



中国科学院自动化研究所
INSTITUTE OF AUTOMATION
CHINESE ACADEMY OF SCIENCES

情感计算 -表情识别



授课教师：陶建华

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

背景和意义

- 人脸表情识别是指从图像或视频序列中分离出特定的表情状态，从而确定被识别对象的心理情绪
 - 面部表情是面部肌肉的一个或多个动作或状态的结果；这些运动能够表达个体对观察者的情绪状态
 - 在人与人的交流中，只有7%的信息是通过语言来传递，而通过面部表情传递的信息量却达到了55%



多个体表情识别



动态表情分析



人脸表情交互

背景和意义

■ 表情分类

- 表情：俗称宏表情，持续时间长
- 微表情：持续时间短，人难以察觉

■ 主要应用

- 辅助医疗、远程教育、疲劳驾驶、测谎、社交机器人



表情自动拍照



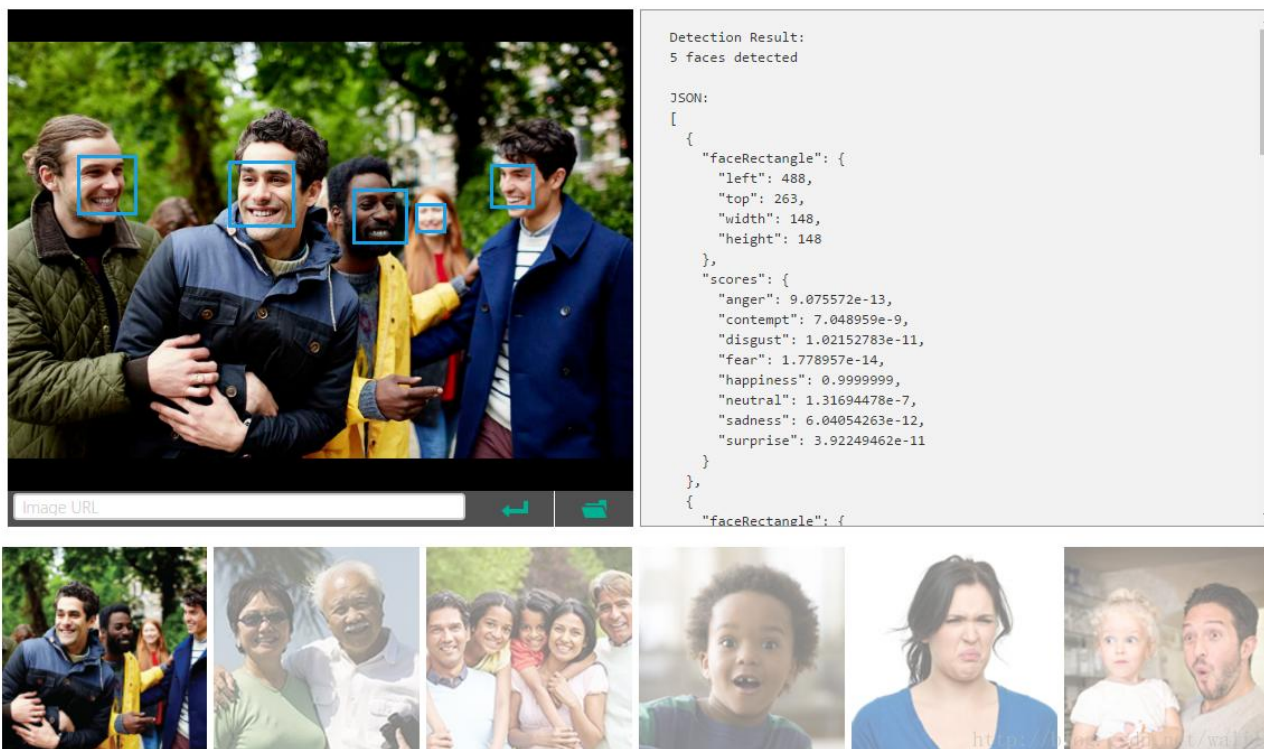
驾驶系统监控



微表情测谎

■ 表情识别的开源工具（API接口）

微软提供了表情识别的API接口，并通过JSON返回识别结果。



The screenshot displays the Microsoft Face API interface. On the left, a group photo of five people is shown with blue bounding boxes around their faces. Below the photo is an input field labeled "Image URL" and two buttons (a green left arrow and a green right arrow). On the right, the "Detection Result" is shown, indicating "5 faces detected". Below this, the JSON output is displayed, showing the bounding box and emotion scores for each face. The JSON structure is as follows:

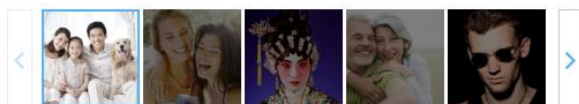
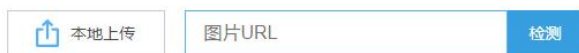
```
Detection Result:
5 faces detected

JSON:
[
  {
    "faceRectangle": {
      "left": 488,
      "top": 263,
      "width": 148,
      "height": 148
    },
    "scores": {
      "anger": 9.075572e-13,
      "contempt": 7.048959e-9,
      "disgust": 1.02152783e-11,
      "fear": 1.778957e-14,
      "happiness": 0.9999999,
      "neutral": 1.31694478e-7,
      "sadness": 6.04054263e-12,
      "surprise": 3.92249462e-11
    }
  },
  {
    "faceRectangle": {
```

Below the main interface, there are six smaller images showing various people with different expressions, illustrating the range of emotions the API can detect.

■ 表情识别的开源工具（API接口）

Face++提供了表情识别的API接口，并通过JSON返回识别结果。



▼ Response JSON

```
{
  "image_id": "8g2nrviNBnpFseprStfyA==",
  "request_id": "1492175342.288cb670-30c9-4a80-8895-4f5cf0f87beb",
  "time_used": 348,
  "faces": [
    {
      "landmark": {
        "mouth_upper_lip_left_contour2": {
          "y": 165,
          "x": 276
        },
        "mouth_upper_lip_top": {
          "y": 164,
          "x": 286
        },
        "mouth_upper_lip_left_contour1": {
          "y": 163,
          "x": 283
        },
        "left_eye_upper_left_quarter": {
          "y": 131,
          "x": 260
        },
        "left_eyebrow_lower_middle": {
          "y": 125,
          "x": 260
        },
        "mouth_upper_lip_left_contour3": {
          "y": 166,
          "x": 278
        },
        "left_eyebrow_lower_left_quarter": {
          "y": 126,
          "x": 254
        }
      }
    }
  ]
}
```

► 人脸关键点 & 人脸属性

<http://blog.csdn.net/waliik>

■ 表情识别的开源工具（API接口）

腾讯优图提供了表情识别的API接口，并通过JSON返回识别结果。



参数设置

人脸属性信息 ☒ 返回 ☐ 不返回

质量检测 ☒ 关闭 ☐ 开启

最大识别人脸数 5

识别结果

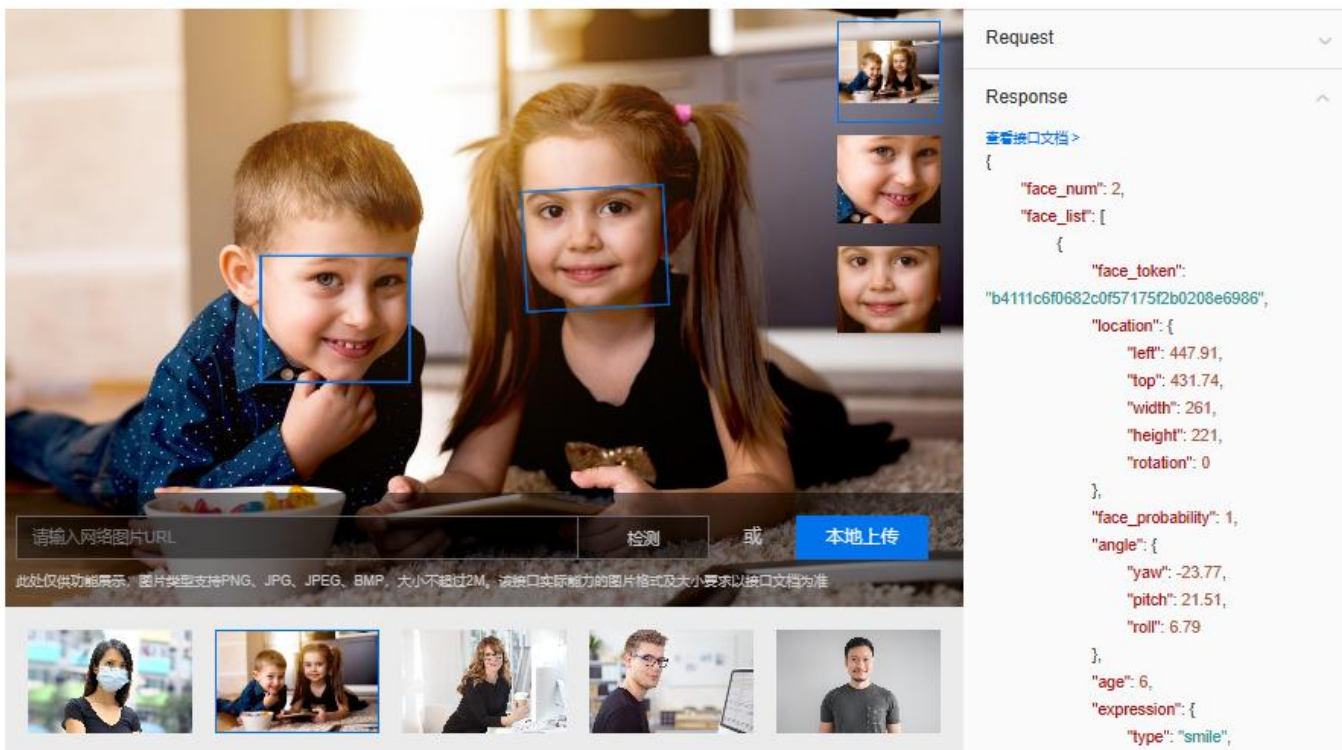
Request

Response

```
DetectFace
{
  "ImageWidth": 650,
  "ImageHeight": 650,
  "FaceInfos": [
    {
      "X": 60,
      "Y": 174,
      "Width": 151,
      "Height": 197,
      "FaceAttributesInfo": {
        "Expression": 91,
        "Glass": false,
        "Pitch": 12,
        "Roll": -13,
```


■ 表情识别的开源工具（API接口）

百度提供了表情识别的API接口，并通过JSON返回识别结果。



The screenshot displays the Baidu facial expression recognition API interface. On the left, a large image of two children is shown with blue bounding boxes around their faces. Below the image, there is a text input field labeled "请输入网络图片URL", a "检测" (Detect) button, and a "本地上传" (Upload from local) button. A small note below the input field states: "此处仅供功能展示，图片类型支持PNG、JPG、JPEG、BMP，大小不超过2M，该接口实际能力的图片格式及大小要求以接口文档为准". To the right of the main image, there are three smaller images showing different facial expressions. On the far right, a JSON response is displayed, showing the detected faces and their attributes.

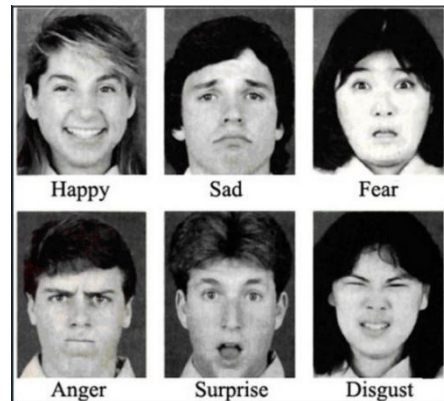
```
Request
Response
查看接口文档 >
{
  "face_num": 2,
  "face_list": [
    {
      "face_token":
      "b4111c6f0682c0f57175f2b0208e6986",
      "location": {
        "left": 447.91,
        "top": 431.74,
        "width": 261,
        "height": 221,
        "rotation": 0
      },
      "face_probability": 1,
      "angle": {
        "yaw": -23.77,
        "pitch": 21.51,
        "roll": 6.79
      },
      "age": 6,
      "expression": {
        "type": "smile",
```

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

研究现状与进展

- 1978年，Suwa等人提出了在图像序列中进行面部表情自动分析，对表情识别进行了初步尝试。
- 从20世纪90年代开始，由Mase和Pentland提出的光流法进行面部表情识别之后，自动面部表情识别进入了新的时期
- 近年来，基于深度学习方法的表情识别性能获得了极大的提升，并得到一系列应用



国外主要研究单位

■ 麻省理工学院

- 提出了情感计算的概念，以表情识别为研究核心

■ 卡内基梅隆大学

- 提出了表情符号的概念，通过有序的排列字符组成表情

■ 匹兹堡大学

- 情感分析实验室研究微表情识别和抑郁检测等领域

■ 帝国理工大学

- 基于表情识别技术研发了多款商业产品

国内主要研究单位

■ 国内

- 中科院自动化所
- 中科院计算所
- 清华大学
- 哈尔滨工业大学
- 中科院心理所
- 复旦大学
- 北京科技大学
- 北京航空航天大学

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

人脸表情数据库

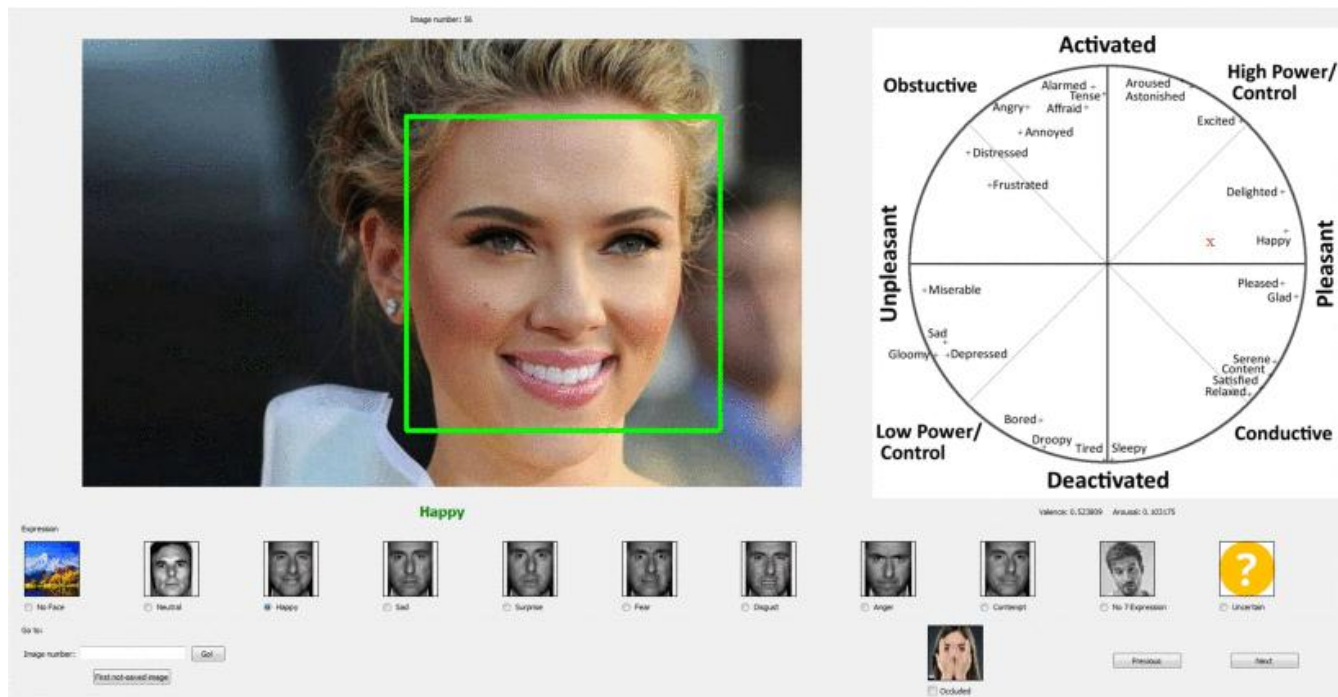
■ 表情数据库

- 建立丰富有效的人脸表情数据库对人脸表情识别技术的发展具有重要意义
- 人脸表情数据库大致分为三类
 - 基于基本情感
 - 基于AU编码的细微表情
 - 其他

人脸表情数据库

■ 野外人脸表情数据库

- 丹佛大学野外人脸表情数据库 (AffectNet) 包含来自互联网的1000000多个面部图像，大约一半的检索图像被手动注释，以确定是否存在七个离散的面部表情以及价效和唤醒的强度。



人脸表情数据库

■ 表情数据库

- 马里兰大学人脸表情数据库包含40名不同种族和文化背景被试者的六种基本情感表情的图像序列，每个图像序列持续约9s并包含一至三种表情；
- 荷兰代尔夫特理工大学MMI人脸表情数据库包含50名被试者的六种基本情感表情的500个静态图像和2000个视频；
- 德国慕尼黑理工大学FGNET数据库包含18名被试者的六种情感表情和中性表情的图像序列；

人脸表情数据库

■ 日本女性人脸表情数据库

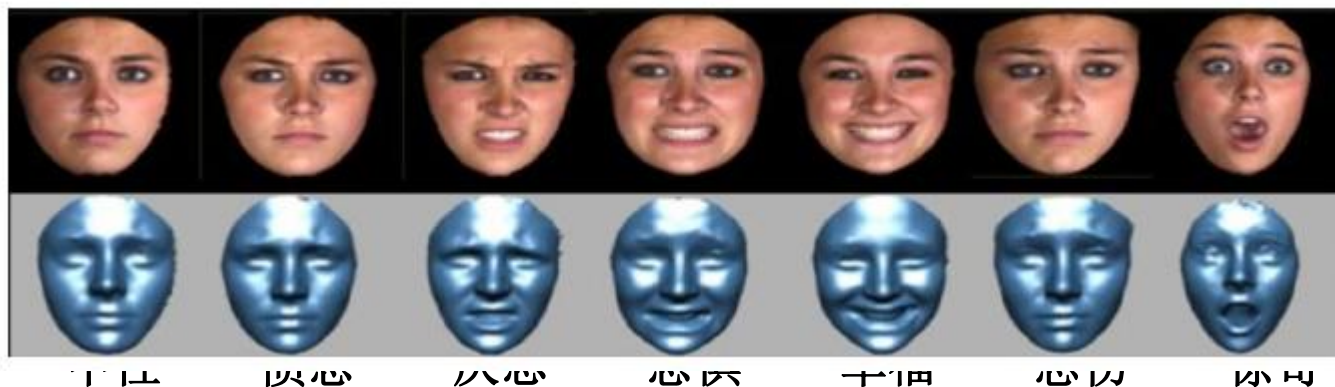
- 日本女性人脸表情数据库 (Japanese Female Facial Expression, JAFFE) 包含10名日本女性的六种情感表情和中性表情共213个图像。



人脸表情数据库

■ BU-3DFE人脸表情数据库

- 由于二维人脸表情数据不能有效用于较大姿势变化的情况，因此纽约州立大学宾汉姆顿分校建立了BU-3DFE人脸表情数据库，包含100名不同种族被试者的六种基本情感表情和中性表情的三维信息。



人脸表情数据库

■ ADSIP3-D动态人脸表情数据库

- 英国中央兰开夏大学建立的ADSIP3-D动态人脸表情数据库包含10名被试者的害怕、嫌恶、生气、高兴、惊讶、悲伤和疼痛七种表情，并分为轻微、一般和极度三种强度。



Happiness - Mild



Happiness - Normal



Happiness - Extreme



Sadness - Mild



Sadness - Normal

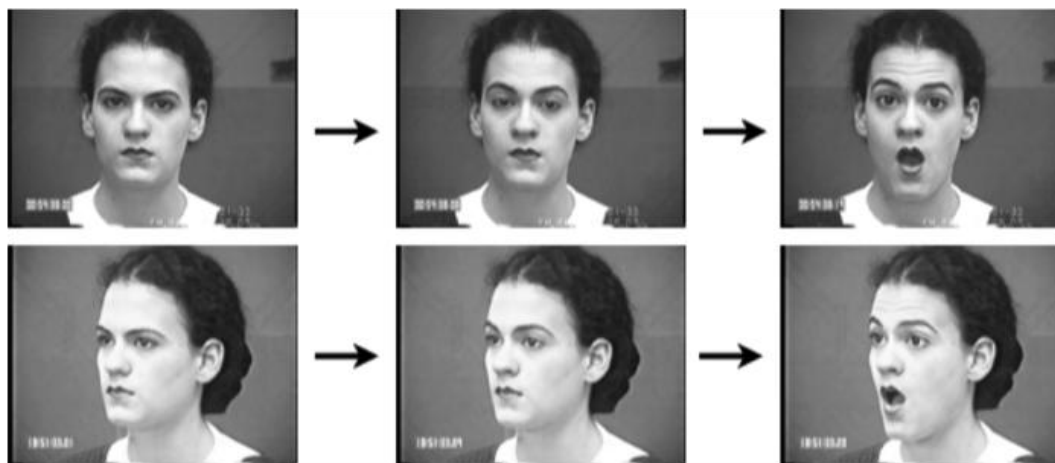


Sadness - Extreme

人脸表情数据库

■ CK人脸表情数据库

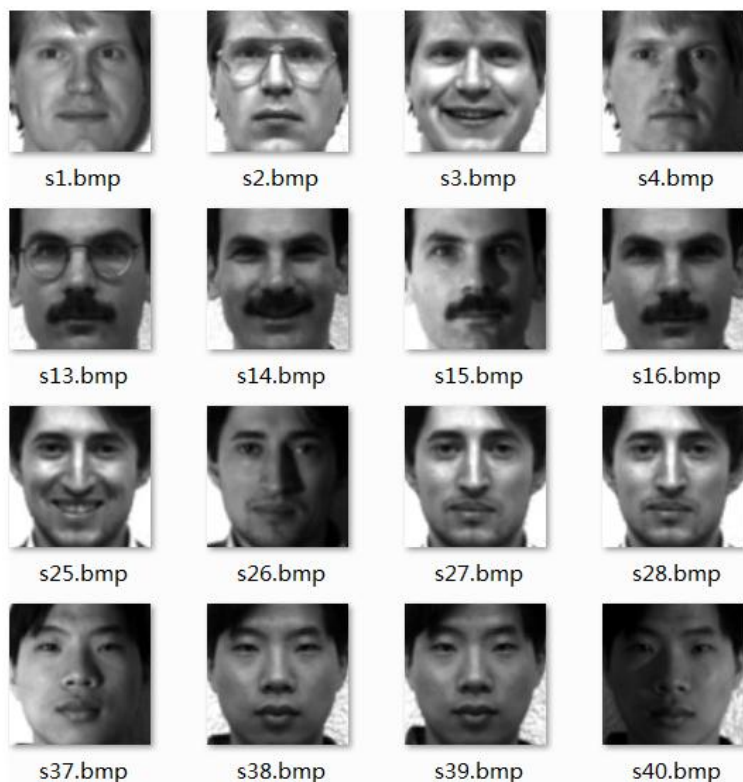
- 卡内基梅隆大学机器人研究所和心理学系共同建立的Cohn-Kanade (CK) AU-Coded人脸表情数据库包含单一运动单元和复合运动单元共23种表情，对210名18-50岁成年人的人脸表情进行了录制，发布了97名被试者的486段人脸表情图像序列。
- 2010年又在CK数据库的基础上建立了CK+ (Extended Cohn-Kanade) 数据库，图像序列增加了22%，被试者增加了27%



人脸表情数据库

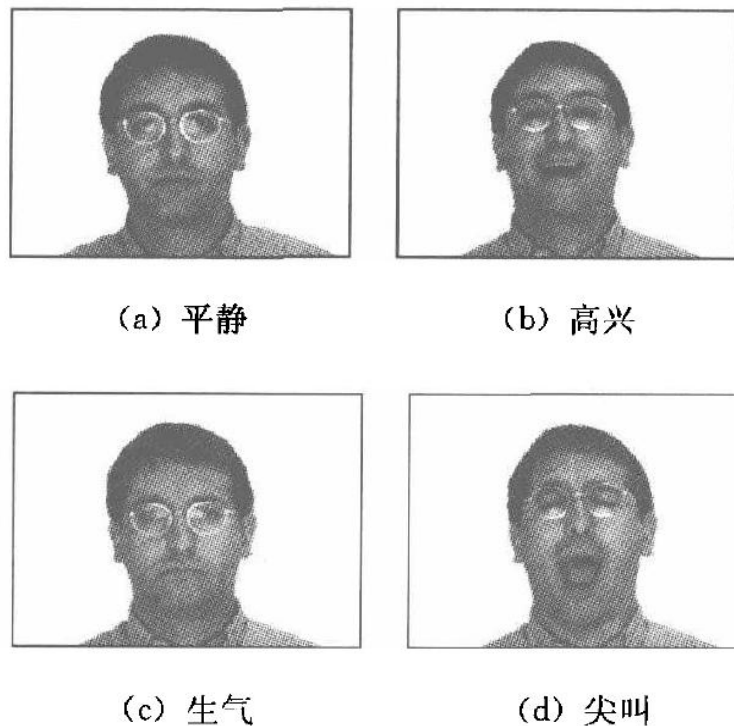
■ Yale人脸表情数据库

■ Yale人脸数据库包含部分非基本表情。如困乏、眨眼等常见表情；



■ AR人脸表情数据库

■ AR人脸数据库，包含尖叫等表情；



人脸表情数据库

■ 韩国人脸表情数据库

- 韩国人脸数据库 (Korean Face Database, KFDB) 包含**眨眼**表情。



(a) 平静

(b) 高兴

(c) 眨眼

(d) 惊讶

(e) 生气

■ RaFD表情数据库

- 荷兰内梅亨大学数据库，**包含轻视表情，每种表情有左视、直视和右视三种注视方向，并从五个角度进行了拍摄**



(a) 悲伤

(b) 平静

(c) 生气

(d) 轻视

(e) 嫌恶

(f) 惊讶

(g) 害怕

(h) 高兴

人脸表情数据库

■ 微表情的数据库：

- 芬兰 Oulu大学的 SMIC (Spontaneous microexpression corpus) 和 SMIC2；
- 中国科学院的 CASME (Chinese Academy of Sciences microexpression)、CASME II和CAS(ME)³；

人脸表情数据库

■ 微表情数据库

- 芬兰 Oulu大学SMIC系列微表情数据库，包含三种不同类型（高帧率、普通、近红外线）相机拍摄的视频子集，总样本数量为306，数据集可分为3种情感类型：positive, negative and surprise。



人脸表情数据库

■ 微表情数据库

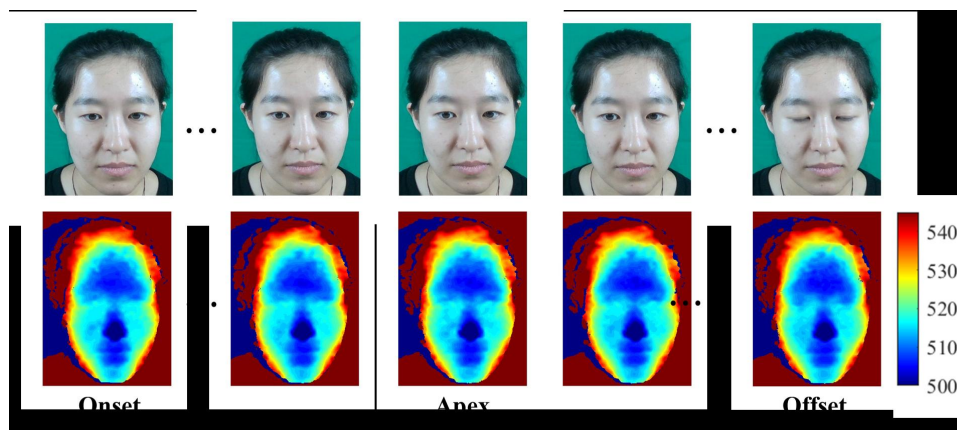
- 中科院心理所微表情数据库CASME，包含微表情视频样本195，受试者人数35，样本标签包含8种情感类型：Amusement (5)，Disgust (88)，Sadness (6)，Contempt (3)，Fear (2)，Tense (28)，Surprise (20)，Repression (40)；
- 中科院心理所微表情数据库CASMEII，包含微表情视频样本 247，受试者26，样本标签包含5种情感类型：Happiness (33 samples)，Repression (27)，Surprise (25)，Disgust (60) and Others (102)。



人脸表情数据库

■ 微表情数据库

- 中科院心理所微表情数据库CAS(ME)³，CAS(ME)³提供了约80小时的视频，超过800万帧，包括1109个人工标注的微表情和3490个宏观表情。此外，CAS(ME)³提供了1508个无标签的视频，超过4000000帧，搭建了基于无监督学习的微表情分析的数据平台。
- 深度流补充运动信息能够提高微表情检测性能，下图为微表情样本的RGB图像和深度信息。



提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

表情编码系统

- 将面部的各种运动和变形进行细致的编码，根据表情字典将编码与某种表情对应或描述各种表情
- 代表性的表情编码系统
 - 面部动作编码系统
 - 最大判别面部运动编码系统
 - MPEG-4 三维人脸动画标准
 - 面部表情编码系统

面部动作编码系统

- 面部动作编码系统FACS由Ekman和Friesen于1978年提出，其基于解剖学定义了44个运动单元(action unit, AU)以及运动描述符(action descriptor, AD)来描述面部运动。2002年确定了改进版本。
- 每个运动单元具有不同的位置和强度，强度分为三或五个等级。例如，AU 4由弱到强的强度可依次用AU 4A、AU 4B、AU 4C、AU 4D和AU 4E表示。

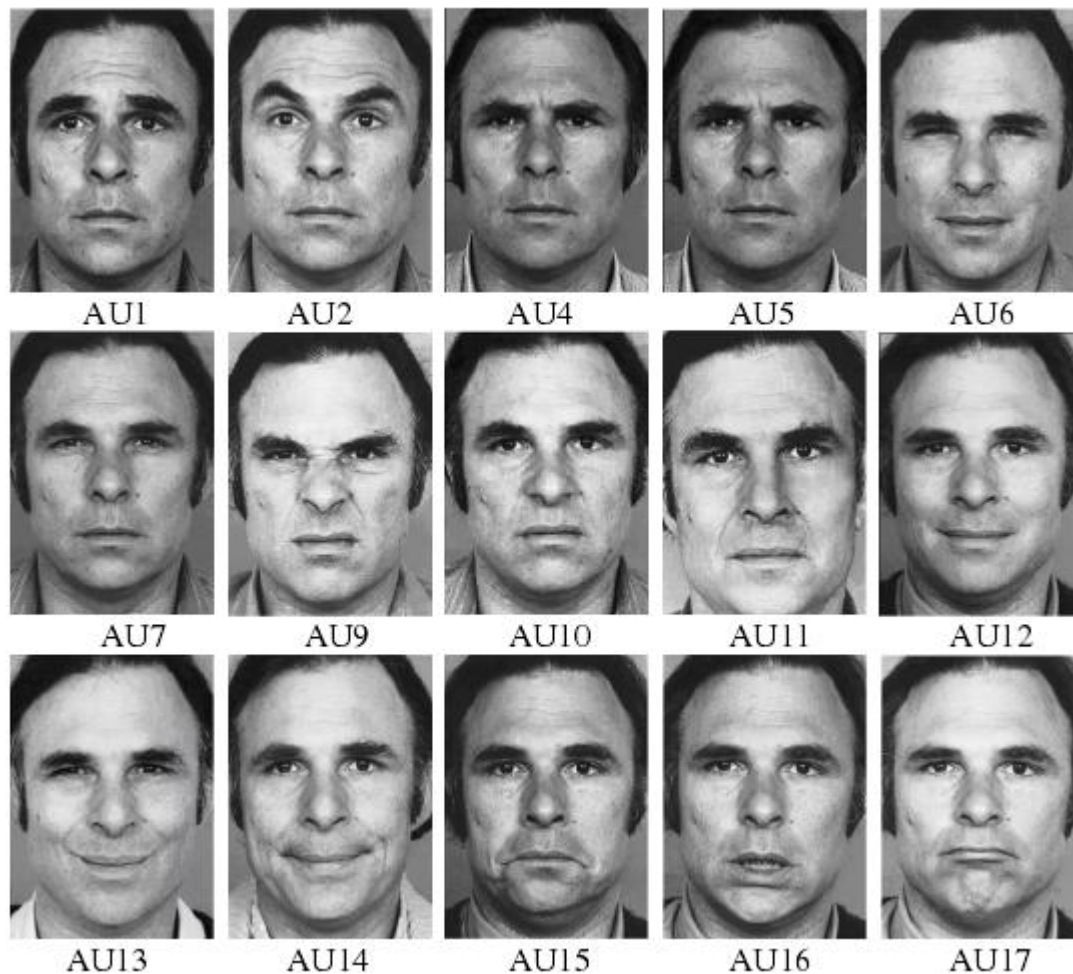
面部动作编码系统

■ 编码方式

名称	描述	名称	描述	名称	描述
AU 1	内眉上扬	AU 21	脖子收紧	AU 39	鼻孔收缩
AU 2	外眉上扬	AU 22	嘴唇呈漏斗状	AU 43	眼睛闭合
AU 4	眉毛下压	AU 23	嘴唇收紧	AU 45	眨眼(blink)
AU 5	上眼睑上扬	AU 24	嘴唇挤压	AU 46	眨眼(wink)
AU 6	脸颊上提	AU 25	上下嘴唇分开	AU 51	头向左
AU 7	眼睑收紧	AU 26	颚下降	AU 52	头向右
AU 9	鼻子起皱	AU 27	嘴巴伸展	AU 53	头上扬
AU 10	上嘴唇上扬	AU 28	嘴唇吸入	AU 54	头向下
AU 11	嘴唇沟纹加深	AD 29	推颚	AU 55	头向左倾斜
AU 12	唇角拉伸	AD 30	颚斜向一边	AU 56	头向右倾斜
AU 13	脸颊鼓起	AU 31	颚钳紧	AU 57	头前倾
AU 14	酒窝	AD 32	咬嘴唇	AU 58	头后倾
AU 15	唇角下压	AD 33	脸颊鼓起	AU 61	眼睛左转
AU 16	下唇下压	AD 34	脸颊膨胀	AU 62	眼睛右转
AU 17	下颚上提	AD 35	脸颊吸入	AU 63	眼睛向上
AU 18	嘴唇褶皱	AD 36	脸颊凸出	AU 64	眼睛向下
AD 19	吐舌	AD 37	舔嘴唇	AU 65	斜视
AU 20	嘴唇延伸	AU 38	鼻孔扩张	AU 66	对视

面部动作编码系统

■ 编码方式



名称	描述
AU1	内眉上扬
AU2	外眉上扬
AU4	眉毛下压
AU5	上眼脸上扬
AU6	脸颊上提
...
AU17	下巴向上

面部动作编码系统

■ FACS扩展

- 佐治亚理工学院的Essa和麻省理工学院的Pentland在FACS的基础上提出了FACS+系统。该系统包含**脸部运动的时间和空间模式**，能够更精确地描述人脸表情。
- 哈尔滨工业大学的金辉和高文提出了FACS' (转换FACS) 系统，将运动单元的运动转化为基于物理结构和肌肉模型的运动特征向量序列，**对眼部和嘴部分别进行表情编码**。

最大判别面部运动编码系统 (MAX)

- 最大判别面部运动编码系统MAX系统由Izard提出，开发该编码系统的主要原因是为了提供一个有效和可靠的系统来测量婴儿和儿童的面部行为中的情感信号。
- MAX系统的标记语言包括68个数字代码。例如，代码0表示没有观察到的运动。
 - 该系统将表情分为三组，分别为：代码1表示眉毛、前额和鼻根；代码2表示眼睛、鼻子和脸颊；代码3表示嘴巴和下颚。
 - 例如，平静的表情可以表示为0/0/0。可由该系统进行解释的表情包括：感兴趣—兴奋、快乐—喜悦、惊奇—惊讶、悲伤—沮丧、生气—愤怒、害怕—恐惧、不适—疼痛。

MPEG-4三维人脸动画标准

■ MPEG-4三维人脸动画标准包含了两部分：

- 人脸定义参数 (facial definition parameter, FDP)，用于描述特定人脸纹理形状模型与通用人脸模型间的差别。
- 人脸动画参数 (facial animation parameter, FAP)，用于描述特定的人脸表情与中性表情的变化关系。

■ 编码方式

- MPEG-4定义了68个FAP，描述人脸某一区域在某个方向上的运动
- 第一个FAP为viseme，决定两个视位(在语音驱动说话人头部动画中一种常用的为口形建立的音频—视频模型)混合值 口型的变化
- 第二个FAP为expression，决定两种表情混合值 表情
- 其他66个FAP分别定义了人脸某一小区域的运动

面部表情编码系统 (FACES) 可计算度低

- 美国加利福尼亚大学的Kring等提出了FACES。
- 该系统提供了**表情的频率、强度、诱力和持续时间**等信息，能够对分布在情感维度模型上的表情进行度量，并且所需时间较短。
- 利用两个评价者对参与者的表情进行编码，并且注意了许多环节(如身体位置的移动、手部的遮挡、眼镜、讲话、面部抽搐等)对表情的影响，使得FACES的可靠性很高。

视频情感特征

- 局部二值模式 (LBP)

- Gabor

- 主动外观模型 (ASM)

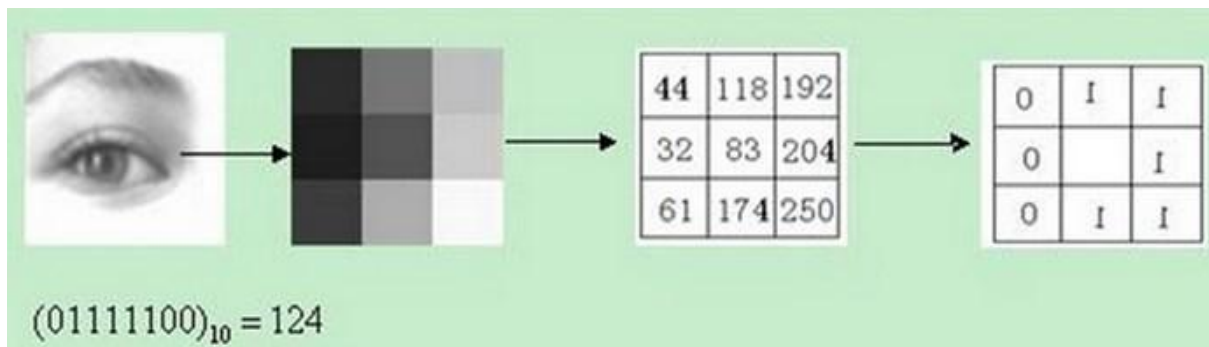
- 光流法

- 深度视频特征

视频情感特征

■ 局部二值模式特征

- 在像素3*3的邻域内，以邻域中心像素为阈值，相邻的8个像素的灰度值与阈值进行比较，若周围像素大于阈值，则该像素点的位置被标记为1，否则为0。
- 将这8位二进制数依次排列形成一个二进制数字，这个二进制数字就是中心像素的LBP值，反映了该像素周围区域的纹理信息。
- 能够有效描述纹理信息，具有良好的旋转不变性。

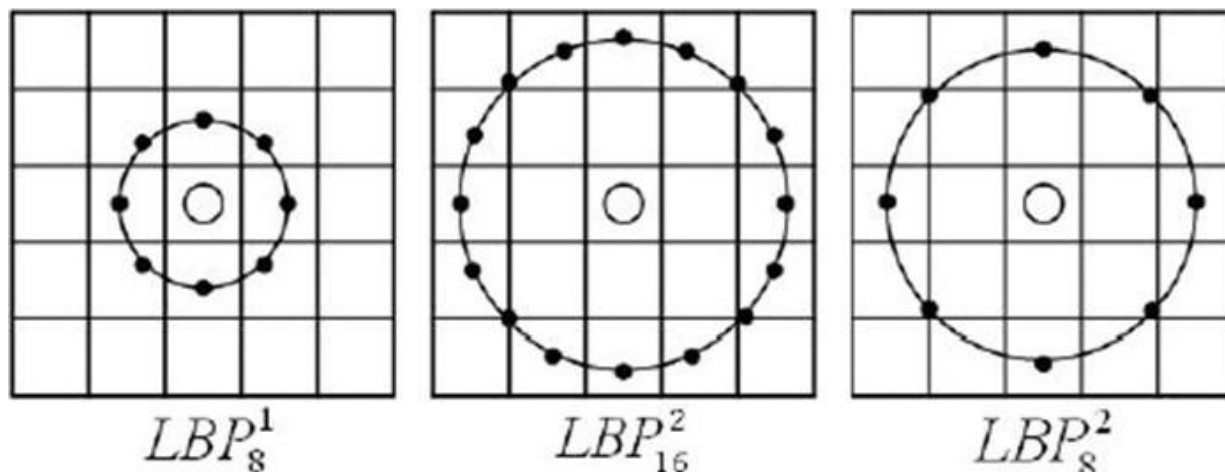


注：如果是彩色图，需要先转换成灰度图

视频情感特征

■ LBP变体——圆形LBP特征

■ 使用可变半径的圆对近邻像素进行编码，可以得到如下的近邻：（LBPPR表示不同的圆形半径R，不同的采用点数目P下的LBP计算方式）

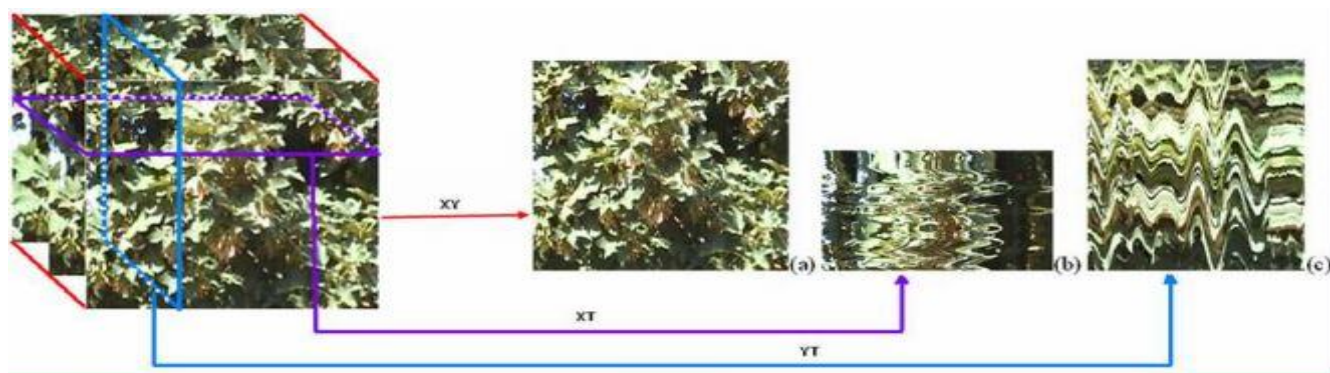


视频情感特征

■ LBP变体——LBP-TOP特征

■ 一个图像序列，给出三个正交平面的纹理图，X-Y 就是正常图像，X-T 是每一行沿着时间轴扫描而过的纹理，Y-T 是每一列沿着时间轴扫描而过的图像。

■ LBP-TOP在这三个平面都提取 LBP 特征，然后把三个平面的 LBP 特征连接起来。



Ahonen T, Hadid A, Pietikainen M. Face recognition with local binary

pattern.

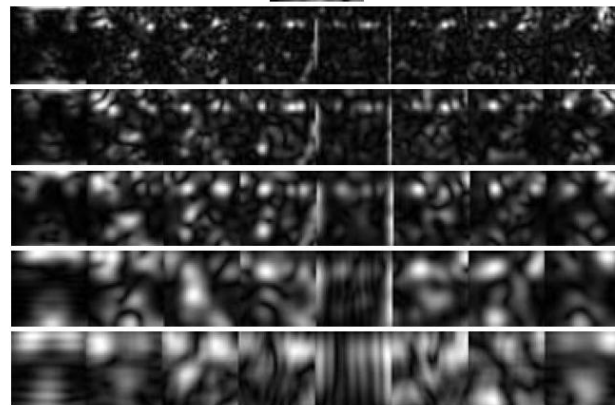
视频情感特征

■ Gabor特征

■ 通过滤波器对输入图像进行滤波，根据结果描述像素之间的关系

■ 从不同尺度、不同方向上提取信息，每个像素点用一组数据描述，计算量大

■ 对光照具有良好的适应性



Gabor滤波器组滤波后的图像

Liu C, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition.

视频情感特征

■ 主动外观模型（ASM）

- 基于点分布模型提取表情轮廓的特征点，根据搜索结果不断优化统计模型的参数，最终使模型和表情轮廓匹配
- ASM模型能够有效的覆盖人脸集合外形子空间，**具有较高的点定位精度；能够反映图像的局部纹理信息**



ASM特征点集合模型

Cootes T F, Taylor C J. Active shape models-their training and application.

视频情感特征

■ 光流法

- 光流是用来测量人脸面部像素点运动的瞬时速度场
- 利用图像序列中像素在时间域上的变化以及相邻帧之间的相关性确定当前帧跟上一帧之间存在的对应关系
- 光流法在微表情识别中广泛应用



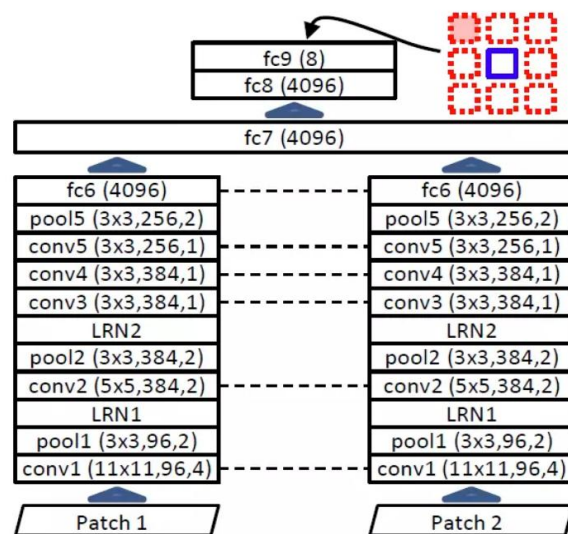
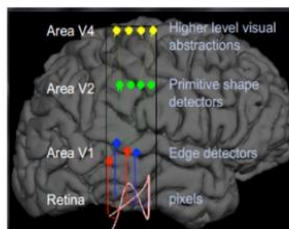
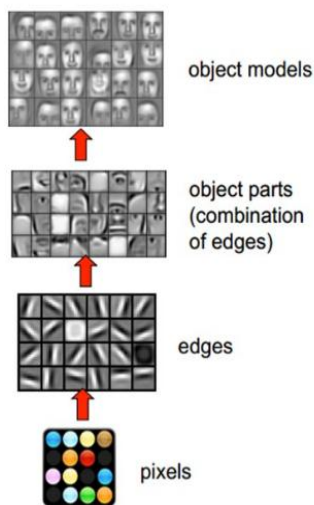
基于光流法对嘴巴特征点跟踪

Bourel F. Robust facial expression recognition using a state-based model of spatially-localized facial dynamics.

视频情感特征

深度视频特征

- 视觉系统的信息处理是分级的
- 从底层到高层的特征表示越来越抽象
- 神经网络的底层刻画边缘信息，中间层将底层边缘信息进行组合，高层能抽象复杂的语义信息



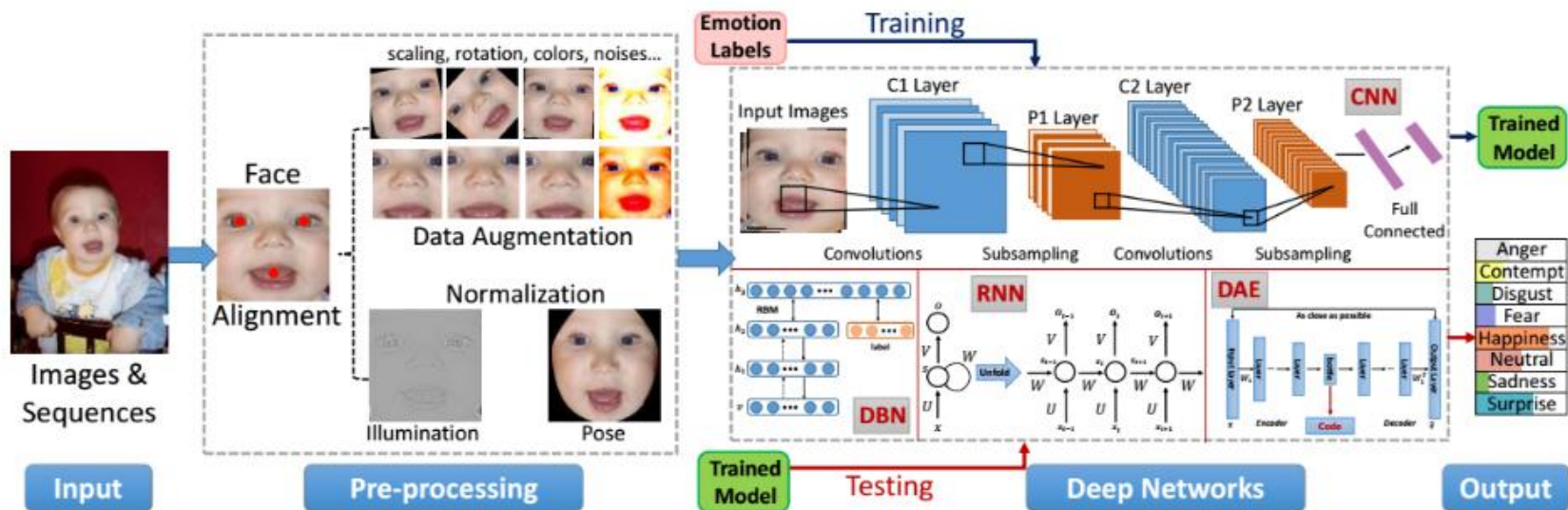
提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

表情识别

表情识别的流程

- 预处理：人脸检测、人脸对齐、数据增强、人脸归一化
- 特征提取
- 面部表情分类



表情识别

■ 主要分类模型

- 隐马尔科夫模型
- 支持向量机
- AdaBoost方法
- 稀疏表示法
- 深度学习方法

表情识别

■ 隐马尔科夫模型

抗干扰能力不足, 泛化能力不足
(光线) (个人表情)

- 适用于动态过程时间序列建模, 具有强大的模式分类能力, 理论上可处理任意长度的时序
- 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
- 运用HMM方法能够比较精确的描绘表情的变化本质和动态性能, 追踪到表情图片中隐含的状态转移特征
- 选取CMU人脸表情数据库实验结果

表情	正确	错误	正确率/(%)
中性	100	0	100
高兴	93	7	93
惊讶	90	10	90
愤怒	87	13	87
恐惧	82	18	82
厌恶	80	20	80
悲伤	78	22	78

Bartlett M S, Littlewort G, Frank M. Recognizing facial expression: machine learning and application to spontaneous behavior.

表情识别

■ 支持向量机

- 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
- 作为分类器做表情识别，在表情识别时一般和Gabor滤波器一起使用作为分类器

+ Gabor滤波器 最优.

- 在JAFPE人脸表情数据库实验结果

表情	正确	错误	正确率/(%)
中性	10	0	100
高兴	8	2	80
惊讶	10	0	100
愤怒	10	0	100
恐惧	6	4	60
厌恶	10	0	100
悲伤	7	3	70

Xu C, Dong C, Feng Z. Facial expression pervasive analysis based on Haar-like feature and SVM.

表情识别

■ AdaBoost方法 (小样本)

- 针对同一个训练集训练不同的分类器（弱分类器），然后把这些弱分类器集合起来，构成一个更强的最终分类器
- 通过融合多分类器有效提高表情识别的准确率

■ 矩阵分解法 (干扰消除)

- 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

Xiang Y, Wu Y, Peng J. An improved Adaboost face detection algorithm based on the weighting parameters of weak classifier.

表情识别

■ 基于卷积神经网络的表情识别

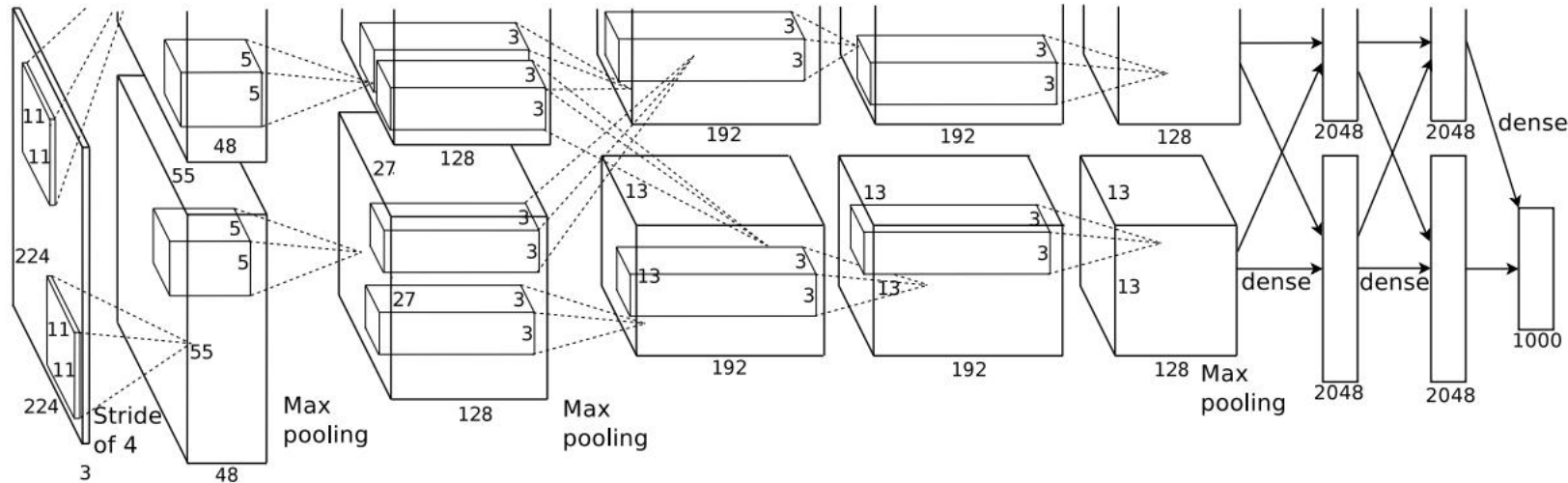
- 直接将人脸图像作为特征输入，在特征自学习能力上具有优势
- 用于表情识别的典型网络结构包括AlexNet、Inception等

Wang J, Yuan C. Facial expression recognition with multi-scale convolution neural network.

表情识别

AlexNet

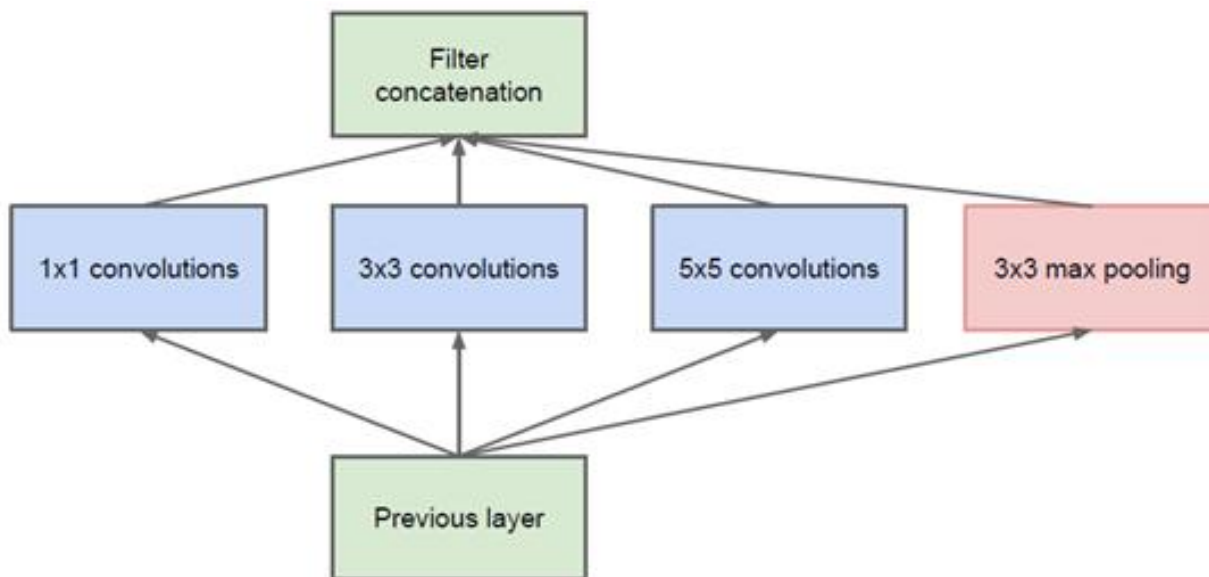
- 它是一种高效的端到端表情识别方法。
- 网络结构分为上下2层，即网络分布在2个GPU上分组处理。
- 在CK+人脸表情数据集中测试达到88.57%的准确率。抗干扰能力更强
数据要求更高



表情识别

■ Inception结构

- 将CNN中常用的卷积（1x1，3x3，5x5）、池化操作（3x3）堆叠在一起，一方面增加了网络的宽度，另一方面也增加了网络对尺度的适应性。

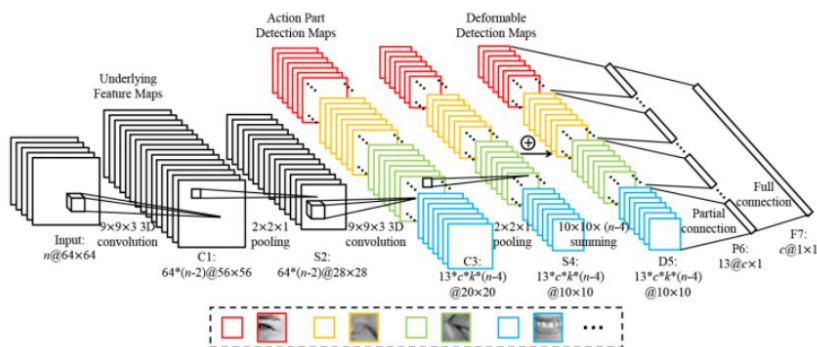


表情识别

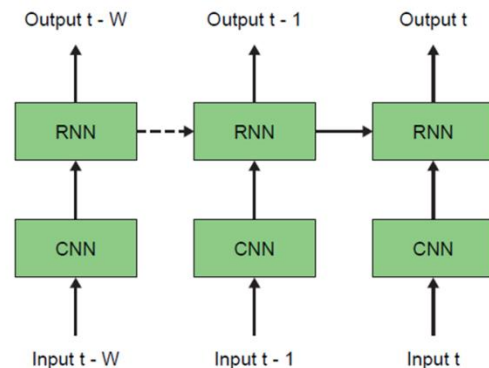
■ 考虑时序性

■ C3D (3D convolutional networks) 对于通常人脸图像上的2D空间卷积，沿着时间轴增加了一个时间维度，就是3D时空卷积。

■ CNN+LSTM: CNN用于人脸特征抽取，LSTM用于视频时序建模



C3D



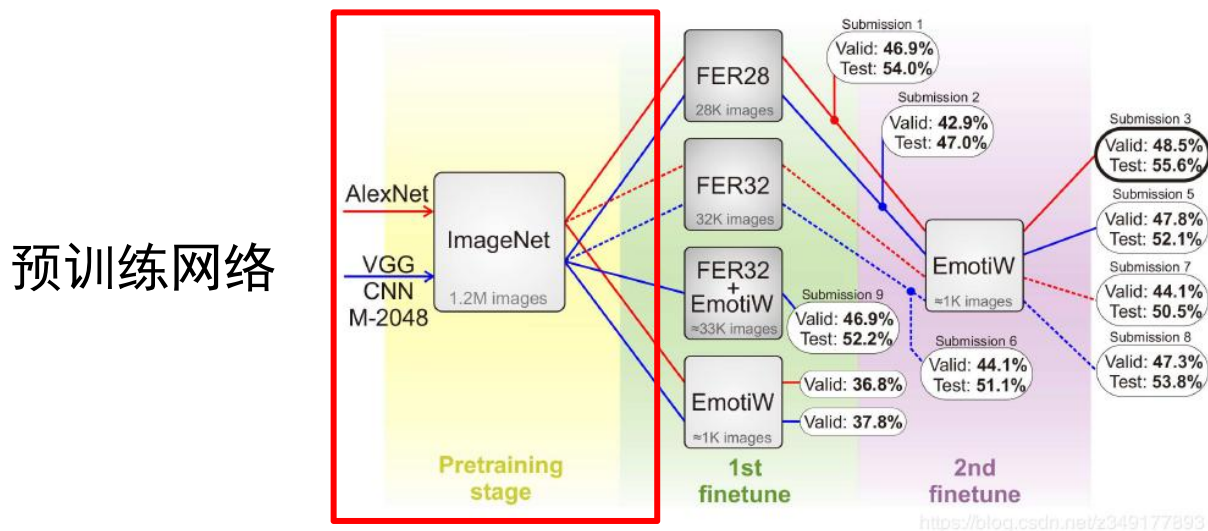
CNN+LSTM

Ambadar Z, Schooler J W, Cohn J F. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions.

表情识别

■ 预训练与微调

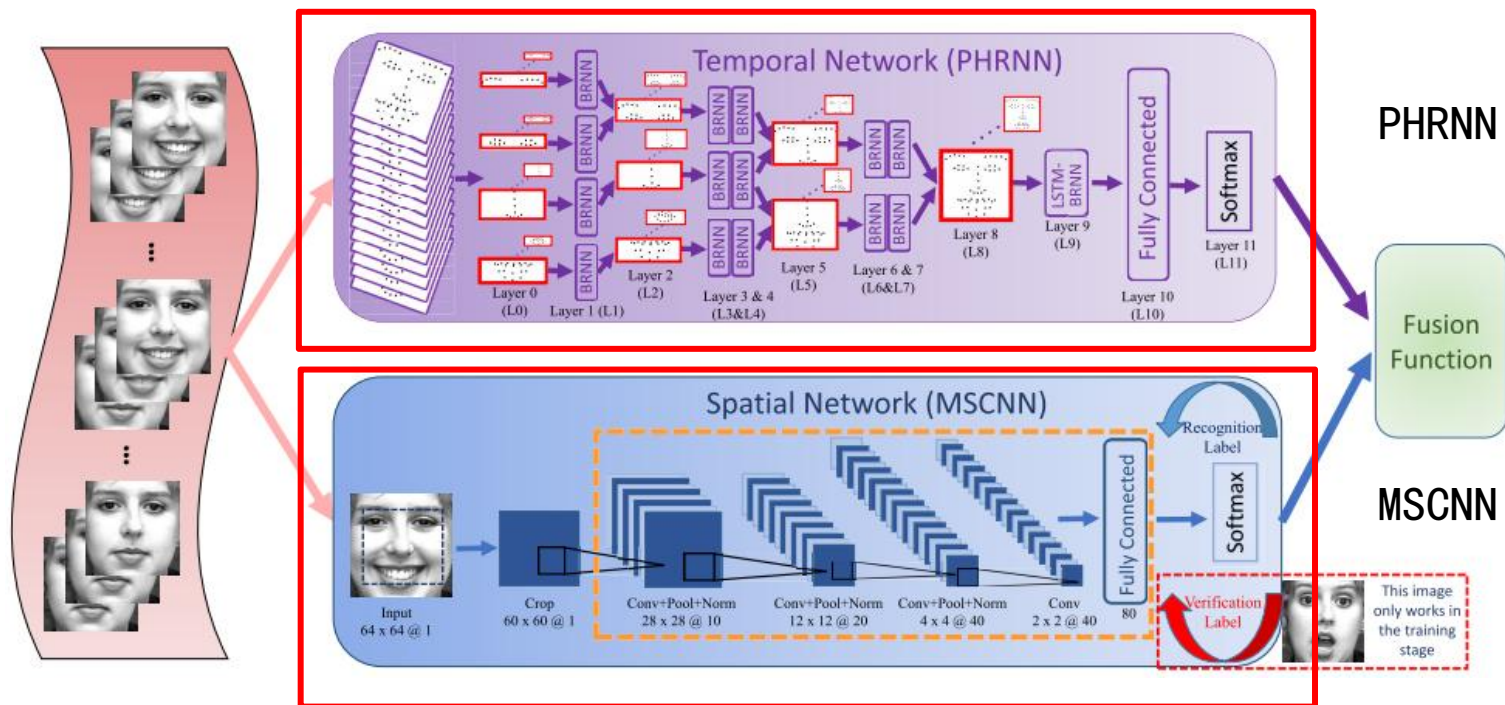
- 解决表情识别数据不足，易过拟合的问题；
- 使用额外的任务导向的数据预训练自定义的网络，或者在已经预训练好的网络模型上进行微调；
- 第一阶段在预训练模型上使用 FER2013 进行微调，第二阶段利用目标数据库的训练数据进行微调。使模型更切合目标数据库



表情识别

网络集成

- 利用PHRNN模块构建时态网络，捕获连续帧的动态特征。
- 构造基于MSCNN的空间网络，从静止帧中提取静态特征。
- 将两种网络相结合，提高人脸表情识别的性能



Li S, Deng W. Deep Facial Expression Recognition: A Survey.

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

微表情识别

- 微表情是一种人类在试图隐藏某种情感时无意识做出的短暂的自发式面部表情；它的**持续时间仅为 1/25 秒至 1/5 秒**，表达的是一个人试图压抑与隐藏的真正情感
- 在一段心理医生与抑郁症病人的对话视频中，经常**微笑的病人有几帧非常痛苦的表情**。他将这种人在经历强烈情绪时产生的快速、无意识的自发式面部动作成为微表情
- 微表情是一种特殊的面部微小动作，可以作为判断人主观情绪的重要依据，在公共安防和心理治疗领域有着广泛的应用价值

微表情识别

■ 微表情特点

- 动作幅度小；
- 持续时间短
- 不会在上半脸和下半脸同时出现。

■ 对微表情的人工识别需要有专业的培训；研究人员利用计算机视觉技术自动识别微表情以提高微表情的应用可行性；**依靠人力识别的准确率只有 47 % 。**

■ 计算机就能够以低廉的成本处理 大规模的微表情识别任务，这显然超过专业人员人 工识别微表情的效率

微表情识别

下面展示数据集中的一些图片：



(a) 取自 SMIC 2/HS 的惊讶的微表情
(a) Surprise microexpression from SMIC 2/HS

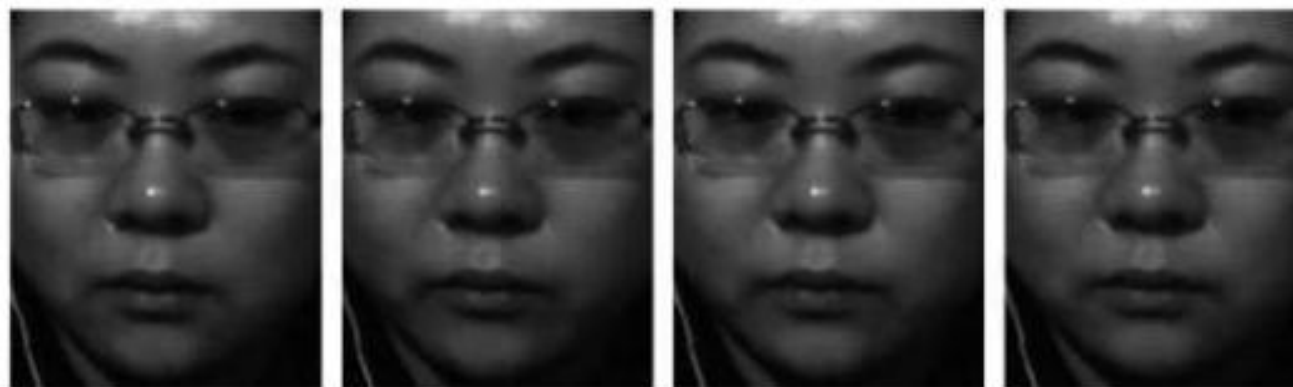


(b) 取自 SMIC 2/VIS 的非微表情
(b) Non-microexpression from SMIC 2/VIS

微表情识别



(c) 取自 SMIC 2/NIR 的惊讶的微表情
(c) Surprise microexpression from SMIC 2/NIR



(d) 取自 SMIC 的消极的微表情
(d) Negative microexpression from SMIC

微表情识别



(e) 取自 CASME 的厌恶的微表情
(e) Disgust microexpression from CASME

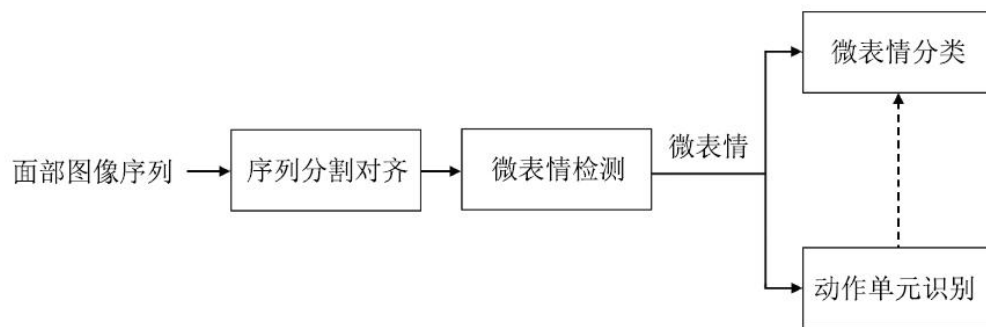


(f) 取自 CASME II 的愉悦的微表情
(f) Happy microexpression from CASME II

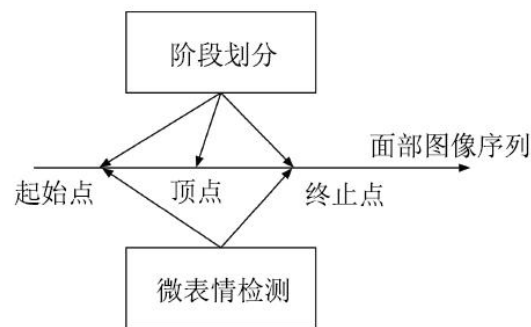
微表情识别

■ 微表情识别流程

- 预处理：序列分割对齐
- 微表情检测：检测图像序列中是否包含微表情
- 微表情识别：识别微表情状态



(a)



(b)

■ 微表情分类模型与表情识别类似，这里重点介绍针对微表情处理的典型方法

■ 主要方法

- 基于LBP-TOP的微表情识别
- 基于MDMO的微表情识别方法 *光流法*
- 基于De launay时域编码的识别方法
- 基于Gabor特征的识别方法
- 基于面部动力谱特征（Facial dynamics map, FDM）的识别方法
- 基于运动模式放大的分类方法
- 基于深度学习的微表情识别
- 基于时空网络的微表情识别
- 基于振动图像的微表情识别

微表情识别

■ 基于LBP-TOP的微表情识别方法

- 该方法首先利用68点主动形状模型（Active shape model, ASM）**确定人脸的关键点**；
- **以得到的关键点为基础**，利用局部加权平均算法（Local weighted mean, LWM）[71]**计算各序列的第1帧中的人脸图像与模型的人脸图像的形变关系**；
- 形变关系作用于对应序列的各帧图像。由此，**不同人脸表情状态下的不同序列之间存在差异**。

Pfister, Tomas, Xiaobai Li, Guoying Zhao and Matti Pietikäinen. “Recognising spontaneous facial micro-expressions.” 2011 International Conference on Computer Vision (2011): 1449–1456.

微表情识别

■ 基于MDMO的微表情识别方法

- 在视频序列中提取了光流主方向，进而计算了面部块中的平均光流特征，提出了主方向平均光流特征；
- 使用人脸检测算法定位各帧的人脸，接着修正自第2帧起的各帧的光流场，寻找仿射变换矩阵，使得各帧的人脸特征点在该矩阵变换下与第1帧的人脸间的差异最小；
- 在特征提取中，基于关键点的面部块规则，将面部分割成不相互重叠的36个区域。同时，每帧提取一个光流场，每一块提取一个主方向，分别计算各分区中最相似的光流场运动矢量的平均值，作为该区域的运动特征。

Barron, John L., David J. Fleet and Steven S. Beauchemin. "Performance of optical flow techniques." International Journal of Computer Vision 12 (1992): 43-77.

微表情识别

■ 基于Delaunay时域编码的识别方法

- 微表情自身变动幅度很小，仅用关键点难以描述表情变化；**该方法利用特征点将图像归一化，得到特征点位置固定的人脸图像序列**
- 根据给定的特征点**把人脸分割为一系列三角区域**，因为特征点已经归一化，每个三角区域的大小和形状是相同的
- 通过对比相同区域随着时间的变化，可以刻画微表情的动态过程

■ 基于Gabor特征的识别方法

- 首先利用Gabor特征计算每帧的情绪，**通过帧率确定每段视频的持续时间确定微表情片段，进一步进行微表情分类**
- 通常对Gabor特征进行PCA和LDA降维

Lu Z Y, Luo Z Q, Zheng H C. A delaunay-based temporal coding model for micro-expression recognition.

Wu W, Shen X B, Fu X L. The machine knows what you are hiding: an automatic micro-expression recognition system.

微表情识别

■ 基于面部动力谱特征（Facial dynamics map, FDM）的识别方法

- 以光流场为描述微表情运动模式的基础特征，提出了面部动力谱特征，将抽取出的光流场进一步分割成小的时空立方体提取特征参数
- 能够较好的反应微表情的运动模式，且计算的面部动力谱特征容易可视化，对微表情深层理解起到辅助作用

■ 基于运动模式放大的分类方法

- 使用欧拉视频放大技术将微表情的动作幅度增强放大，再通过常规的识别技术对放大的微表情进行识别

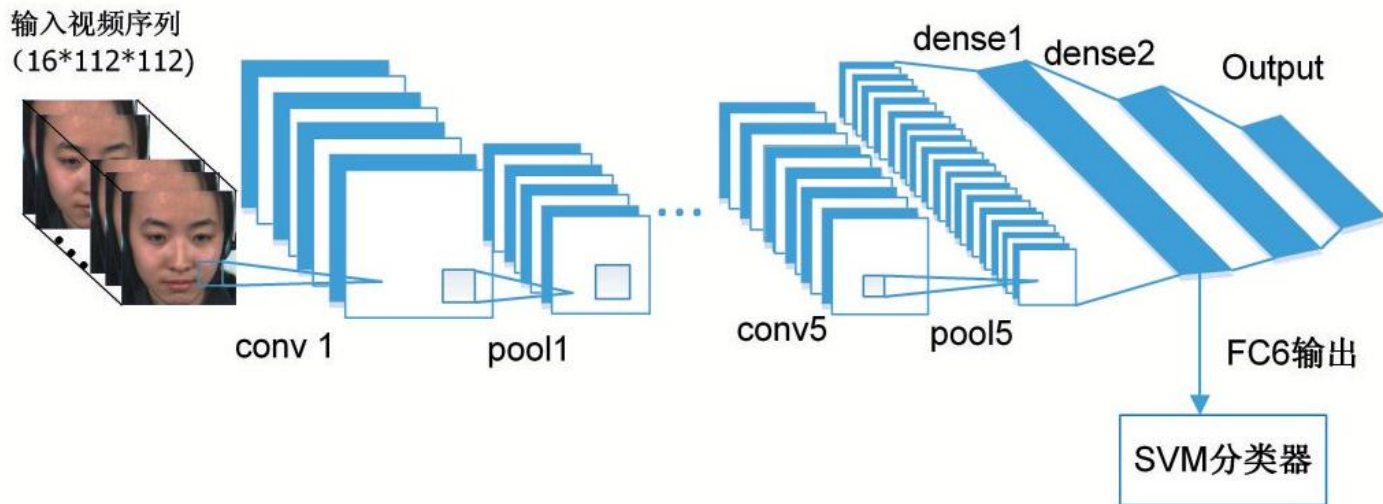
Xu F, Zhang J P, Wang J Z. Micro-expression identification and categorization using a facial dynamics map.

Li X B, Hong X P. Reading hidden emotions:spontaneous micro-expression spotting and recognition.

微表情识别

■ 基于深度学习的微表情识别

- 由于微表情识别的训练样本有限，通常利用深度神经网络抽取深度特征，然后基于深度特征采用传统分类器模型进行训练。
- 下图中利用3D-CNN提取dense1这个中间层输出特征并送入SVM分类器进行微表情识别。



Santos J, Seixas D, Brandao S. Investigating the role of the ventromedial Prefrontal Cortex in the Assessment of Brands.

微表情识别

■ 基于时空网络的微表情识别方法

- CNN能够有效地提取出图像的空间特征，而RNN能够有效地提取出序列中的时间特征，如果能够结合CNN和RNN的特点，就能有效提取出微表情序列中的时空特征；
- 首先从微表情序列中提取具有代表性的5帧：onset（起始帧），apex（顶点帧），offset（结束帧）以及onset和apex中间帧、apex和offset的中间帧，将每一帧使用CNN提取出特征，然后将5帧特征输入到LSTM中训练以提取时间特征，最后进行分类。
- 引进迁移学习的方法，先使用宏观表情训练卷积神经网络，然后在微表情数据集上进行微调，以此提升小样本学习的能力。

Liu, Yong-Jin, Jinkai Zhang, Wen-Jing Yan, Sujing Wang, Guoying Zhao and Xiaolan Fu. “A Main Directional Mean Optical Flow Feature for Spontaneous Micro-Expression Recognition.” IEEE Transactions on Affective Computing 7 (2016): 299–310.

微表情识别

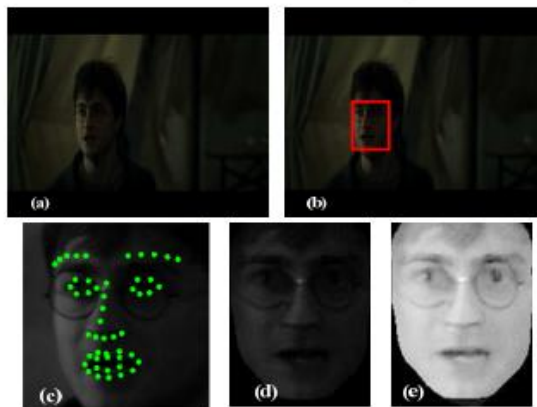
■ 振动图像的微表情识别

- 一个人的面部图像振动分量（VC）包含关于面部运动以及面部区域的不自主颤抖的信息。后一种成分通常与神经紧张程度有关，通常表现为嘴唇颤抖、眉毛、面部左右侧抽搐。为了识别上述VC，通常可以分析相邻的图像帧，揭示微小的变化，以这种方式获得的数据能评估面部不同部分的当前震颤水平，并对人的当前心理情绪状态的变化做出可靠的结论。
判断有没有微表情.
- 面部振动图像微表情识别技术可用于潜在危险物体（PDO）控制操作员的功能和心理情绪状态（FPS）。危险物体包括热电站、核电站、化工厂、水利建筑、高速运输等。除了上述危险物体外，在医学、体育和教育领域，对心理情绪状态的监控也是非常必要的。

微表情识别

■ 振动图像的微表情识别

- 在照明不均匀和不稳定的条件下分析人的面部振动图像时，为了增强对人的当前心理情绪状态的估计的可靠性，提出的方法本质是**补偿由照明的空间不均匀性或照明的时间不稳定性引起的人脸振动图像失真**；
- (a) 源图像；(b) 面部检测 (用红色矩形示出)；(c) 面部特征定位 (用 49 个绿色点示出) 和人脸裁剪；(d) 面部正面化；(e) 照明标准化。



Kolobashkina L, Alyushin V. Increasing the informativeness content of human face vibration image through application principles of cognitive psychology[J].

Procedia Computer Science, 2020, 169: 876-880.

■ 振动图像的微表情识别

- 2018年俄罗斯国立核研究大学Mikhail Alyushin提出了一种处理人脸热图像的振动分量的方法，目的是评估其当前的心理情绪状态。该方法的本质是**创建一个人的热图像的个体模型 (IMTIP)**。该模型包括图形组件，这些组件包含以下信息：
 - 1. 信息最丰富的面部无意识振动和面部表情 (IMTIP1)；
 - 2. 由于人的心脏和呼吸活动而明显表现出调节效应的面部区域 (IMTIPM)；
 - 3. 用于将上述IMTIP组件链接到以不同图形格式获得的面部图像的参考点。
- 以这种方式形成的IMTIP使得选择性地处理人脸的热图像成为可能。

提纲

- 背景及意义
- 研究现状与进展
- 人脸表情数据库
- 人脸表情特征
- 表情识别
- 微表情识别
- 展望

■ 主要局限

数据不足

- 表情识别的鲁棒性不足，容易受到光照、遮挡物的影响；(NN降低)
- 不同个体的外貌、肤色、文化背景存在差异，需要提高模型的泛化能力；
- 表情的刻画不够细致，主要集中在有限的几种表情，缺乏对表情的细粒度刻画；
动作方面
- 表情识别的性能依赖于大数据支撑，资源受限条件下的识别性能不足；

■ 未来的发展方向

- **纳入深度信息：**目前的表情识别大多依赖于2D图像；基于深度成像技术能够支持3D脸部分析，且有助于识别2D外观不易察觉的表情；
- **表情细微变化的建模：**在更为复杂的情感空间中实现不同个体的表情识别与跟踪；
- **融入非视觉信息：**通过整合多源信息判断面部表情，如通过热成像探测不同面部区域的辐射，根据血液的流动变化判断表情；

Thanks