



(12) 发明专利申请

(10) 申请公布号 CN 112766112 A

(43) 申请公布日 2021. 05. 07

(21) 申请号 202110023841.4

(22) 申请日 2021.01.08

(71) 申请人 山东大学

地址 266237 山东省青岛市即墨滨海路72号

(72) 发明人 杨明强 闫汝玉

(74) 专利代理机构 济南圣达知识产权代理有限公司 37221

代理人 黄海丽

(51) Int. Cl.

G06K 9/00 (2006.01)

G06K 9/62 (2006.01)

权利要求书2页 说明书11页 附图6页

(54) 发明名称

基于时空多特征融合的动态表情识别方法及系统

(57) 摘要

本发明公开了基于时空多特征融合的动态表情识别方法及系统,包括:获取待识别的人脸表情图像序列;对待识别的人脸表情图像序列进行预处理;对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。本发明方法对LBP-TOP算子进行改进,增强了特征的表达能力和特征提取速率。本发明提出了一种新的几何特征表达方式,直接将归一化后的人脸关键点作为几何特征,有效地降低特征维度。



1. 基于时空多特征融合的动态表情识别方法,其特征是,包括:
获取待识别的人脸表情图像序列;
对待识别的人脸表情图像序列进行预处理;
对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;
对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;
对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。
2. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对待识别的人脸表情图像序列进行预处理;具体包括:
对待识别的人脸表情图像序列中的所有帧人脸表情图像,均进行灰度归一化处理;
对灰度归一化处理后的脸表情图像,进行裁剪;
对裁剪得到的人脸表情图像进行几何归一化处理。
3. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对灰度归一化处理后的脸表情图像,进行裁剪;具体包括:
对灰度归一化后的图像采用SDM算法检测到人脸49个关键点,确定两只眼睛瞳孔的坐标,这两只眼睛瞳孔之间的距离为 d ,以两只眼睛瞳孔距离的中点为中心点,将动态人脸图像序列中的每帧人脸图像进行裁剪,裁剪后图片的宽为 $1.2 \times d$,高为 $2.25 \times d$ 。
4. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对预处理后的人脸表情图像序列,提取时空特征;其中,时空特征是指改进的LBP-TOP直方图特征;时空特征的提取步骤具体包括:
S1031: 对于一个宽为 X 、高为 Y 、帧数为 T 的图像序列,分别计算图像序列 XT 平面、 YT 平面以及峰值帧图像 XY 平面的LBP直方图向量;
S1032: 将表情序列 XT 和 YT 平面的LBP直方图向量与 XY 平面表情峰值帧的LBP直方图向量级联,从而得到改进的LBP-TOP直方图特征。
5. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对表情峰值关键帧图像,分别采集纹理特征和几何特征;其中,纹理特征是指Gabor特征;纹理特征的提取步骤具体包括:
采用5个尺度8个方向的Gabor滤波器,对归一化后的表情峰值帧图像提取Gabor幅值特征作为纹理特征。
6. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对表情峰值关键帧图像,分别采集纹理特征和几何特征;几何特征的提取步骤包括:
对表情的峰值帧图像,采用SDM算法检测到人脸49个关键点,将 X 轴和 Y 轴的坐标分别进行均值为0,方差为1的标准化操作;将98维特征向量作为人脸表情的几何特征。
7. 如权利要求1所述的基于时空多特征融合的动态表情识别方法,其特征是,对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果;具体包括:
对第一、第二和第三识别结果进行多数投票,通过集成学习的多数投票法得到分类结果。
8. 基于时空多特征融合的动态表情识别系统,其特征是,包括:

获取模块,其被配置为:获取待识别的人脸表情图像序列;

预处理模块,其被配置为:对待识别的人脸表情图像序列进行预处理;

特征提取模块,其被配置为:对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;

识别模块,其被配置为:对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;

输出模块,其被配置为:对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。

9. 一种电子设备,其特征是,包括:一个或多个处理器、一个或多个存储器、以及一个或多个计算机程序;其中,处理器与存储器连接,上述一个或多个计算机程序被存储在存储器中,当电子设备运行时,该处理器执行该存储器存储的一个或多个计算机程序,以使电子设备执行上述权利要求1-7任一项所述的方法。

10. 一种计算机可读存储介质,其特征是,用于存储计算机指令,所述计算机指令被处理器执行时,完成权利要求1-7任一项所述的方法。

基于时空多特征融合的动态表情识别方法及系统

技术领域

[0001] 本申请涉及图像处理技术领域,特别是涉及基于时空多特征融合的动态表情识别方法及系统。

背景技术

[0002] 本部分的陈述仅仅是提到了与本申请相关的背景技术,并不必然构成现有技术。

[0003] 随着人工智能和计算机技术的不断发展,人脸表情识别也成为研究热点。机器通过识别人脸面部表情的变化,分析人类的内在心理活动,从而更好地为人类服务,促进人工智能的进一步发展。通过人脸表情进行情感识别是情感计算中最有效、最直接的方法之一。面部表情以非语言的形式反映人类情感状态,是人类情感表达系统中不可缺少的组成部分。研究表明,在口语交流中,说话者想要传递的情感信息有55%来自面部表情,面部表情包含最多的情感信息。心理学家将人类的情绪划分为六种基本情感,即愤怒、厌恶、恐惧、快乐、悲伤和惊讶,每种情感有其相应的表情。静态和动态方法的融合可以突破各自的局限性,面部表情的峰值帧具有很强的表达能力,而时间信息在特殊的视频分类任务中是必不可少的。而且,单个特征不够全面和丰富,不足以捕获全局信息,因此需要设计一个鲁棒的特征提取算子将多个互补特征进行融合。

[0004] 目前,表情识别的研究有两种主流方法:基于图像的方法和基于视频的方法。基于图像的方法直接对表情的峰值帧进行分析和提取特征,是一种静态的方法,该方法通常分为两类:基于几何特征的方法和基于外观特征的方法;而基于视频的方法是对图像序列进行处理,提取表情序列中的时间和运动信息,是一种动态的方法。然而,很少有工作将这两种方法结合在一起。

[0005] 虽然表情识别在人机交互、安全驾驶、医疗诊断等领域已经有了广泛的应用,但是受个体差异、面部遮挡和面部姿势变化的影响,表情识别的精度依然有待于提高。由于人脸表情的变化是一个动态的过程,单一特征的提取具有局限性,导致特征提取不充分,识别准确率不理想。

发明内容

[0006] 为了解决现有技术的不足,本申请提供了基于时空多特征融合的动态表情识别方法及系统;该方法将时间特征、纹理特征和几何特征相结合,采用集成学习的相对多数投票法,将特征在决策层进行融合,解决单一特征描述子具有局限性的问题,从而提高表情识别准确率。

[0007] 第一方面,本申请提供了基于时空多特征融合的动态表情识别方法;

[0008] 基于时空多特征融合的动态表情识别方法,包括:

[0009] 获取待识别的人脸表情图像序列;

[0010] 对待识别的人脸表情图像序列进行预处理;

[0011] 对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采

集纹理特征和几何特征；

[0012] 对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;

[0013] 对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。

[0014] 第二方面,本申请提供了基于时空多特征融合的动态表情识别系统;

[0015] 基于时空多特征融合的动态表情识别系统,包括:

[0016] 获取模块,其被配置为:获取待识别的人脸表情图像序列;

[0017] 预处理模块,其被配置为:对待识别的人脸表情图像序列进行预处理;

[0018] 特征提取模块,其被配置为:对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;

[0019] 识别模块,其被配置为:对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;

[0020] 输出模块,其被配置为:对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。

[0021] 第三方面,本申请还提供了一种电子设备,包括:一个或多个处理器、一个或多个存储器、以及一个或多个计算机程序;其中,处理器与存储器连接,上述一个或多个计算机程序被存储在存储器中,当电子设备运行时,该处理器执行该存储器存储的一个或多个计算机程序,以使电子设备执行上述第一方面所述的方法。

[0022] 第四方面,本申请还提供了一种计算机可读存储介质,用于存储计算机指令,所述计算机指令被处理器执行时,完成第一方面所述的方法。

[0023] 第五方面,本申请还提供了一种计算机程序(产品),包括计算机程序,所述计算机程序当在一个或多个处理器上运行的时候用于实现前述第一方面任意一项的方法。

[0024] 与现有技术相比,本申请的有益效果是:

[0025] (1) 本发明方法对LBP-TOP算子进行改进,增强了特征的表达能力和特征提取速率。

[0026] (2) 本发明提出了一种新的几何特征表达方式,直接将归一化后的人脸关键点作为几何特征,有效地降低特征维度。

[0027] (3) 本发明在决策层进行特征融合,同时考虑到几何外观特征以及动态静态特征,显著地提高表情识别准确率。

[0028] 本发明附加方面的优点将在下面的描述中部分给出,部分将从下面的描述中变得明显,或通过本发明的实践了解到。

附图说明

[0029] 构成本申请的一部分的说明书附图用来提供对本申请的进一步理解,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。

[0030] 图1为本发明的系统流程图;

[0031] 图2为本发明采用SDM(Supervised Descent Method)算法检测到的人脸49个关键点位置信息示意图。

[0032] 图3为本发明对原始图像进行预处理的示意图;

- [0033] 图4为本发明提出的得到改进的LBP-TOP统计直方图特征的过程示意图；
- [0034] 图5为改进的LBP-TOP算子的识别准确率与表情序列所分块数以及在分块过程中是否采用70%重叠率的变化关系示意图；
- [0035] 图6为实施例1中改进的LBP-TOP算子与原始LBP-TOP算子以及其他各个分量识别准确率的比较示意图；
- [0036] 图7(a)-图7(f)为同一人的6种基本表情关键点示意图；
- [0037] 图7(g)-图7(l)为不同女人的6种基本表情关键点示意图；
- [0038] 图7(m)-图7(r)为不同男人的6种基本表情关键点示意图。

具体实施方式

[0039] 应该指出,以下详细说明都是示例性的,旨在对本申请提供进一步的说明。除非另有指明,本文使用的所有技术和科学术语具有与本申请所属技术领域的普通技术人员通常理解的相同含义。

[0040] 需要注意的是,这里所使用的术语仅是为了描述具体实施方式,而非意图限制根据本申请的示例性实施方式。如在这里所使用的,除非上下文另外明确指出,否则单数形式也意图包括复数形式,此外,还应当理解的是,术语“包括”和“具有”以及他们的任何变形,意图在于覆盖不排他的包含,例如,包含了一系列步骤或单元的过程、方法、系统、产品或设备不必限于清楚地列出的那些步骤或单元,而是可包括没有清楚地列出的或对于这些过程、方法、产品或设备固有的其它步骤或单元。

[0041] 在不冲突的情况下,本发明中的实施例及实施例中的特征可以相互组合。

[0042] 术语解释:

[0043] LBP-TOP是Local Binary Pattern from Three Orthogonal Planes的缩写,中文含义是三个正交平面上的局部二值模式。

[0044] 实施例一

[0045] 本实施例提供了基于时空多特征融合的动态表情识别方法;

[0046] 如图1所示,基于时空多特征融合的动态表情识别方法,包括:

[0047] S101:获取待识别的人脸表情图像序列;

[0048] S102:对待识别的人脸表情图像序列进行预处理;

[0049] S103:对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;

[0050] S104:对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;

[0051] S105:对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。

[0052] 作为一个或多个实施例,所述获取待识别的人脸表情图像序列,通过摄像头进行采集。

[0053] 示例性的,本实施例采用扩展的Cohn-Kanade (CK+) 数据集进行实验,使用该数据集中包含6种基本表情的309个表情序列。

[0054] 作为一个或多个实施例,所述S102:对待识别的人脸表情图像序列进行预处理;具体包括:

[0055] S1021:对待识别的人脸表情图像序列中的所有帧人脸表情图像,均进行灰度归一化处理;

[0056] S1022:对灰度归一化处理后的脸表情图像,进行裁剪;

[0057] S1023:对裁剪得到的人脸表情图像进行几何归一化处理。

[0058] 示例性的,所述S1021:对待识别的人脸表情图像序列中的所有帧人脸表情图像,均进行灰度归一化处理;具体包括:

[0059] 采用如下公式(1)对人脸表情图像进行灰度归一化:

$$[0060] \quad I_{\text{gray}} = I_B \times 0.114 + I_G \times 0.587 + I_R \times 0.299, \quad (1)$$

[0061] 其中, I_B 、 I_G 、 I_R 分别代表图像的蓝、绿、红三个颜色通道。

[0062] 示例性的,所述S1022:对灰度归一化处理后的脸表情图像,进行裁剪;具体包括:

[0063] 对灰度归一化后的图像采用SDM算法检测到人脸49个关键点,位置信息如图2所示,确定两只眼睛瞳孔的坐标,这两只眼睛瞳孔之间的距离为d,以两只眼睛瞳孔距离的中点为中心点,将动态人脸图像序列中的每帧人脸图像进行裁剪,裁剪后图片的宽为 $1.2 \times d$,高为 $2.25 \times d$,如图3所示。

[0064] 示例性的,所述S1023:对裁剪得到的人脸表情图像进行几何归一化处理;具体包括:

[0065] 最后将裁剪后的人脸表情图像几何归一化为 $M \times N$ 。对于改进的LBP-TOP特征,M取256,N取128;对于Gabor特征,为了降低维度,M取112,N取96。

[0066] 作为一个或多个实施例,所述S103:对预处理后的人脸表情图像序列,提取时空特征;其中,时空特征是指改进的LBP-TOP直方图特征;时空特征的提取步骤具体包括:

[0067] S1031:对于一个 $X \times Y \times T$ 的图像序列,分别计算图像序列XT平面、YT平面以及峰值帧图像XY平面的LBP直方图向量;

[0068] S1032:将表情序列XT和YT平面的LBP直方图向量与XY平面表情峰值帧的LBP直方图向量级联,从而得到改进的LBP-TOP直方图特征。

[0069] 进一步地,所述峰值帧图像的确定方式为:

[0070] (1) 对于一个 $X \times Y \times T$ 的图像序列,将每一帧灰度图像转化为二维矩阵;

[0071] (2) 将一个图像序列中的所有二维矩阵按顺序连接为一个三维矩阵;

[0072] (3) 在三维矩阵中,用N来控制滑动窗口的长度,以每帧的前后 $-N/2$ 到 $N/2$ 为一个区间,计算每一个区间的频率;

[0073] (4) 选取频率最高的区间,此区间的中间帧即为峰值帧图像。

[0074] 示例性的,所述S1031:对于一个 $X \times Y \times T$ 的图像序列,分别计算图像序列XT平面、YT平面以及峰值帧图像XY平面的LBP直方图向量;具体步骤包括:

[0075] 对于图像序列,用 g_c 表示表示中心像素(x_c, y_c, t_c)的灰度值, g_p 表示第p个邻域像素的灰度值,P为邻域像素的个数,R决定邻域半径的大小。局部纹理特征LBP(Local Binary Pattern)的计算方式如下:

$$[0076] \quad LBP_{P,R}(x_c, y_c, t_c) = \sum_{i=0}^{P-1} s(g_p, g_c) 2^p, \quad (2)$$

[0077] s是二值函数,定义如下:

$$[0078] \quad s(x, y) = \begin{cases} 1 & \text{如果 } x - y \geq 0 \\ 0 & \text{其它} \end{cases}, \quad (3)$$

[0079] 将LBP特征的统计直方图作为特征向量从而将全局信息考虑进去。对于一个 $X \times Y \times T$ 的图像序列,分别计算图像序列XT平面、YT平面以及峰值帧图像XY平面的LBP直方图向量:

$$[0080] \quad H_{XY} = I(LBP_{P,R}(x_c, y_c, t_c) = i) \quad (4),$$

$$[0081] \quad H_{XT_{seq}} = \sum_1^Y I\{LBP_{P,R}(x_c, y_c, t_c) = i\} \quad (5),$$

$$[0082] \quad H_{YT_{seq}} = \sum_1^X I\{LBP_{P,R}(x_c, y_c, t_c) = i\} \quad (6),$$

[0083] I 是二值函数,定义如下:

$$[0084] \quad I(A) = \begin{cases} 1 & \text{如果 } A \text{ 为真} \\ 0 & \text{如果 } A \text{ 为假} \end{cases} \quad (7),$$

[0085] i 表示统计直方图的模式,采用LBP均值模式,因此这里 $i \in [0, 58]$ 。

[0086] 示例性的,所述S1032:将表情序列XT和YT平面的LBP直方图向量与XY平面表情峰值帧的LBP直方图向量级联,从而得到改进的LBP-TOP直方图特征;具体步骤包括:

[0087] 将上述三部分的直方图特征分别进行归一化后,级联成一个特征向量。 H 即为人脸表情序列进行改进的LBP-TOP算子计算后得到的特征向量。

$$[0088] \quad H = \{H_{XY}, H_{XT_{seq}}, H_{YT_{seq}}\} \quad (8)。$$

[0089] 应理解的,对得到的归一化后的动态人脸图像序列利用改进的LBP-TOP算子提取时空特征。人脸表情序列是一个三维立体图像,因此可以看作多个二维平面在某个维度上的堆叠。XT和YT平面的图像有更多时间维度信息,并且更好地反应面部肌肉的运动方向,而XY平面的图像只有在表情峰值时才具有最大的表达能力。本发明将原始LBP-TOP算子的XY平面的LBP直方图特征舍弃,保留表情序列XT和YT平面的LBP直方图特征,将其与XY平面表情峰值帧的LBP直方图特征级联,从而得到改进的LBP-TOP算子。该算子不仅保留了原始LBP-TOP在时间维度上的时空纹理特征,而且利用了峰值帧图像表示性强的优点。

[0090] 图4展示了对于一个局部人脸表情序列块,每个中心像素都在三个平面上,改进的LBP-TOP特征直方图就是分别计算峰值帧图像XY平面的LBP统计直方图特征级联XT和YT平面的LBP统计直方图特征。

[0091] 改进的LBP-TOP特征识别准确率与表情序列所分块数以及在分块过程中是否采用70%的重叠率有较强的关系,如图5所示。图中的“ $n_1 n_2$ ”形式,“ n_1 ”取“0”表示XY平面峰值帧的直方图特征; n_1 取“1”表示XT和YT平面表情序列的直方图特征;“ n_2 ”取“0”表示不采用70%的重叠率,取“1”表示采用。例如,“00”表示XY平面峰值帧特征及XT和YT平面表情序列的特征均不采用70%的重叠率。根据实验结果,本发明将裁剪后的表情峰值帧图像采用70%的重叠率进行 8×4 的分块,对分块后的图像提取LBP直方图特征;将裁剪后的动态人脸图像序列进行 8×4 的分块,不采用70%的重叠率,对分块后的图像序列提取XT和YT平面的LBP-TOP直方图特征。在选择上述参数的情况下,改进的LBP-TOP算子识别准确率最高。

[0092] 从图6可以看出,在相同条件下,改进的LBP-TOP算子较原始LBP-TOP算子以及其他

分量,识别准确率有较大提升。

[0093] 此外,本实施例还比较了原始LBP-TOP算子和改进的LBP-TOP算子的特征提取速度。实验表明,特征提取时间随表情序列长度和分块数而变化。在表情序列长度(19)和块大小(8×4)相同的条件下,原始LBP-TOP算子的计算时间为3.91s,而改进的LBP-TOP算子的计算时间为3.61s,并且当表情序列长度为39时,原始LBP-TOP算子和改进的LBP-TOP算子的计算时间分别为7.54s和6.36s。随着序列长度和块数的增加,改进的LBP-TOP算子的时间优势更加明显。

[0094] 作为一个或多个实施例,所述S103:对表情峰值关键帧图像,分别采集纹理特征和几何特征;其中,纹理特征是指Gabor特征;纹理特征的提取步骤具体包括:

[0095] 采用5个尺度8个方向的Gabor滤波器,对归一化后的表情峰值帧图像提取Gabor幅值特征作为纹理特征。

[0096] 示例性的,Gabor滤波器的定义如下:

$$[0097] \quad G_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right) \times [\exp(ik_{u,v} \cdot z) - \exp(-\frac{\sigma^2}{2})] \quad (9),$$

[0098] 其中, $z=(x,y)$,表示图像的像素点, $k_{u,v}=k_v e^{i\varphi_u}$, $k_v=2^{-\frac{v+2}{2}}\pi$, $\varphi_u=u\frac{\pi}{K}$, u 表示Gabor滤波器的方向, $u \in (0, \dots, 7)$, v 表示Gabor滤波器的尺度, $v \in (0, \dots, 4)$, K 表示总的方向数。因此,5个尺度8个方向的Gabor滤波可以通过 $k_{u,v}$ 来描述。其中, k_v 分别为 $\frac{\pi}{2}$ 、 $\frac{\sqrt{2}\pi}{4}$ 、 $\frac{\pi}{4}$ 、 $\frac{\sqrt{2}\pi}{8}$ 、 $\frac{\pi}{8}$;对应每一个 k_v , φ_u 分别为 0 、 $\frac{\pi}{8}$ 、 $\frac{\pi}{4}$ 、 $\frac{3\pi}{8}$ 、 $\frac{\pi}{2}$ 、 $\frac{5\pi}{8}$ 、 $\frac{3\pi}{4}$ 、 $\frac{7\pi}{8}$ 。

[0099] 图像的Gabor表示是图像 $I(z)$ 与二维Gabor滤波器的卷积:

$$[0100] \quad F_{u,v}(z) = I * G_{u,v}(z) \quad (10),$$

[0101] $F_{u,v}(z)$ 是复数,由于Gabor的幅值信息相对相位信息稳定,因此,这里采用 $F_{u,v}(z)$ 的幅值作为图像的Gabor纹理特征。

[0102] 作为一个或多个实施例,所述S103:对表情峰值关键帧图像,分别采集纹理特征和几何特征;几何特征的提取步骤包括:

[0103] 对表情的峰值帧图像,采用SDM算法检测到人脸49个关键点,将X轴和Y轴的坐标分别进行均值为0,方差为1的标准化操作。

[0104] 用 (x_i, y_i) 表示第 i 个人脸关键点的坐标,情感 e 的几何特征向量为:

$$[0105] \quad V_e = [x_1, y_1, x_2, y_2, \dots, x_n, y_n], n=49 \quad (11),$$

[0106] 将这98维特征向量作为人脸表情的几何特征。

[0107] 作为一个或多个实施例,所述S104:对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;具体包括:

[0108] 基于训练后的第一基分类器,对时空特征进行表情识别,得到第一识别结果;

[0109] 基于训练后的第二基分类器,对纹理特征进行表情识别,得到第二识别结果;

[0110] 基于训练后的第三基分类器,对几何特征进行识别,得到第三识别结果。

[0111] 进一步地,所述训练后的第一基分类器,其训练步骤包括:

[0112] 构建基分类器;构建第一训练集,所述第一训练集为已知表情识别结果的时空特征;

[0113] 将第一训练集输入到基分类器中,输出第一分类结果。

[0114] 同样的,第二基分类器与第三基分类器的训练步骤与第一基分类器的训练步骤是一样的,不一样的仅仅是训练集不一样;第二基分类器的训练集为已知表情识别结果的纹理特征;第三基分类器的训练集为已知表情识别结果的几何特征。将得到的时空特征、Gabor纹理特征、几何特征分别使用SVM(Support Vector Machine)的多核学习方法得到三个基分类器。

[0115] 作为一个或多个实施例,所述S105:对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果;具体包括:

[0116] 通过集成学习的多数投票法得到样本的分类结果。对于每一个样本 x ,多数投票法的计算方式如下:

$$[0117] \quad H(x) = c_k, \quad k = \operatorname{argmax}_j \{ \sum_{t=1}^T h_t^j(x) \}, \quad j \in \{1,2,3,4,5,6\} \quad (12),$$

[0118] T 是基分类器的个数, $h_t^j(x) \in \{0,1\}$ 代表分类标记,若 h_t 将样本 x 预测为 c_j 类,则 $h_t^j(x)$ 的值为1,否则为0。对样本进行分类时,将票数最多的类别作为最终的分类结果。

[0119] 采用留一法进行验证。对于 N 个样本的数据,取1个样本作为测试集,其它 $N-1$ 个样本作为训练集,将 N 次测试的平均值作为最终结果。

[0120] 基于时空多特征融合的动态表情识别方法,是一种将几何-外观特征和动态-静态特征相结合的方法。具体来说,采用改进的LBP-TOP算子作为时空特征,Gabor幅值特征作为纹理特征,人脸关键点作为几何特征。

[0121] 图7(a)-图7(f)为同一人的6种基本表情关键点示意图;图7(g)-图7(l)为不同女人的6种基本表情关键点示意图;图7(m)-图7(r)为不同男人的6种基本表情关键点示意图,不管是同一个人,不同的女人或不同的男人,同一表情的坐标表示是非常相似的。另外,在嘴巴、眼睛和眉毛这些位置,不同表情的坐标表示存在较大差异。这表明人脸关键点可以去除人脸图像的共同底层结构,消除脸型影响,有效地提取表情的形状属性。

[0122] 采用留一法进行验证。对于 N 个样本的数据,取1个样本作为测试集,其它 $N-1$ 个样本作为训练集,将 N 次测试的平均值作为最终结果。

[0123] 表1、在CK+数据集上使用改进的LBP-TOP算子的表情识别准确率

[0124]

表情	生气	厌恶	害怕	高兴	悲伤	惊讶	准确率(%)
生气	42	0	0	0	3	0	93.3
厌恶	1	58	0	0	0	0	98.3
害怕	0	0	17	1	3	4	68
高兴	0	0	0	69	0	0	100
悲伤	4	0	0	0	22	2	78.6
惊讶	1	0	0	0	0	82	98.8
总体							93.85

[0125]

表2、在CK+数据集上使用Gabor算子的表情识别准确率

[0126]

表情	生气	厌恶	害怕	高兴	悲伤	惊讶	准确率(%)
生气	40	1	0	0	4	0	88.9
厌恶	2	57	0	0	0	0	96.6
害怕	1	0	22	0	0	2	88
高兴	0	0	0	69	0	0	100
悲伤	1	0	1	0	26	0	92.9

[0127]

惊讶	0	0	1	0	0	82	98.8
总体							95.79

[0128]

表3、在CK+数据集上使用几何关键点特征的表情识别准确率

	表情	生气	厌恶	害怕	高兴	悲伤	惊讶	准确率(%)
[0129]	生气	42	3	0	0	0	0	93.3
	厌恶	2	56	0	0	1	0	94.9
	害怕	0	0	19	5	0	1	76
	高兴	0	0	0	69	0	0	100
	悲伤	6	0	0	0	22	0	78.6
	惊讶	0	0	0	0	0	83	100
	总体							94.17

[0130] 表4、在CK+数据集上进行决策层特征融合后的表情识别准确率

	表情	生气	厌恶	害怕	高兴	悲伤	惊讶	准确率(%)
[0131]	生气	45	0	0	0	0	0	100
	厌恶	0	58	0	0	1	0	98.3
	害怕	0	0	23	1	0	1	92
	高兴	0	0	0	69	0	0	100
	悲伤	3	0	0	0	25	0	89.3
	惊讶	0	0	0	0	0	83	100
	总体							98.06

[0132] 从表1、表2、表3和表4可以看出,在识别某种特定表情时,总会有一种特征提取算子的性能优于其他。例如,改进的LBP-TOP算子能更好地识别厌恶;Gabor特征在识别悲伤和恐惧方面具有更大的优势;几何特征对惊讶的识别率可以达到100%。因此,在决策级融合时,由于各特征提取算子的优势表情不同,可以有效提高各类表情的识别准确率。

[0133] 实施例二

[0134] 本实施例提供了基于时空多特征融合的动态表情识别系统;

[0135] 基于时空多特征融合的动态表情识别系统,包括:

[0136] 获取模块,其被配置为:获取待识别的人脸表情图像序列;

[0137] 预处理模块,其被配置为:对待识别的人脸表情图像序列进行预处理;

[0138] 特征提取模块,其被配置为:对预处理后的人脸表情图像序列,提取时空特征;对表情峰值关键帧图像,分别采集纹理特征和几何特征;

[0139] 识别模块,其被配置为:对时空特征进行表情识别,得到第一识别结果,对纹理特征进行表情识别,得到第二识别结果;对几何特征进行识别,得到第三识别结果;

[0140] 输出模块,其被配置为:对第一、第二和第三识别结果进行多数投票,得到最终的表情识别结果。

[0141] 此处需要说明的是,上述获取模块、预处理模块、特征提取模块、识别模块和输出模块对应于实施例一中的步骤S101至S105,上述模块与对应的步骤所实现的示例和应用场景相同,但不限于上述实施例一所公开的内容。需要说明的是,上述模块作为系统的一部分可以在诸如一组计算机可执行指令的计算机系统中执行。

[0142] 上述实施例中对各个实施例的描述各有侧重,某个实施例中未详述的部分可以参见其他实施例的相关描述。

[0143] 所提出的系统,可以通过其他方式实现。例如以上所描述的系统实施例仅仅是示意性的,例如上述模块的划分,仅仅为一种逻辑功能划分,实际实现时,可以有另外的划分方式,例如多个模块可以结合或者可以集成到另外一个系统,或一些特征可以忽略,或不执行。

[0144] 实施例三

[0145] 本实施例还提供了一种电子设备,包括:一个或多个处理器、一个或多个存储器、以及一个或多个计算机程序;其中,处理器与存储器连接,上述一个或多个计算机程序被存储在存储器中,当电子设备运行时,该处理器执行该存储器存储的一个或多个计算机程序,以使电子设备执行上述实施例一所述的方法。

[0146] 应理解,本实施例中,处理器可以是中央处理单元CPU,处理器还可以是其他通用处理器、数字信号处理器DSP、专用集成电路ASIC,现成可编程门阵列FPGA或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件等。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。

[0147] 存储器可以包括只读存储器和随机存取存储器,并向处理器提供指令和数据、存储器的一部分还可以包括非易失性随机存储器。例如,存储器还可以存储设备类型的信息。

[0148] 在实现过程中,上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。

[0149] 实施例一中的方法可以直接体现为硬件处理器执行完成,或者用处理器中的硬件及软件模块组合执行完成。软件模块可以位于随机存储器、闪存、只读存储器、可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器,处理器读取存储器中的信息,结合其硬件完成上述方法的步骤。为避免重复,这里不再详细描述。

[0150] 本领域普通技术人员可以意识到,结合本实施例描述的各示例的单元及算法步骤,能够以电子硬件或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行,取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能,但是这种实现不应认为超出本申请的范围。

[0151] 实施例四

[0152] 本实施例还提供了一种计算机可读存储介质,用于存储计算机指令,所述计算机

指令被处理器执行时,完成实施例一所述的方法。

[0153] 以上所述仅为本申请的优选实施例而已,并不用于限制本申请,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

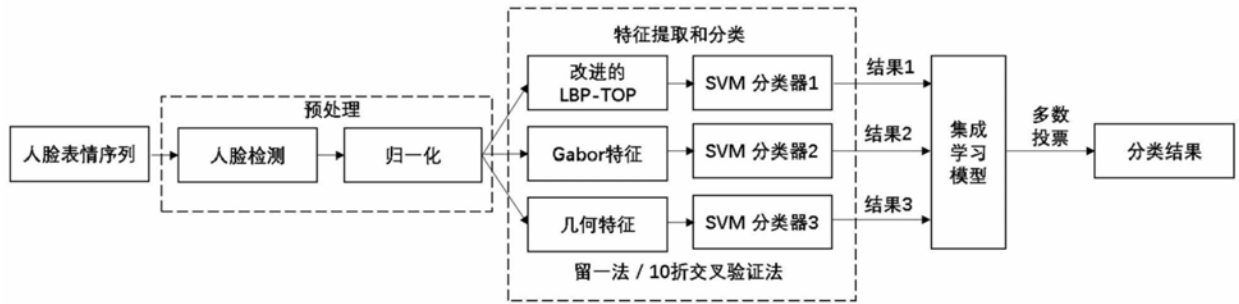


图1

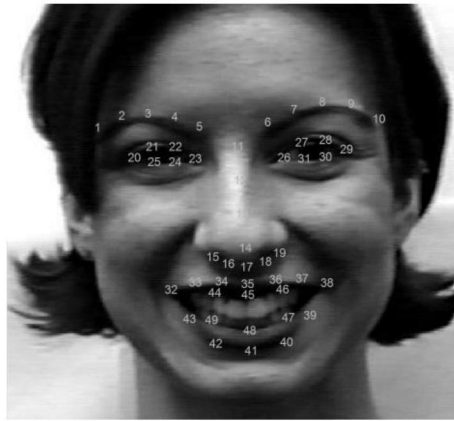


图2

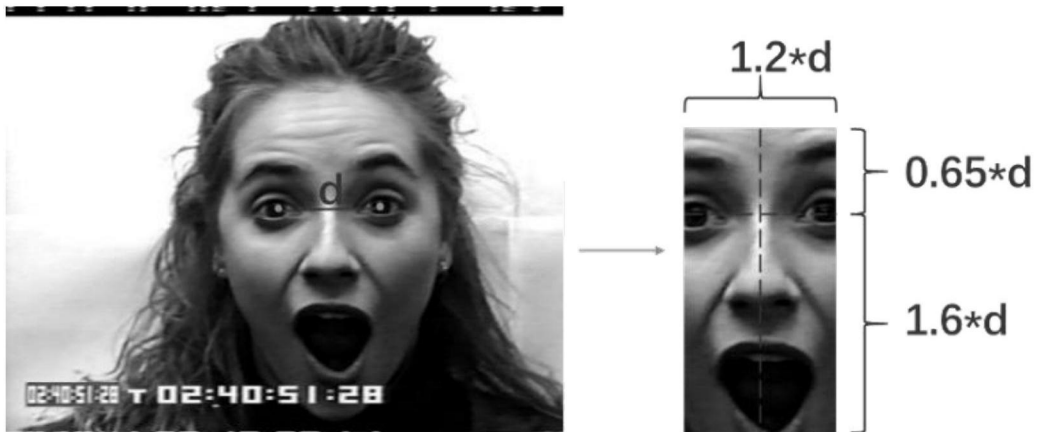


图3

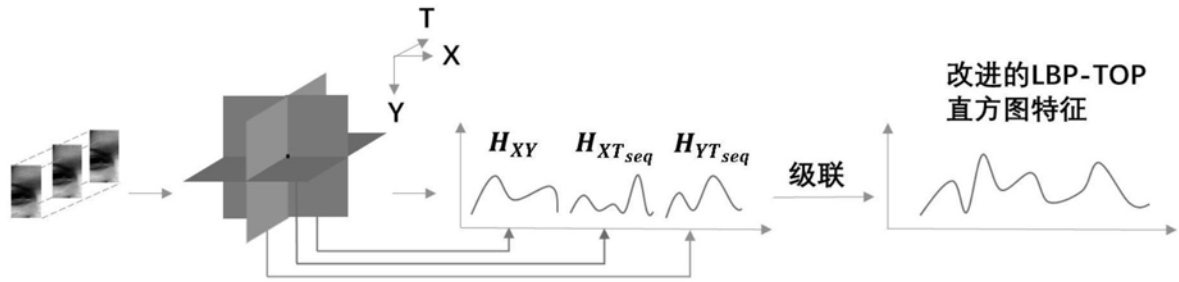


图4

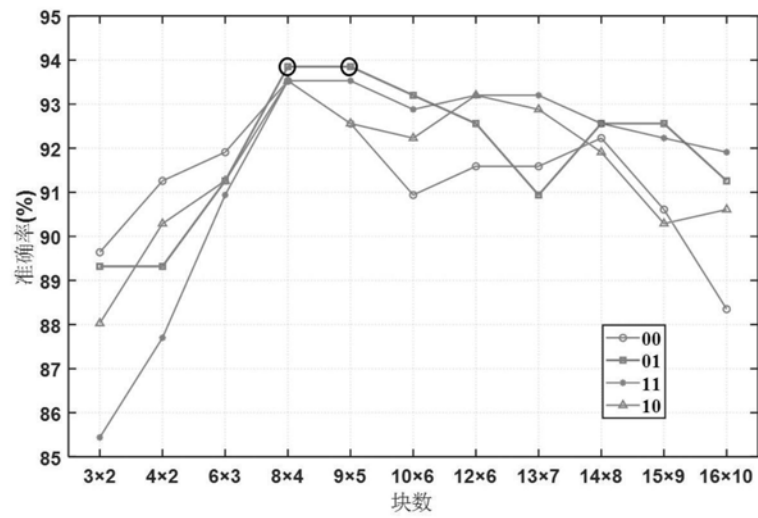


图5

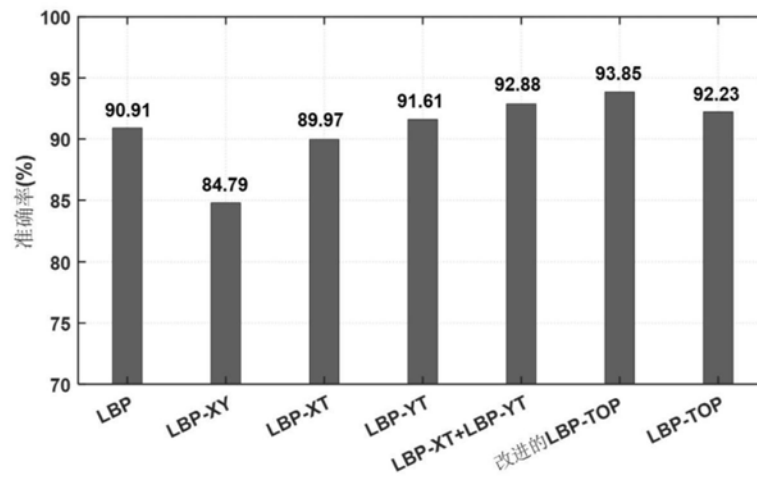


图6



图7 (a)



图7 (b)



图7 (c)



图7 (d)



图7 (e)

惊讶



图7 (f)

生气



图7 (g)

厌恶



图7 (h)

害怕



图7 (i)

高兴



图7 (j)

悲伤



图7 (k)

惊讶



图7 (l)

生气

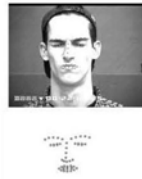


图7 (m)

厌恶

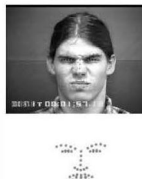


图7 (n)

害怕



图7 (o)

高兴



图7 (p)

悲伤



图7 (q)

惊讶



图7 (r)