-1344.
- [24] Huang Y, Yan Y, Chen S, *et al.* Expression-targeted feature learning for effective facial expression recognition[J]. Journal of Visual Communication and Image Representation, 2018, 55: 677-687.
- [25] Li Y Q, Li Y J, Li H B, *et al.* Fusion of global and local various feature for facial expression recognition [J]. Acta Optica Sinica, 2014, 34(5): 0515001.
- 李雅倩, 李颖杰, 李海滨, 等. 融合全局与局部多样性特征的人脸表情识别[J]. 光学学报, 2014, 34(5): 0515001.
- [26] Wang Z, Ruan Q Q, An G Y. Facial expression recognition using sparse local fisher discriminant analysis[J]. Neurocomputing, 2016, 174: 756-766.
- [27] Li Y Q, Wang S F, Zhao Y P, *et al.* Simultaneous facial feature tracking and facial expression recognition [J]. IEEE Transactions on Image Processing, 2013, 22(7): 2559-2573.
- [28] Fan X J, Yang X B, Ye Q L, *et al.* A discriminative dynamic framework for facial expression recognition in video sequences [J]. Journal of Visual Communication and Image Representation, 2018, 56: 182-187.