

## 实验题五、情感语音的声学属性分析

### 背景知识：

语音是个人情感的一种重要表达方式（其他表达方式还有表情、姿态等）。通过分析语音的一些信号特性/声学属性，可以从中获取说话人的情感状态。将情感数字化以使之可计算的办法，有基于情感类别的“情感范畴观”思想（类似于颜色的红绿蓝等类别），也有基于多维空间坐标的“情感维度观”思想（类似于表示颜色的 RGB 值）。考虑到后者在测量和理解上的困难，前者在声音信号的分析中应用较多。

### 任务内容：

从“1-高兴、2-着急、3-愤怒、4-伤心”等4种情感类别中，**挑选2种语音情感类别**，从互联网音视频网站发布的访谈栏目、对话聊天、现场采访、电话采访等形式的视频中，收集并抽取带有指定情感的语音片段。然后，**对其声音属性进行分析，主要声学属性包括：音量大小（信号幅度）、各频带的平均能量（按20ms为单位进行计算DFT频谱，然后平均划分成30个频带）。**

为了使分析在统计上有意义，要求：每段情感语音片段应该仅包含一种情感类别；**每种情感至少收集10小段；片段长度2秒~6秒，语义内容基本完整**；语音的情感类别应保证在分割成小片段后仍然能比较容易地确定其类别，并与它在原始音频流中时的判断结果一致。

### 数据标注：

每个小段的话者信息（某甲、某男等即可）、情感种类（自己选择的2种之一）、性别、大致年龄（1-少、2-青、3-中、4-老）；统一保存为16K采样，16位精度，单声道的WAV文件格式（有些软件工具或设备录制的是MP3，可以通过一些音频工具转成WAV格式）。

命名格式：X-G-A.wav，其中：X为1-4是情感编号；G为性别，1男声，2女声；A为年龄段标号。如2-1-3.wav表示着急语音、男声、中年人。

本次实验收集数据，将在后续实验中继续被使用，**请自己妥善保存。**

### 提交要求：

1. 考虑到语音情感表达的真实性，应从任务中指定的音频形式中收集，**不能从影视剧、话剧、小品等表演形式的音频录音中截取**；
2. 不得借用他人收集提交的音频素材；（这个算抄袭）
3. 允许请他人协助评判自己搜集语音片段的情感类别是否符合要求。