**面部表情识别系统——文献综述**

1. **引言**

计算机科学技术的迅速发展对人类社会产生巨大的影响，人性化的人机交互模式成为人类和机器交流方式的发展趋势。情感计算作为人性化的人机交互的基础逐渐成为一个研究热点，而人脸表情识别作为情感计算的前提也受到了广泛的关注。人脸表情识别是当前计算机视觉、模式识别、人工智能等领域的热点研究课题。它是智能人机交互技术中的一个重要组成部分，不同领域的研究者提出了许多新方法。

在人类的社会交往中，信息的传递和情感的表达主要是通过自然语言和形体语言来完成的，形体语言指的是通过表情、手势等身体部位的动作来传达主体思想和情感的一种语言。而形体语言中最具代表的是面部表情。面部表情不但可以表现出主体的情绪状态，还能传达主体的思想意图。

时至今日，随着计算机技术、通信技术及人工智能技术的不断发展，工业自动化、智能化程度的不断提高，人机交互等智能产品也逐渐走近人们的日常生活中。如果可以利用计算机、手机等智能设备及时获取并正确分析理解人脸面部表情所代表的情感状态和心理信息，从而对用户情绪做出实时反应与交互，实现产品的智能化。从技术的创新和社会发展的方向来看这都极具意义，并意味着进一步推动人工智能时代的发展。

**关键字：面部表情识别、特征提取、表情分类、数据库、预处理**

1. **研究意义**

面部表情识别技术作为一门图像信号处理、心理学科、生理学科、模式识别和人工智能等相关领域的交叉学科，目前主要被应用于安全、医疗、交通、通信以及人工智能等相关领域。例如：通过人脸表情识别技术对机场、火车站、地铁站等公共安全场所进行全天候监测，可有效预测和预防公共场所的突发事件；通过人脸表情识别技术实时监测驾驶员情绪状态，可有效预防驾驶员出现疲劳驾驶、酒驾等状态并及时提醒，从源头上降低交通事故发生的可能性；通过人脸表情识别技术实时监测病患者的面部表情变化，不仅能判断出一些不能言语的病患者的情感状态，还能避免一些人为的致痛性操作给病患者带来的痛苦，并及时改善护理质量，特别是在新生儿领域，对新生儿的病情预测与治疗提供依据，可以发挥极大的医护作用。另外，在多模式人机交互界面(MMHCI)中,表情与声音、视线、体态等结合起来，可以获得更高效、更人性化的人机交流。在如面部神经瘫痪诊断、人脸图像实时传输、人脸图像合成与动画、智能机器人、智能监控等多个领域,表情识别都有着广泛的应用。

我们目前已经迈入人工智能时代，面部表情识别的研究将在未来产生举足轻重的意义与作用。因此，我认为，对面部表情识别的研究与发展成为了每一个时代开拓者身上义不容辞的责任与义务。

1. **研究现状**

**3.1面部表情识别的难点**

1.建立人脸表情模型和情绪分类,并把它们同面部特征与表情的变化联系起来。

2.人脸表情丰富(包括表情的各种倾向),人类对表情的控制能力,表情的变化细微而复杂,对表情变化特点的概括等诸多方面因素。

3.同种表情的表现程度不一定。

4.采集到图像数据中的光照与脸部旋转角度等,对识别结果的影响颇大。

5.数据来源方面,许多实验的数据的表情来自实验室中人工采集数据,这样的数据与人类自然流露的表情会有差异。

6.对人脸的熟悉程度:即指训练数据和测试数据之间的相关度,对识别结果也会有影响。

7.浅层特征模型,信息丢失制约识别准确性

8.应用场景复杂,受环境因素影响鲁棒性差

9.动态模型较少,序列信息缺乏有效利用等

**3.2识别系统概述：**人 脸 检 测 
与 足 位 
表 情 特 征 
提 取 
人 脸 表 情 
分 类 

从表情识别过程来看,表情识别可分为三部分:人脸图像的获取与预处理、表情特征提取和表情分类。建立一个 FER 系统,第 1 步需要对人脸进行检测与定位;第 2 步从人脸图像或图像序列中提取能够表征输入表情本质的信息,在提取特征数据的过程中,为了避免维数危机,可能还需要特征降维、特征分解等进一步处理;第3 步分析特征之间的关系,将输入的人脸表情分类到相应的类别。

**3.3人脸表情识别的框架流程图：**

****

**表情数据库或者摄像头采取**

**表情图像**

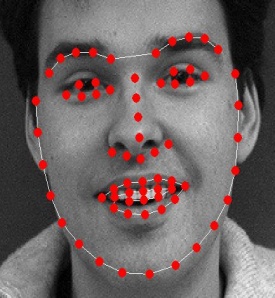
经摄像设备采集

****

**人脸识别**

**几何归一化和直方图均衡化**

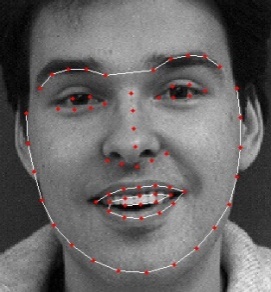
**图像预处理**

****

**ASM与AAM算法**

**PCA（Principal Component Analysis，主成分分析算法等进行特征降维**

**表情特征提取**

****

**Happy**

**BR神经网络、SVM等分类器**

**人脸表情分类**

**3.4人脸检测方法：**

**(1) 基于统计的人脸检测——是将人脸图像视为一个高维向量,将人脸检测问题转化为高维空间中分布信号的检测问题。**

**1. 样本学习**:将人脸检测视为从非人脸样本中识别人脸样本的模式分类问题,通过对人脸样本集和非人脸样本集进行学习以产生分类器。目前国际上普遍采用人工神经网络。

**2. 子空间方法**:将KL 变换引入了人脸检测,利用主元子空间(特征脸),而人脸检测利用的是次元子空间(特征脸空间的补空间)。用待检测区域在次元子空间上的投影能量,也即待检测区域到特征脸子空间的距离作为检测统计量,距离越小,表明越像人脸。子空间方法的特点在于简便易行,但由于没有利用反例样本信息,对与人脸类似的物体辨别能力不足。

**3. 模板法**:模板分为固定模板和变形模板。固定模板是求取测试样本与参考模板之间的某种度量,由阈值大小的定义来判断测试样本是否是人脸。它多用于粗检测和预处理过程。变形模板包含一些非固定的元素,加入了惩罚机制,以参数化或自适应的曲线和曲面来构成人脸模板。

**(2) 基于知识建模的人脸检测——是利用人的知识建立若干规则,从而将人脸检测问题转化为假设、验证问题。**

**(1) 人脸规则**:人脸规则是人脸遵循的一些几乎是普遍适用的空间相关性。它包括:①灰度分布规则。如五官的空间位置分布大致符合“三停五眼”等,人脸不同区域的明暗关系不变,眼睛的灰度总是比前额和颧骨低,鼻梁的灰度一般比两侧亮等。②轮廓规则。人脸的轮廓可以简单地看成一个近似椭圆,而人脸检测可以通过椭圆检测来完成。③运动规则。通常相对背景人总是在运动的,利用运动信息可以简单有效地将人从任意复杂背景中分割出来。

**(2) 颜色、纹理信息**:同一种族人的面部肤色在颜色空间中的分布相对比较集中,颜色信息在一定程度上可以将人脸从大部分背景区分开来。

**(3) 对称性**:人脸具有一定的轴对称性,各器官也具有一定的对称性。Reisfeld提出广义对称变换方法检测局部对称性强的点来进行人脸器官定位。

**3.5人脸表情图像特征提取**

如何对人脸表情特征进行有效描述,如何获取人脸表情关键的区分特征是人脸表情识别系统中最重要的步骤。表情特征的表示与提取就是将表情图像从一个高维特征向量变换为一个低维的、具有较大判别信息的向量,新的特征向量和原图像互为表示,损失信息尽可能少。一般来讲,新的特征需要具备以下特点:具有较低的数据维度,并能较好的体现人脸表情的本质特征,相比原图像尽可能减少损失信息;图像关键特征向量易于计算,方便把人脸表情图像快速转换为特征向量;新生成的人脸表情特征向量体现数据的关联关系和结构,方便进行分类器训练;方便发现潜在的人脸图像数据间关系,从而实现对人脸表情类别信息的理解。

1.**基于局部纹理的特征提取方法**

（1）**Gabor**

Gabor变换是一种局部频率信息描述方法。Gabor变换被广泛应用于视觉信息处理,主要**优点**是:(1)具有良好的边缘敏感特性,能够描述图像边缘方向选择和尺度选择;(2)具有良好的光照适应性,对光照变化不敏感。

(2)**LBP**

局部二值模式(Local Binary Pattems,LBP)是一种局部特征描述算子,具有很强的纹理信息描述能力。方便度量和提取图像局部纹理信息,且对光照具有不变性。由于其算法复杂度低,消耗内存小,原理简单,具有较强的纹理特征描述能力。LBP成功应用于人脸检测,唇语识别,表情检测,动态纹理等领域。LBP特征也被用于面部表情识别。

(3)**Haar**

Haar是一种局部特征描述工具,结合全局信息统计积分图,常用于进行人脸表示与识别。基本的Haar特征分为:边缘特征、线性特征和对角线特征三种类型,三种基本特征组合成特征模板。

2.**基于梯度特征的特征提取方法**

**1.HOG**

方向梯度直方图( Histogram of OrientedGradient,HOG)特征是一种常用于物体检测、对象跟踪的特征描述算子。基本思想是目标对的局部表象和形状可用梯度或边缘方向密度的分布来描述。

1. **SIFT**

尺度不变特征变换(Scale-invariant feature transform,SIFT)是基于物体关键点的梯度大小、梯度方向的一种局部描述算法。该算法具有尺度不变性,常用于图像处理领域。

**3.基于模板的特征提取方法**

基于模板的方法是一种基于统计思想的模型,主要包含主动形状模型(Active Shape Model,ASM)算法和主动外观模型(Active Appearance Model,AAM)算法。

**4.基于运动特征的提取方法**

**（1）图像差分法**

图像差分法是指利用图像序列像素差分来描述物体运动的一种方式。常用于进行运动目标检测,其主要的优点是:(1)算法建立在相邻时间帧图像之间,对运动变化敏感,具有很高的实时性;(2)算法基于差值计算,算法简单,运算量小。同时,缺点也很明显,主要是:(1)对环境噪声较为敏感,容易受到干扰;(2)阈值选择较为关键,人工确定阈值难以取得最佳效果。

**（2）光流法**

光流法是一种基于亮度模式的运动特征表示方法。光流法通常分为有全局光流场法和特征点光流场法。光流法主要优点是不仅描述了物体运动信息,同时也描述了物体结构信息。缺点是运算量大,耗时多,且实时性差。

**（3）特征点跟踪法**

特征点跟踪法是选择脸部的显著特征点来进行运动估计,为了防止在跟踪过程中丢失特征点,特征点一般都选取在脸部的永久特征上。这种方法只考虑感兴趣的特征点来分析表情,而不必理会背景等无关信息,减少了运算量,但由于只提取了部分特征点,而忽略了其他的一些脸部活动,可能会丢失有用信息。

**3.6表情识别方法分类：**

**(1) 基于模板的匹配方法**

为每一个要识别的表情建立一个模板,将待测表情与每种表情模板进行匹配,匹配度越高,则待测表情与某种表情越相似。对6种典型的基本人脸表情可以建立相应的模板,但对于识别非典型脸部表情比如混合表情时较困难,很难用有限的模板来代表无限多的人脸表情的情况。

**(2) 基于神经网络的方法**

神经网络方法的研究具有悠久的历史,它很适合于解决非线性问题。网络的输出节点对应6种基本表情和中性表情。但是,如果对自然表情类型而不是 6 中基本表情分类,那么神经网络方法的训练过程的工作量会非常大。另外,神经网络除了可以当作分类器外还可以作为特征提取器。

**(3) 基于概率模型的方法**

估计表情图像的参数分布模型,分别计算被测表情属于每个类的概率,取最大概率的类别为识别结果。隐马尔可夫模型是一个统计模型,具有十分健壮的数学结构,适用于动态过程时间序列建模并具有强大的时序模式分类能力,理论上可以处理任意长度的时序。

**(4) 基于支持向量机的方法**

支持向量机SV 是一种两类分类器,其基本思想是将训练样本经非线性变换映射到另一高维空间中,在变换后的高维空间内寻找一最优的分界面,在类间空隙最大的情况下将样本无错误或接近无错误分开。由于表情识别一般是多类别的分类问题,因此需要将 SVM 方法扩展到多类别的情形。主要有两种策略,一种是将多类别问题分解成一系列二类别问题来求解;另外一种是将包含一类不同样本的多类别问题分解成一个二类别问题,每次任选 C 类样本作为二类别问题的一类,剩下的 C-1 类样本作为另一类。

**3.7人脸表情数据库及预处理**

目前国内外现有的人脸表情数据库中的表情图像数量级比较小，无法满足深度网络对输入图像数量级的要求。所以，对人脸表情库通过一些方法来扩增是必要的。

**1.人脸表情数据库简介**

目前，国内外现有的人脸表情数据库主要分为两类，一类为静态图像组成的人脸表情数据库，如：JAFFE；另一类为视频或从视频帧提取的表情动态图像序列组成的人脸表情数据库，如：CK+、AFEW 6.0、FER2013、BHU。

**2.人脸表情数据库的预处理**

从数据库的实例样本图像中，可以发现，有些数据库并非统一的灰度图像，且有些图像还包含了拍摄日期等其他无助于识别分类的噪声信息，所以需要对人脸表情数据库进行以下几个方面的预处理。

**（1）        人脸检测**

从数据库的实例样本可以看到，表情图像还包含了拍摄日期干扰噪声，而且非人脸部分占据了表情图像的很大一部分，这会影响到后续实验中的表情特征提取过程，可能会引入非表情特征的干扰，不利于分类识别，所以，先对图像进行人脸检测处理。人脸检测旨在确定输入图像中存在的所有人脸的位置和大小。

**(2)        数据集扩增**

对原始图像稍作变换，得到新的图像来扩增数据集，实践证明，这种方法是简单且有效的。称这种扩增叫数据集扩增（Data Augmentation），通常有两种方法，一种 crop，一种基于原始图像的几何变换。

**(3)        归一化**

人脸图像之所以在训练前先进行归一化，是因为这样处理后的图像对光照强度、姿势等。

不同的成像条件具有一致性。人脸归一化包括几何归一化和灰度归一化，而几何归一化又包含尺度归一化、歪头矫正和扭脸矫正。

**3.8现阶段具体的人脸表情识别方法**

|  |  |  |  |
| --- | --- | --- | --- |
| 人脸表情识别方法 | 方法简单描述 | 优点 | 缺点 |
| 稀疏表示 | 用稀疏表示法对样本库进行描述，建立超完备子空间，重构并观察残差，最后通过稀疏系数进行分类 | 操作简单，可以做前期的基础实验，有一定的鲁棒性 | 描述对象必须要是稀疏的，降低了实际应用价值，对于样本要求也比较高 |
| Gabor变换 | 通过定义不同的核频率、带宽和方向对图像进行多分辨率分析，能有效提取不同方向不同细节程度的图像特征并相对稳定, 常与ANN 或SVM 分类器结合使用，提高表情识别的准确率 | 在频域和空间域都有较好的分辨能力，有明显的方向选择性和频率选择特性 | 作为低层次的特征，不易直接用于匹配和识别,识别准确率也不是很高，样本较少的条件下识别准确率也较低 |
| 主成分分析和线性判别 | 尽可能多的保留原始人脸表情图像中的信息,并允许分类器发现表情图像中相关特征,通过对整幅人脸表情图像进行变换,获取特征进行识别 | 具有较好的可重建性 缺点：可分性较差 | 外来因素的干扰（光照、角度、复杂背景等）将导致识别率下降 |
| 支持向量机 | 作为分类器做人脸识别，在表情识别时一般和Gabor滤波器一起使用作为分类器 | 在小样本下的识别效果较为理想，可以做实时性的表情识别 | 样本较大时，计算量和存储量都很大，识别器的学习也很复杂 |
| 光流法对运动特征提取 | 是将运动图像函数f (x,y,t)作为基本函数，根据图像强度守恒原理建立光流约束方程，通过求解约束方程，计算运动参数 | 反映了人脸表情变化的实际规律，受外界环境的影响较小，比如光照条件变化时，识别率不会有太大变化 | 识别模型和算法较复杂，计算量大 |
| 图像匹配法 | 通过使用弹性图匹配的方法将标记图和输入人脸图像进行匹配 | 允许人脸旋转，和能够实时处理 | 会受到其他部位特征的影响，如眼镜，头发等 |
| 隐马尔可夫模型 | 由观察的面部表情序列及模型去计算观察面部表情序列的概率，选用最佳准则来决定状态的转移;据观察的面部表情序列计算给定的模型参数 | 识别准确率较高，平均在97%以上 | 对前期的面部表情序列模型要求较高，这对表情识别算法的准确率影响也较大 |
| 其他方法如：矩阵分解法 | 以NMF为例，分解后的基图像矩阵和系数矩阵中的元素均是非负的。将表征人脸各部分的基图像进行线性组合从而表征整个表情图像。 | 需要的样本较少，在无遮挡时识别准确率90%以上 | 受外界环境影响较大，识别准确率在嘴巴受到遮挡时，准确率只有80%左右 |

1. **结论**

人工智能时代已经来临，为了使面部表情识别的应用更具有前瞻性和利用价值，在算法的改进中不能只局限于现有的已知模式下的数据库，而应该转变数据思维，在结构化、半结构化、非结构化的复杂数据源上以更具有实用价值的指导下设计新的算法或改进。

**五.附录**

**参考文献：**

【1】崔洁，面部表情识别的研究，硕士学位论文，西北工业大学，2006.3

【2】周晓彦，面部表情识别的研究，硕士学位论文，东南大学，2005.3

【3】张利伟，张航，张玉英，面部表情识别综述，中南大学信息科学与工程学院，湖南·长沙

【4】杨雨浓，基于深度学习的人脸表情识别方法的研究，博士学位论文，西北大学，2017.6

【5】何良华，人脸表情识别中若干关键技术的研究，博士学位论文，东南大学，2005.9

【6】陆慧聪，面部表情识别系统中表情特征提取与识别算法的研究，硕士学位论文，东南大学，2006.3

【7】刘小军，人脸识别技术研究，博士学位论文，中国科学院，2001.5

【8】赵明华，人脸检测和识别技术的研究，博士学位论文，四川大学，2006.10

【9】林妙真，基于深度学习的人脸识别研究，硕士学位论文，大连理工大学，2013.6

【10】黄雪琳，人脸的几何特征提取与查询，硕士学位论文，华侨大学，2004.4

【11】赵武峰，人脸识别中特征提取方法的研究，博士学位论文，浙江大学，2009.6

【12】李刚，基于特征脸法的人脸识别的研究，硕士学位论文，国防科学技术大学研究生院，2002.11