atitit 音频 项目 系列功能表 音乐 v3 t67.docx

Atitit 音频 项目 系列功能表

音频 音乐 语言领域的功能表

听歌识曲功能 酷我功能。

铃声 功能。。

音频切割（按照副歌部分。从歌词副歌部分拉取。。Nlp识别副歌

音效功能

歌曲与语音内容识别

Dj调速

Dj机循环功能

混音(dj版 音效版

音频混音特效

CUE点管理功能(多个cue点

视频音频分离

音频音乐人声类型分类

音乐去人声，得到伴奏

音乐歌词识别（获取到歌词，爬虫法与语音识别法

音频处理算法、音视频编解码算法

音乐信息结构化方法

音乐分类（seoin cintswi ，，dwicheolei

音乐信息检索算法,（根据歌名，歌手 ，年代，歌词搜索

音乐歌词类统计分析（高频词云分析

语音识别（asr）

Midi类，简谱类

格式研究mp3 m4a

音频混音

目录

[1.1. ACR(自动内容识别) > 2](#_Toc30886)

[1.2.  3. 音频切割： 3](#_Toc1417)

[1.2.1. 原理 3](#_Toc4139)

[2. 歌曲识别 4](#_Toc17875)

[3. 常用类库 7](#_Toc10212)

[3.1. ACRCloud ffmpeg 7](#_Toc11655)

[4. 常见项目要求的可能功能 7](#_Toc14921)

[4.1. 音频类功能缩写术语 NS,VAD,AGC,AEC ANS 7](#_Toc5998)

[4.2. DRC Dynamic Range Control(DRC)动态范围控制 8](#_Toc8304)

[5. ref 8](#_Toc25613)

## [ACR(自动内容识别)](http://baike.baidu.com/subview/267733/18948545.htm" \t "http://cn-console.acrcloud.com/_blank) >

优秀的企业和开发者提供最先进的 [ACR(自动内容识别)](http://baike.baidu.com/subview/267733/18948545.htm" \t "http://cn-console.acrcloud.com/_blank) >技术， 使客户的产品获得诸如"听音识曲"，"哼唱识曲"，["摇电视"](http://baike.baidu.com/item/%E6%91%87%E7%94%B5%E8%A7%86" \t "http://cn-console.acrcloud.com/_blank)，"版权检测"，"视频识别"，"电台监控"，"广告监控" 等能力。

## 3. 音频切割：

　　ffmpeg -i "sourceFile" -ss startTime -to endTime -y "targetFile"（按起点和终点切割）

　　ffpmeg -i "sourceFile" -ss startTime -t duration -y "targetFile"（按起点和持续时间切割）

VAD音频切分工具 静音区域获取

### 原理

一个FRAME时长默认10s，通过webrtc的vad计算这个FRAME是否是活动（ACTIVE: 有声音， INACTIVE：静音）。

这样可以获得音频的所有FRAME的活动值（ACTIVE或者INACTIVE）。从而在静音（INACTIVE）的音频段上的切分音频。

// [5150, 5220) [5230, 6380) [6520, 6970) [7000, 8040) [8080, 8670) // 表示 [5150, 5220) 等区间内是有声音的，其它区间均为静音， // 如果需要切割的话，应该在静音区间选择合适的位置切割

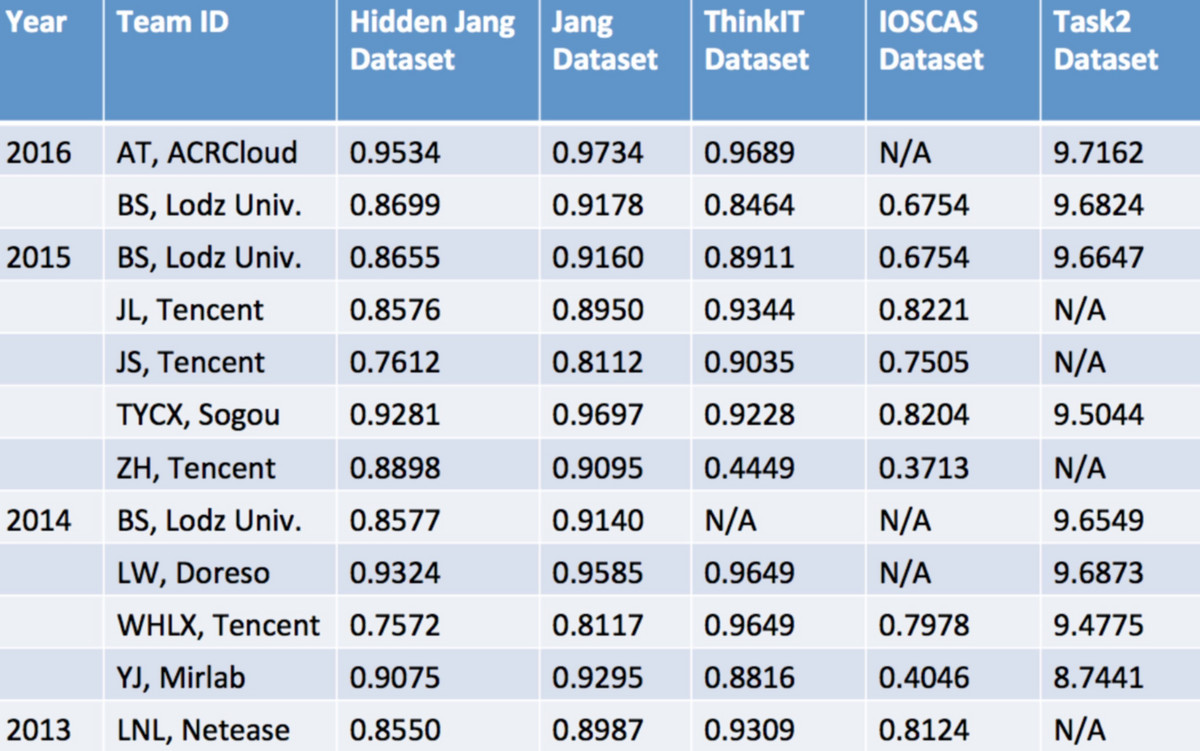
# 歌曲识别

歌曲识别技术分为歌曲原声识别以及哼唱识别。歌曲原声识别通过听筒收集音乐播放信息，生成音频指纹，在曲库中识别到对应的歌曲。 哼唱识别通过用户对着话筒哼唱小段歌曲，系统自动识别并检索出所哼唱的歌曲

音频识别与“哼唱识别”早已不是新鲜技术，早在 2008 年，盛大游戏推出的音乐类竞技网游《巨星》便已经将这类技术搬到了游戏产品中，但由于时间点不合适且运营、游戏模式太重等原因，这款游戏仅运营三年左右便宣布停服。

36氪近期接触到的****「ACRCloud」成立于 2015 年，是一家音频识别技术研发商，****先后研发了****音乐识别、哼唱识别、直播互动触发、点播视频识别****等多项自动识别服务，为各大互联网运营商、广电合作伙伴提供内容识别平台。

事实上，早在十年前便产品化的音频识别技术并非新鲜事，全球音乐信息检索竞赛 MIREX（Music Information Retrieval Evaluation eXchange）也是每年举办一次，历年参与比赛的公司也包括酷狗、腾讯、搜狗、网易等，而 ACRCloud 多次拿下过哼唱识别、音频指纹识别的第一名，同时也保持着该比赛历年最好成绩的记录。



MIREX比赛结果图表

联合创始人李蕴博告诉36氪，****ACRCloud 的核心团队来自于 Shazam、当当网、中科院、华为、电影网等公司及科研院校****，除了****算法****方面有多年储备以外，在****音乐、视频行业****也有多年工作经验，所以比较了解客户的痛点和需求。

例如，网易云音乐、唱吧、小米通过使用 ACRCloud 的服务实现哼唱识别歌曲；Deezer, Anghami, KKBOX 等众多海外音乐服务商使用 ACRCloud 的听歌识曲功能；电视台会基于音频识别完成广告和音乐版权内容的监测、收视调研、电视节目和广告互动等工作。



值得一提的是，近期****针对国内外音乐社交、音乐游戏等客户的关于用户演唱水平自动评价的集中需求****（如音遇的AI识别等功能）****，ACRCloud 技术团队经研发，已完成**[支持基于乐句哼唱 AI 识别、客观评价打分的新引擎](http://www.acrcloud.cn/humming-sdk" \t "https://36kr.com/p/_blank)**（点击内链可查看 DEMO 测试效果及合作方式），并可根据客户产品需要，为其定制其专有曲库，帮助产品快速上线并降低曲库运营成本。

ACRCloud 的服务是面向全球用户的，公司也为此在中国、亚洲地区、欧洲地区、美国地区分别建设了一个服务器节点。****截至目前，ACRCloud 开放注册用户量达 3 万，付费用户超过 2000 例，覆盖了超过 180 个国家。****

ACRCloud 定位于企业服务公司，所以公司主要以 SaaS 形式为各行业用户提供音频识别服务，而用户会利用这种技术设计制作不同产品，而 ACRCloud 会收取相关服务费。李蕴博表示，****“随着音遇的走红，可能会给音频识别行业带来新的增长点，所以明年的营收增长率有希望更高。”****

ACRCloud 创立初期的启动资金 100 万，团队拥有 10 多名员工。公司创立第一年便实现收支平衡，收入年增速约为 100%，现阶段暂无融资计划。

——————————

# 常用类库

## ACRCloud ffmpeg

# 常见项目要求的可能功能

## 音频类功能缩写术语 NS,VAD,AGC,AEC ANS

AGC是自动增益补偿功能（Automatic Gain Control），AGC可以自动调麦克风的收音量，使与会者收到一定的音量水平，不会因发言者与麦克风的距离改变时，声音有忽大忽小声的缺点。

ANS是背景噪音抑制功能（Automatic Noise Suppression），ANS可探测出背景固定频率的杂音并消除背景噪音，例如：风扇、空调声自动滤除。呈现出与会者清晰的声音。

AEC是回声消除器（Acoustic Echo Canceller）,AEC是对扬声器信号与由它产生的多路径回声的相关性为基础，建立远端信号的语音模型，利用它对回声进行估计，并不断地修改滤波器的系数，使得估计值更加逼近真实的回声。然后，将回声估计值从话筒的输入信号中减去，从而达到消除回声的目的，AEC还将话筒的输入与扬声器过去的值相比较，从而消除延长延迟的多次反射的声学回声。根椐存储器存放的过去的扬声器的输出值的多少，AEC可以消除各种延迟的回声。

WebRTC中的音视频引擎、音频处理算法、音视频编解码算法的研究与实现

有扎实的网络技术基础，对socket通信、UDP/TCP/IP和HTTP有深刻的理解

5、熟悉WebRTC、FFmpeg、licode、kurento、janus、mediasoup等音视频工具

6、熟悉H264、H265、Opus、VP8等编解码，熟悉rtp、rtmp、rtsp、sip等传输协议

7、有语音的相关算法优化经验者优先，如NS，VAD，AGC，AEC等

## DRC Dynamic Range Control(DRC)动态范围控制

提供压缩和放大能力，可以使声音听起来更柔和或者更大声，即一种信号幅度调节方式。

# ref

歌曲识别 - 语音识别 - 讯飞开放平台.html