



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

# 情感计算 ——课程总结



中国科学院自动化研究所

刘斌

liubin@nlpr.ia.ac.cn

# 课程大纲

## ■ 第一部分 前言

- 1.1 情感计算内涵
- 1.2 情感计算历史
- 1.3 情感计算主要方法

## ■ 第二部分 情感认知模型

- 2.1 情感的心理学基础
- 2.2 离散情感模型
- 2.3 连续情感模型

## ■ 第三部分 情感识别

- 3.1 音频情感识别
- 3.2 表情情感识别
- 3.3 微表情情感识别
- 3.4 肢体动作情感识别
- 3.5 生理信号情感识别
- 3.6 多模态融合情感识别

## ■ 第四部分 倾向性分析

- 4.1 文本情感分析
- 4.2 舆情分析

## ■ 第五部分 情感生成

- 5.1 情感语音生成
- 5.2 表情生成
- 5.3 多模态情感生成

## ■ 第六部分 情感交互

- 6.1 情感意图理解
- 6.2 多模态情感交互方法
- 6.3 情感交互的技术标准

## ■ 第七部分 主要应用

- 7.1 情感机器人
- 7.2 医疗健康
- 7.3 舆情分析
- 7.4 安全领域
- 7.5 金融领域
- 7.6 教育领域

# 参考资料

---

- Jianhua Tao, Tieniu Tan, Affective Information Processing, Springer, 2010
- Rosalind W. Picard, Affective Computing, MIT Press, 2000
- Marvin Minsky, 情感机器 (The Emotion Machine: Commonsense Thinking, Artific) (王文革等译), 浙江人民出版社, 2015
- 吴敏, 刘振焄, 陈略峰, 情感计算与情感机器人系统, 科学出版社, 2018
- 詹永照, 毛启容, 林庆, 成科扬, 视觉语音情感识别, 科学出版社, 2018
- 刘光远, 人体生理信号的情感计算方法, 科学出版社, 2014
- 杜坤坤, 刘欣, 王志良, 解仑, 情感机器人, 机械工业出版社, 2013
- 刘兵, 情感分析: 挖掘观点、情感和情绪 (刘康译), 机械工业出版社, 2017

## ■ 情感是人对客观事物的态度体验及相应的行为反应

- 客观事物或情境 与 个体的需要和愿望
  - 符合，引起积极的、肯定的情感
  - 不符合，引起消极的、否定的情感
- 情感是个体与环境间某种关系的维持或改变

# 情感概述

---

## ■ 主观体验 (subjective experience)

- 个体对不同情感状态的自我感受

## ■ 外部表现，即表情 (emotional expressions)

- 在情感状态发生时身体各部分的动作量化形式
  - 面部表情：面部肌肉变化所组成的模式
  - 姿态表情：身体其他部分的表情动作
  - 语调表情：言语的声调、节奏和速度等方面的变化

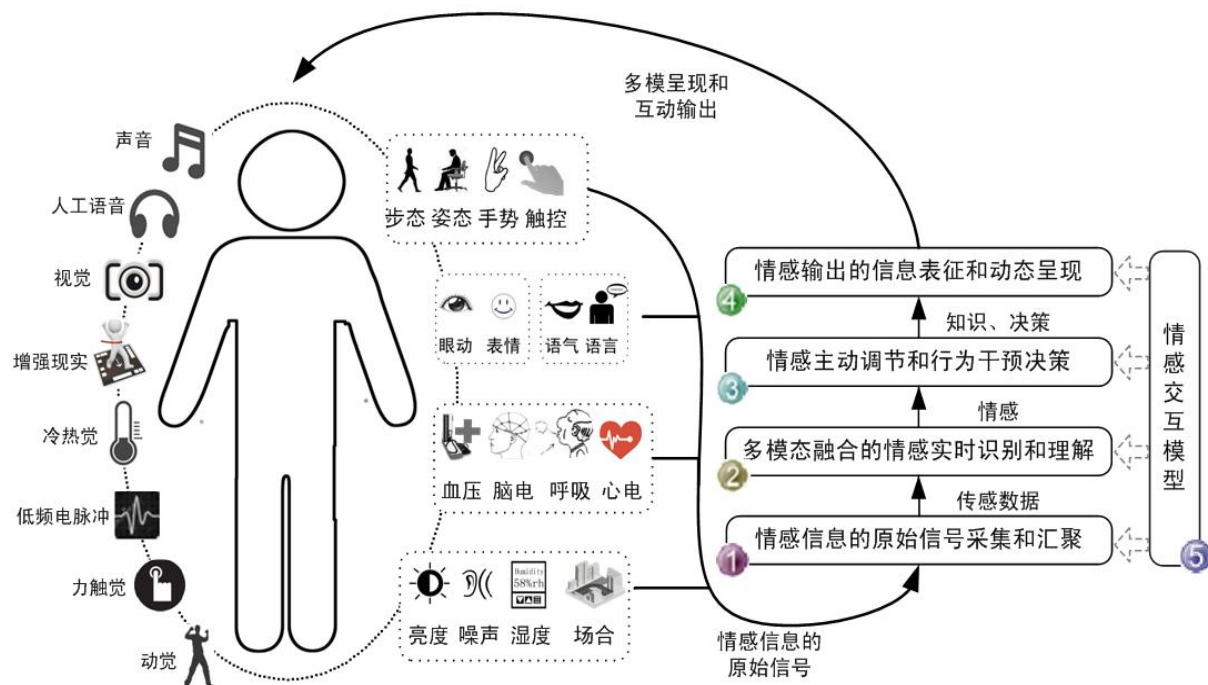
## ■ 生理唤醒 (physical arousal)

- 情感产生的生理反应，是一种生理的激活水平，有不同的反应模式

# 情感概述

- 人操作机器 → 机器辅助人
- 人围着计算机转 → 计算机围着人转
- 认知型 → 直觉型

在人计交互中，计算机应具有情感能力！



# 情感认知模型

---

- 基于基本情感论的情感模型
- 基于维度空间论的情感模型
- 基于认知机制的情感模型
- 基于个性化的情感模型
- 其他情感模型

# 情感认知模型

- 基本情感论：情感具有原型模式(prototype)，即存在着数种基本情感类型。

学者	基本情感
Arnold	Anger, aversion, courage, dejection, desire, despair, dear, hate, hope, love, sadness
Ekman, Friesen, Ellsworth	Anger, disgust, fear, joy, sadness, surprise
Fridja	Desire, happiness, interest, surprise, wonder, sorrow
Gray	Desire, happiness, interest, surprise, wonder, sorrow
Izard	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise
James	Fear, grief, love, rage
McDougall	Fear, disgust, elation, fear, subjection, tender-emotion, wonder
Mower	Pain, pleasure
Oatley, Johnson-Laird	Anger, disgust, anxiety, happiness, sadness
Panksepp	Anger, disgust, anxiety, happiness, sadness
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise
Watson	Fear, love, rage
Weiner, Graham	Happiness, sadness



# 情感认知模型

---

- 基本情感论的缺陷催生了维度空间论。
- 维度空间论采用直观的低维欧式空间来定量地描述、理解人类的情感。
- 情感的维度是指情感在某种特性上，可以用一根实数轴来度量其变化。例如，情感的正性和负性属性。

# 情感认知模型

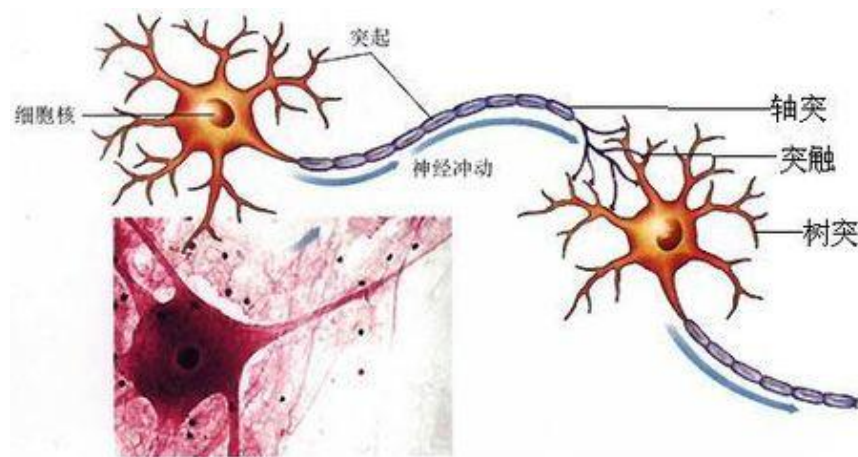
## ■ 基本情感论 vs. 维度空间论

- 基本情感论将情感表示成离散的状态，维度空间论则立足于人类情感体验的欧式空间描述。
- 两者之间的边界并不是严格的，基本情感通过一定方法可以映射到连续的维度空间中，反之亦然。

考察点	离散情感	维度空间情感
情感描述方式	形容词标签	欧式空间的坐标点
情感描述能力	有限的几种情感类别	任意情感变化
被应用到情感识别的时期	1980s	2000s
优点	简洁、易懂、容易着手	无限的情感描述能力
缺点	单一、有限的情感描述能力无法满足对情感细微变化的描述	将主观情感量化为欧式空间上的坐标值对普通人来说有一定的挑战

# 情感认知模型

- 情感是认知神经科学的重要组成部分。 **交感与副交感神经活动是情感产生的重要生理基础**。情感的变化会引起肾上腺素升高，血糖升高，心跳加快，肌肉紧张，瞳孔散大等。



# 情感认知模型

---

## ■ 情感相关联的神经结构

- 认知神经科学认为人体的情感必须以人脑的神经结构和神经活动模式为基础，试图建立情感的神经机制图，以便把握情感的神经活动机制。
- 人脑在处理情感相关信息时，某些特定的脑神经结构的水平确因情感信息变化而有显著的差异。
- 目前已知脑区中，杏仁核、眶额叶皮层、前后扣带回皮层、脑岛、伏核、丘脑、腹侧被盖区等7个神经结构均与情感密切相关。

# 情感认知模型

---

- 大五模型
- Chittaro行为模型
- EFA性格空间的构造方法
- 情绪—心情—性格模型

# 情感认知模型

---

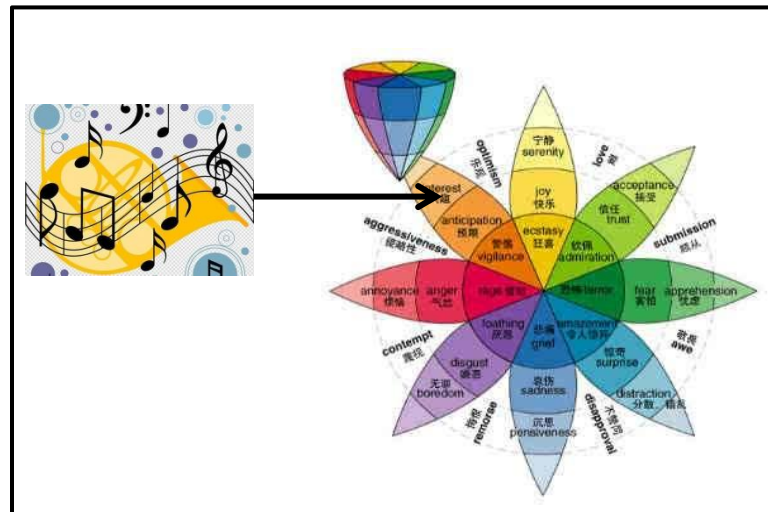
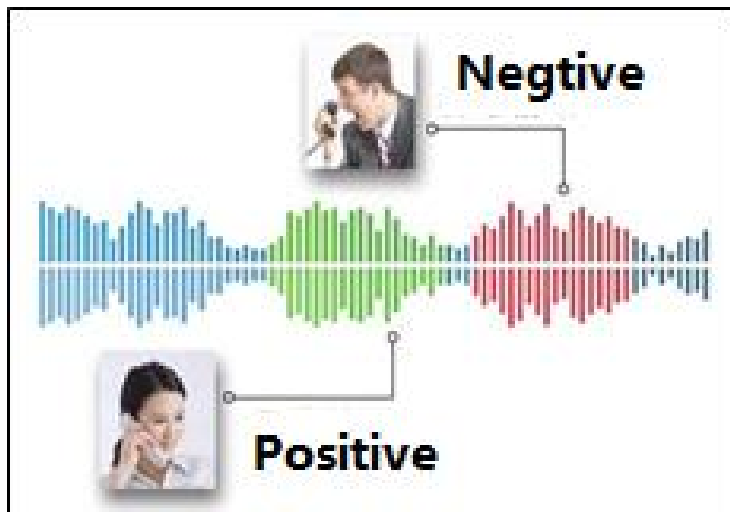
- Picard的情感HMM模型
- 分布式情感模型
- 基于人工心理的状态空间情感模型

# 音频情感识别

## ■ 音频：人类能够听到的所有声音

- 人声
- 音乐
- 环境音效

## ■ 音频中包含丰富的情感



## ■ 语音情感特征种类

- 韵律特征：最主要的语音情感特征，如语速、音量和音调等，例如发怒时都会增加；
- 音质特征：音频抖动（Jitter）和振幅抖动（shimmer），谐波噪声率，共振峰；
- 频谱特征：MFCC, LPCC。



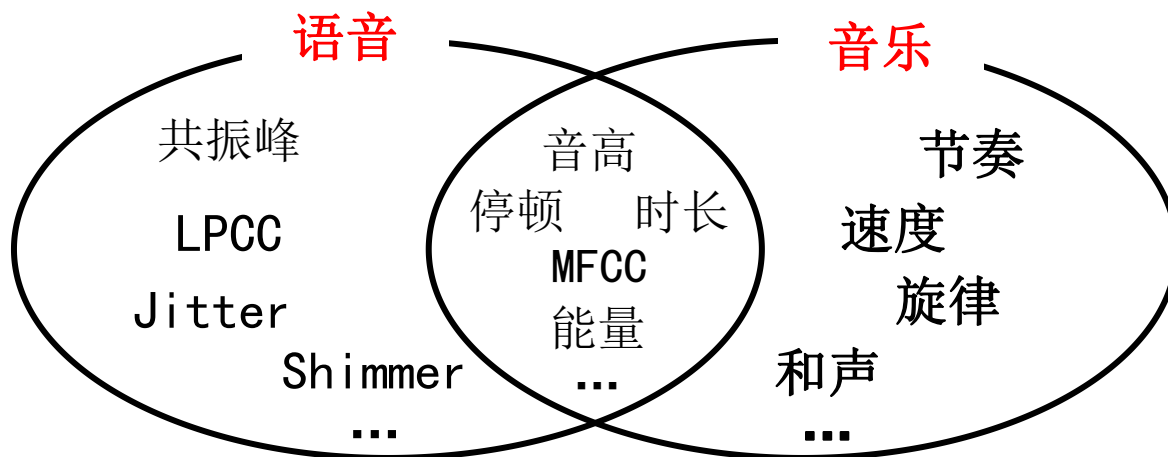
■ 功能性副语言中携带了大量情感信息

功能性副语言	高兴	伤心	惊讶	生气	害怕	厌恶
笑声	Y	N	N	N	N	N
伤心的哭声	N	Y	N	N	N	N
质疑声	N	N	Y	N	N	N
叫喊声	N	N	N	Y	N	N
害怕的哭声	N	N	N	N	Y	N
叹息声	N	N	N	N	N	Y

# 音频情感识别

## ■ 音乐情感特征

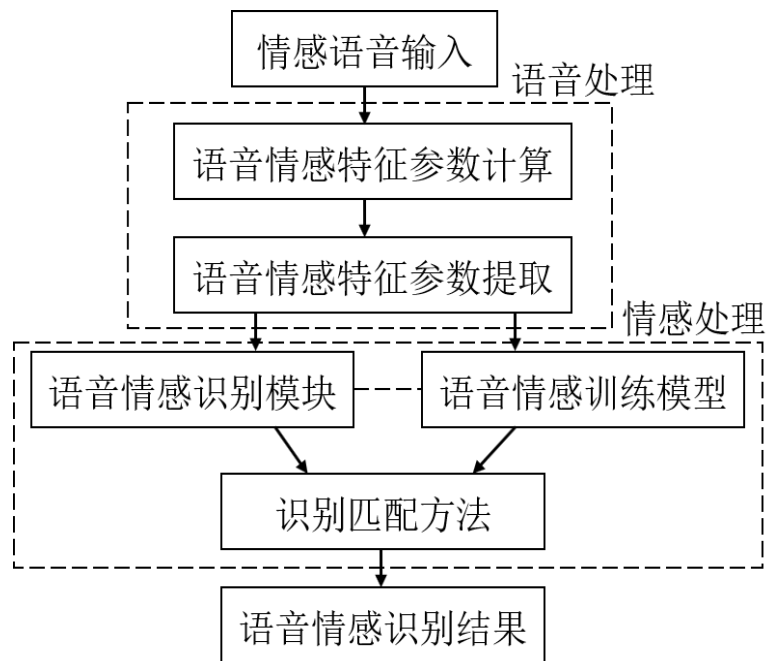
- 节奏 (Rhythm)
- 速度 (Tempo)
- 旋律 (Melody)
- 和声 (harmony)



# 音频情感识别

## ■ 音频情感识别本质上属于模式识别

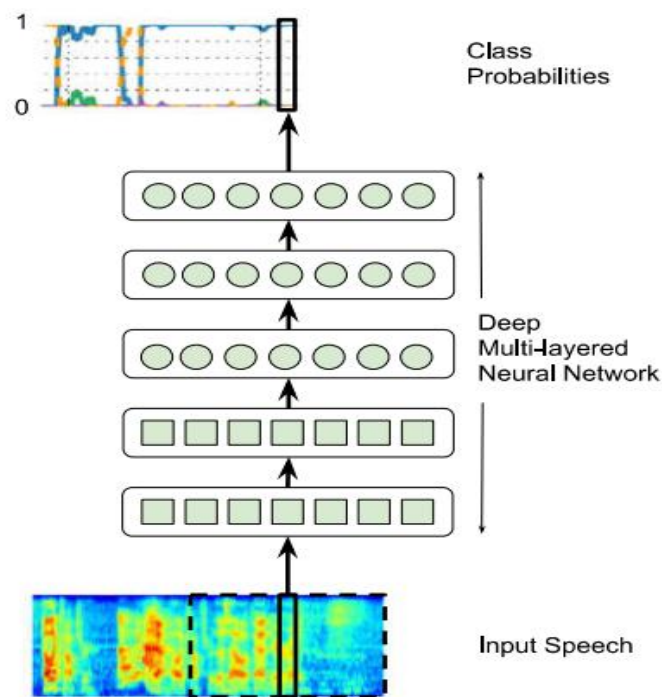
- 神经网络
- 高斯混合模型
- 隐马尔科夫模型
- 支持向量机
- 集成学习算法
- 深度神经网络



# 音频情感识别

## ■ 基于深度学习的音频情感识别模型

- 基于深度学习的音频情感识别模型输入是语音信号，经过多层网络结构，输出各个情感类别概率，得到预测结果。中间网络层可以是不同的网络结构，如**深度神经网络、卷积神经网络或循环神经网络**等。



# 表情识别

- 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

# 表情识别

---

- 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

# 表情识别

---

- 局部二值模式 (LBP)

- Gabor

- 主动外观模型 (ASM)

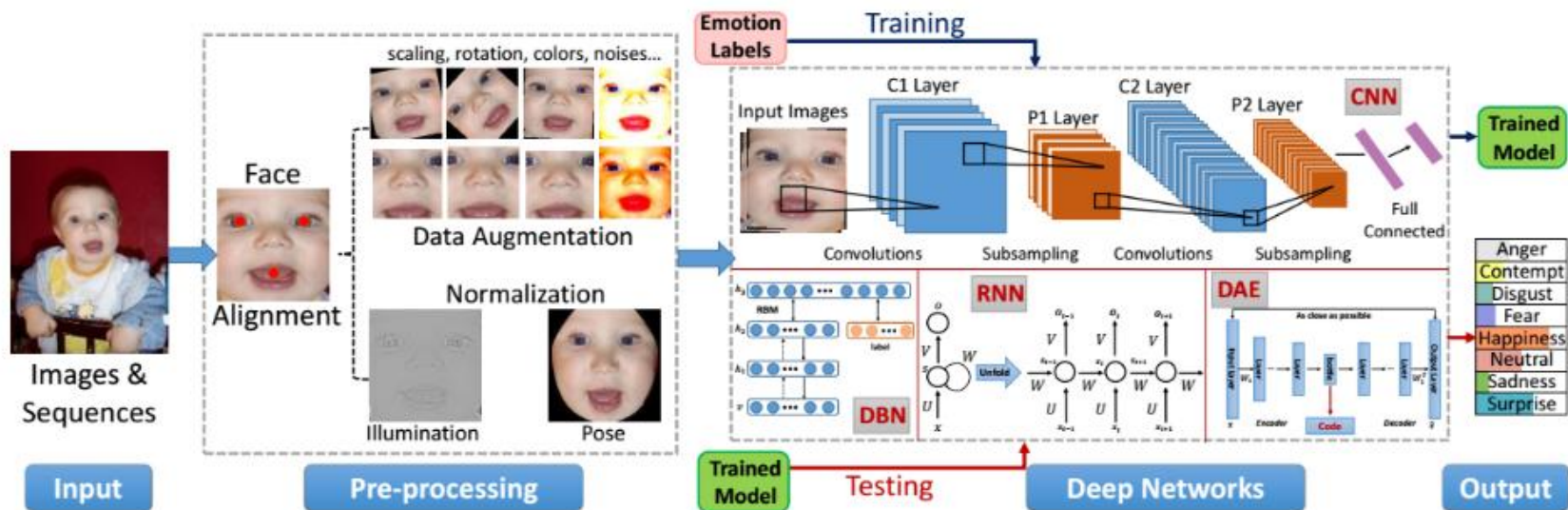
- 光流法

- 深度视频特征

# 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

# 微表情识别

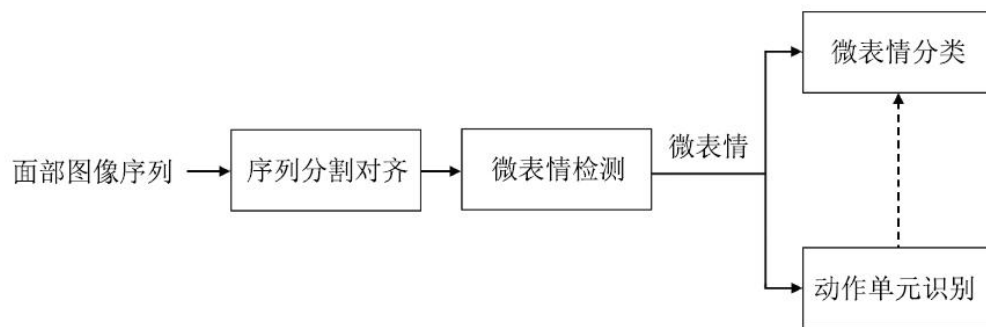
---

- 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

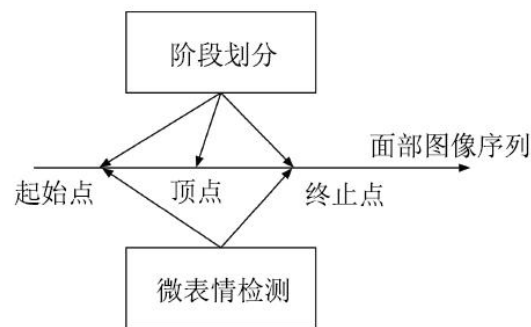
# 微表情识别

## ■ 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)

# 微表情识别

---

## ■ 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

# 肢体动作的情感识别

## ■ 肢体动作的情感表现

- 身体的位置和运动在感兴趣/厌倦和同意/不同意中有连续的表现并且可以识别。
- 相对于语音和人脸表情变化来说，**姿态变化的规律性较难获取，但由于人的姿态变化会使情感表述更加生动**
- 随着对情感计算及智能人机交互研究的不断深入，肢体动作识别也开始作为获取情感的一种手段被用于情感识别中。



# 生理信号的情感识别

---

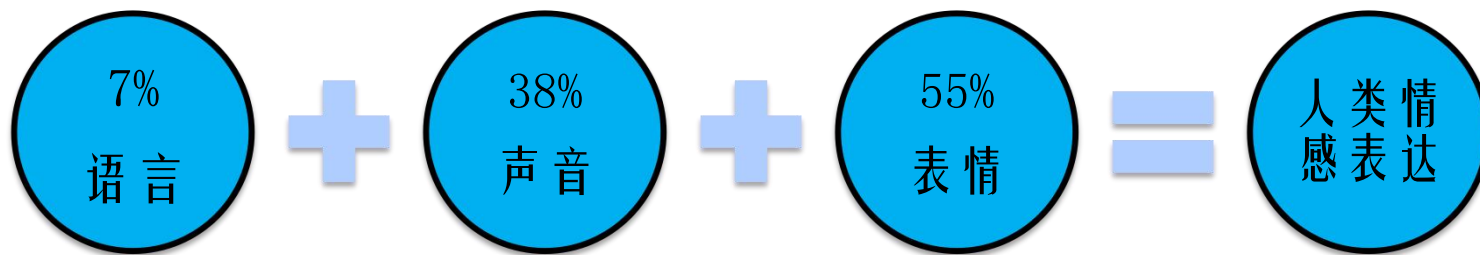
## ■ 基于生理信号的情感识别

- 从音视频中提取的特征难以排除被试者的主观因素的影响。
- 生理变化主要由人的自主神经系统和内分泌系统支配，**反应的信息更为客观。**
- 生理信号虽然能够比较客观精确的表现情感信息，但是多数**生理信号并不容易获取。**
- 生理信号主要包括皮肤电、肌电、血压、脑电、呼吸、脉搏、心电图等

# 多模态情感识别

## ■ 情感是通过多种模态的形式进行表达的

- 人们在高兴时说话节奏欢快，表现在说话的音调和语速上，同时面部会微笑，眯眼，此时语音和表情同时表达出高兴的情感状态。
- 当一个人难过时，往往不会怎么说话，情感识别难以单靠语音单模态信息，难过体现在表情上往往伴随着面部嘴角下垂、皱眉等。



# 多模态情感识别

---

## ■ 单模态情感识别鲁棒性不足、识别率不高

- 当人类主观上对情感信号加以掩饰或者单一通道的情感信号受到其它信号的影响时，情感识别性能将会明显下降，如面部表情容易被遮挡、语音容易受噪声干扰。

## ■ 多模态情感识别信息更加完整，更加符合人类对情感的感知模式

- 人在表达情感时，表现的方式往往是多种，不同表现方式在表达情感信息时存在一定的互补作用。



# 文本倾向性分析

## ■ 概念

- 情感分析（Sentiment analysis），又称倾向性分析，它是对带有情感色彩的主观性文本进行分析、处理、归纳和推理的过程；
- 识别用户对事物或人或一句话的看法、态度，即判别用户对评价对象所持有的情感倾向。



# 文本倾向性分析

---

## ■ 情感信息分类

### ■ 主客观信息的二元分类

### ■ 主观信息的情感分类

## ■ 主观信息的情感分类

### ■ 最常见的褒贬二元分类以及更细致的多元分类

### ■ 按照极性分类：正向，负向，中性

# 文本倾向性分析

---

- 文本情感特征分析方法主要包含四种
  - 基于情感词典的文本情感识别方法
  - 基于统计的机器学习文本情感识别方法
  - 基于深度学习的文本情感识别方法
  - 基于预训练模型的文本情感识别方法

# 文本倾向性分析

---

## ■ 舆情分析与情感分析的区别

- **用途不同：**基于产品评论的情感分析多用于商业，舆情分析多用于政府部门；
- **复杂性不同：**舆情分析是个比较复杂的系统，涉及更多的技术；
- **信息来源更广泛：**例如，新闻评论、BBS、聊天室、博客、RSS等。

# 情感语音生成

---

- 语音合成包括了“说了什么”即文本内容以及“怎么说的”即合成语音的态度与情感状态；
- 现有语音合成系统缺乏情感表现力；
- 情感语音合成方法
  - 情感焦点词
  - 基于场景驱动的情感语音合成
  - 基于韵律转换的情感语音合成
  - 基于说话人自适应的情感语音合成
  - 基于深度学习的情感语音合成

# 表情生成

---

- 表情生成的目的是通过某种表情计算方法产生出有表情的人脸图像
- 表情生成得到了计算机图形学、计算机视觉和模式识别领域的广泛关注
- 表情生成在人脸编辑、影视制作、社交网络和数据扩增方面应用广泛
- 合成高逼真度的人脸图像仍然是一个挑战性难题

# 表情生成

---

## ■ 表情生成方法

- 渐变法
- 表情映射
- 几何驱动
- 表情系数
- 五官移植
- 统计学方法

# 表情生成

---

## ■ 基于深度学习的表情生成方法

- PixelRNN
- GAN
- GAN部分变体



# 多模态情感生成

---

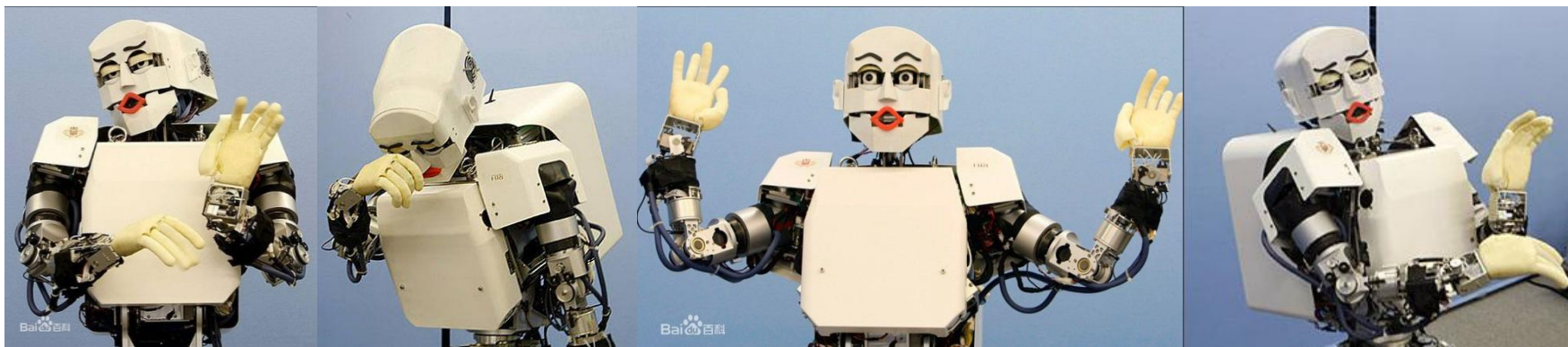
- 个体情感的表达是可以从多个模态（语音、面部表情以及生理信号等等）感知出来的
- 语音和面部表情动作是最容易感知的模态形式
- 单模态情感生成是多模态情感生成的基础
- 多模态情感生成需要根据情感的变化使得音视频信息呈现一致
- 多模态情感生成需要考虑不同模态之间的同步问题

- 情感交互是高级信息时代人机交互的主要发展趋势；
- 情感交互使人机交互可以像人与人交互一样自然、亲切、生动、富有情感；
- 情感交互可以从人脸表情交互、语音情感交互、肢体行为情感交互、生理信号情感交互以及文本情感交互等方面进行探索。
- 当前在情感交互领域，我国已经率先发布了ISO国际标准。

# 情感计算的应用

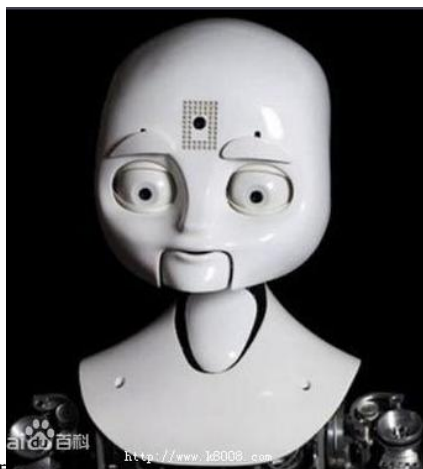
## ■ 情感机器人

2009年，日本早稻田大学， KOBIAN， 能够表达喜怒哀乐等7种表情



麻省理工学院媒体实验室 Nexi

日本软银 Pepper, 2014



# 情感计算的应用

## ■ 教育业

- 情感识别技术应用于远程教学手段中可以扩大“专家”教学范围。
- 情感识别技术的教学软件通过监控学习者对软件内容的反应，在远程教学过程中，适当地调整教学进度和难度。



# 情感计算的应用

## ■ 信息业

- 在信息爆炸的时代，信息检索尤其重要。除了现在常用的文本检索，情感识别和标注将是多媒体检索中不可缺少的一项重要技术。

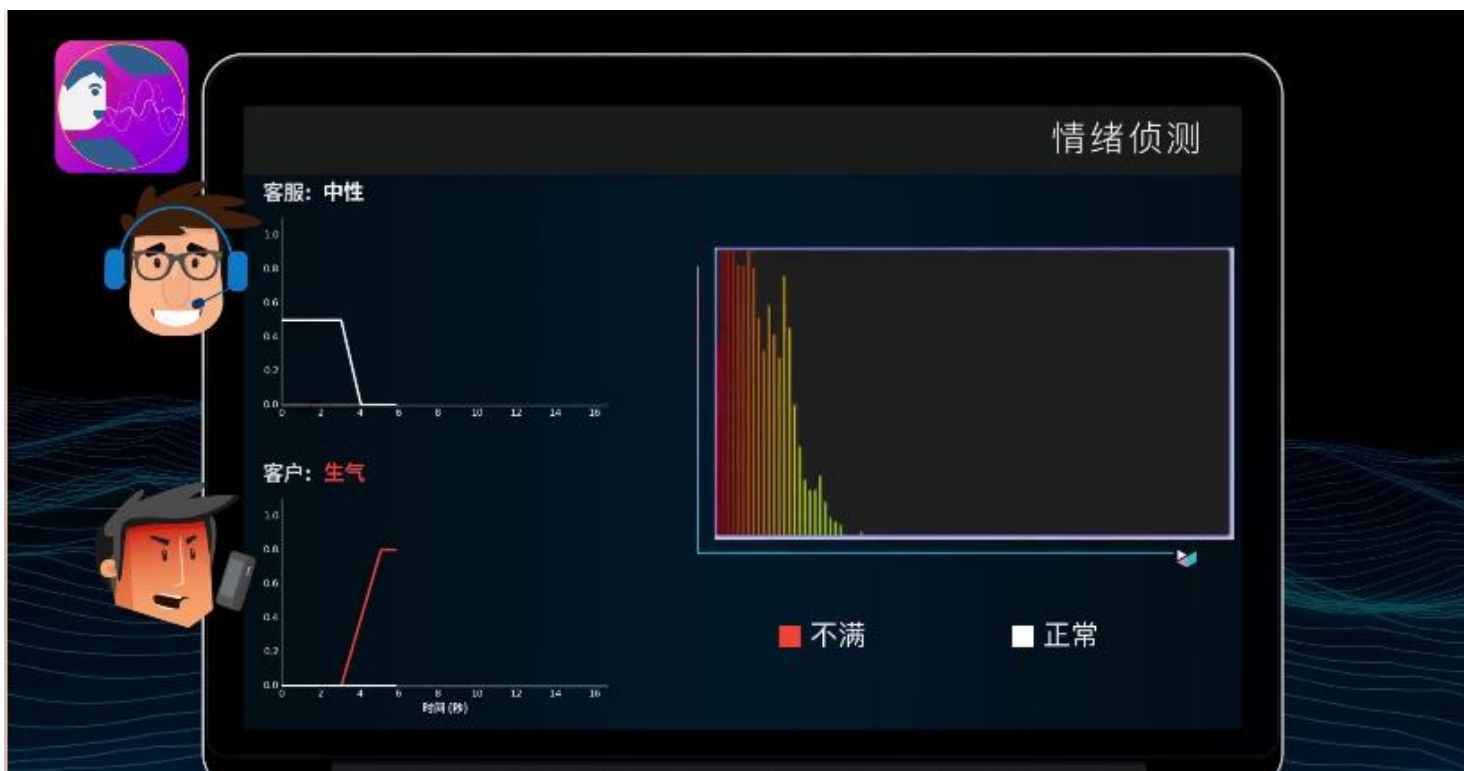




# 情感计算的应用

## ■ 服务业

- 情感识别技术应用于自动呼叫中心，将更加人性化，它能够通过理解客户的“画外音”及时有效地做出变化，最大限度地保留住可能的客户资源。



# 情感计算的应用

## ■ 娱乐业

- 情感识别技术应用于点播系统，能对广播电视节目进行情感标注，根据用户提交的情感需求做出合理的响应。用户能随心所欲地看到“高兴”或“难过”的节目。



苏宁联合竹间智能在CES上发布了虚拟AI购物助手Sue

# 情感计算的应用

## ■ 刑事与安全

- 语音情感识别技术应用于测谎系统，从语音信号中分析被测者的情绪变化，进而判断其言语的可靠性，辅助刑事与安全专家分析。

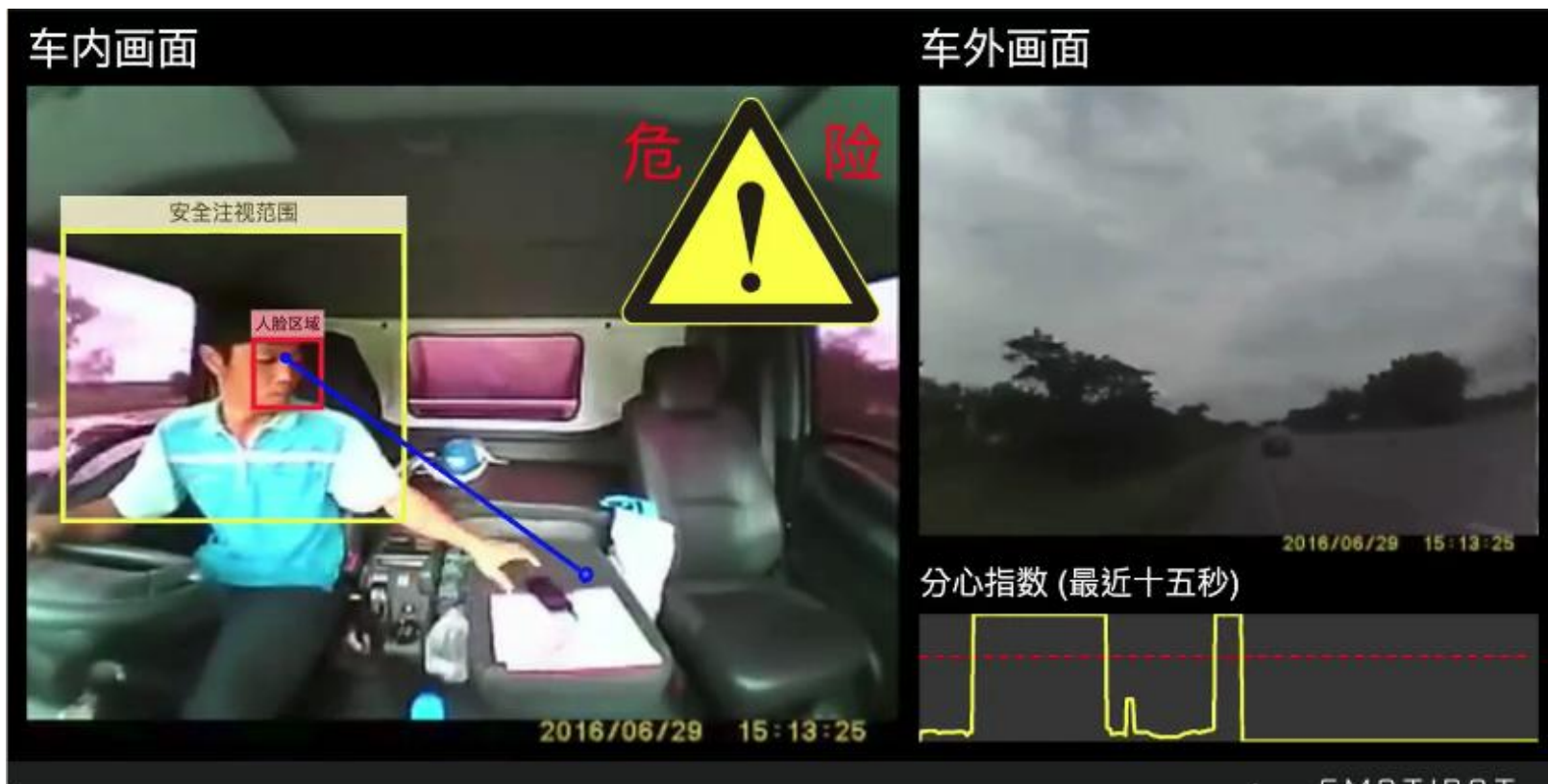




# 情感计算的应用

## ■ 工业

- 情感识别技术应用于汽车驾驶员疲劳检测，可以做出相应的保护和提醒措施。



# 情感计算的应用

## ■ 舆情分析

- 及时掌握社会的情感动态对于维护社会稳定具有重要的意义，对企业而言，及时了解和掌握消费者对信息的反应也可以从中找到机会创造商机。



---

# 讨论&问答