



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

音频指纹技术在流媒体音乐平台的应用

网易云音乐 刘华平
2019.8.23-24

出品: LiveVideoStack CSDN
— 音视频技术社区 —



深圳
2019

遨游“视”界 做你所想
Explore World, Do What You Want

LiveVideoStackCon 2019 深圳

2019.12.13-14



成为讲师: speaker@livevideostack.com

成为志愿者: volunteer@livevideostack.com

赞助、商务合作: kathy@livevideostack.com

出品:  **LiveVideoStack** 
—— 音视频技术社区 ——

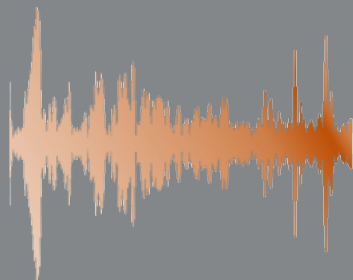


北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

- 音频指纹算法介绍
- 音频指纹工程化
- 应用场景
- 技术进阶展望

音频指纹算法介绍



Audio signal



Digital fingerprint

代表音乐重要声学特征的基于内容的紧致数字签名，其主要目的是建立一种有效机制来比较两个音频数据的感知听觉质量。

音频指纹与哈希函数区别：按照听觉相似性原则应该视为相等，由波形数据计算出的Hash值完全不同。音频指纹希望听觉相似的音频数据产生相似的音频指纹。 音频指纹 & MD5



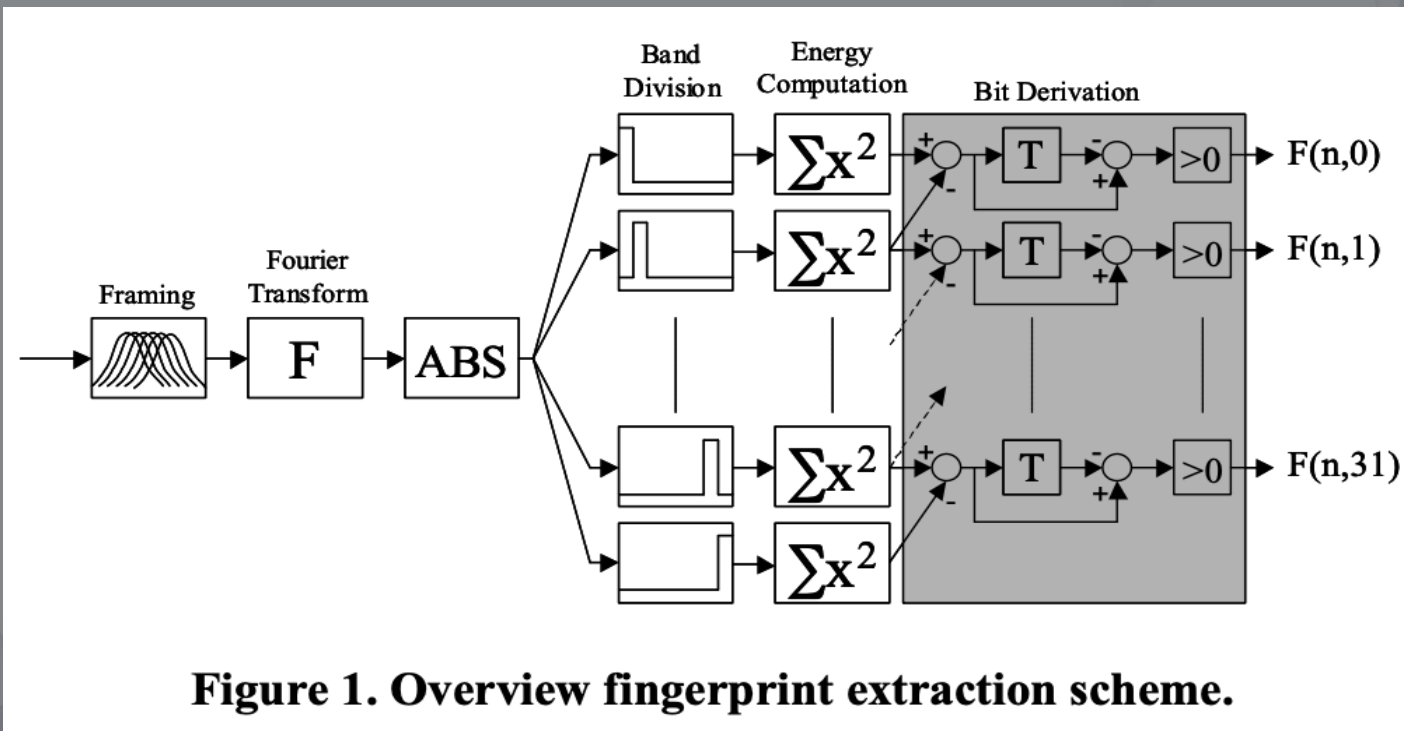
北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

音频指纹算法性质

算法研究层面	用户体验层面	产品运营
准确度	准确	
鲁棒性		
区分度		
可靠性		
粒度	速度	
指纹尺寸		
搜索速度		
可伸缩性		系统建设难度和成本
通用性		
安全性		

主流音频指纹算法1: 2002, Jaap Haitsma A Highly Robust Audio Fingerprinting System (Philips Audio Fingerprinting)

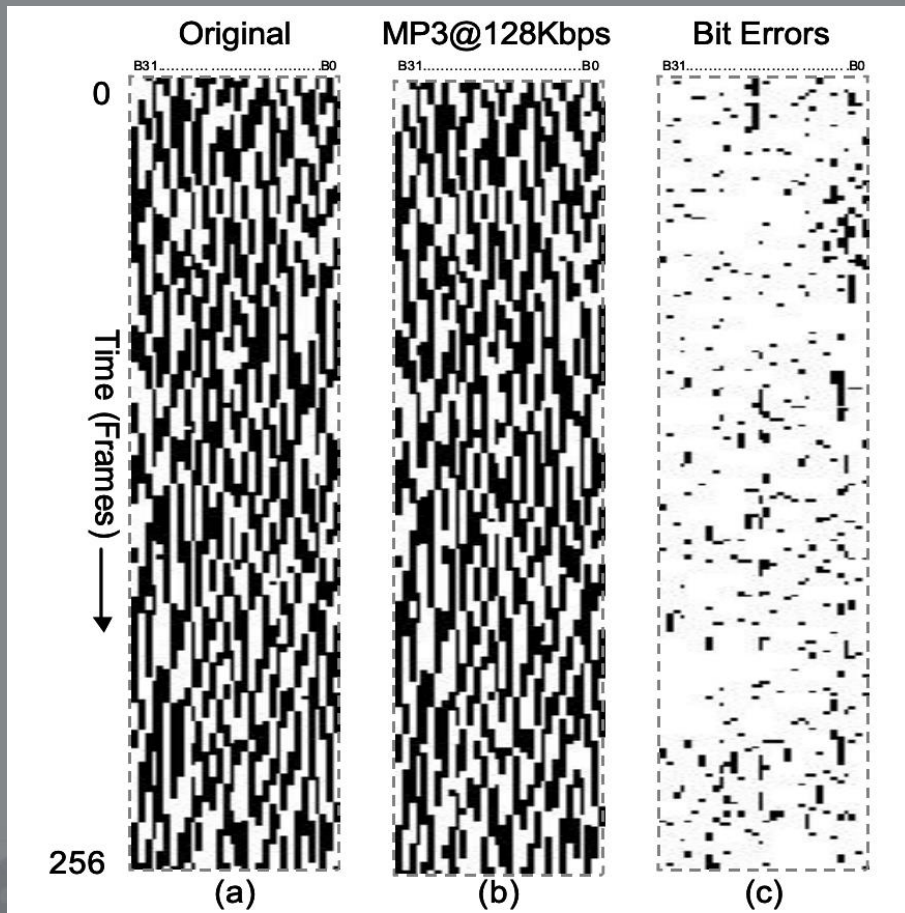


音频指纹算法介绍



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want



音频指纹算法介绍



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

主流音频指纹算法2: Avery Wang, An Industrial Strength Audio Search Algorithm (Shazam Audio Fingerprinting)

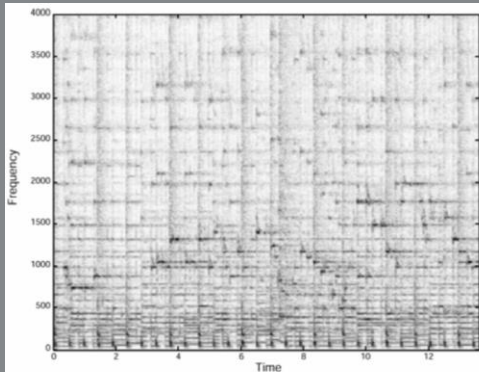


Fig. 1A - Spectrogram

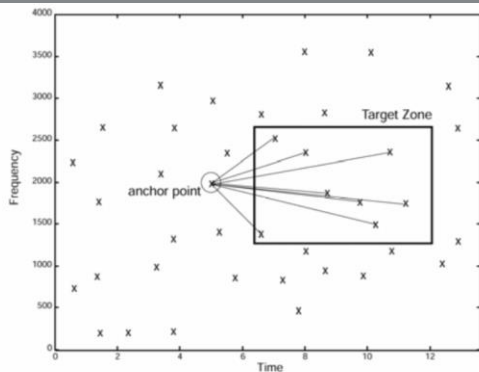


Fig. 1C - Combinatorial Hash Generation

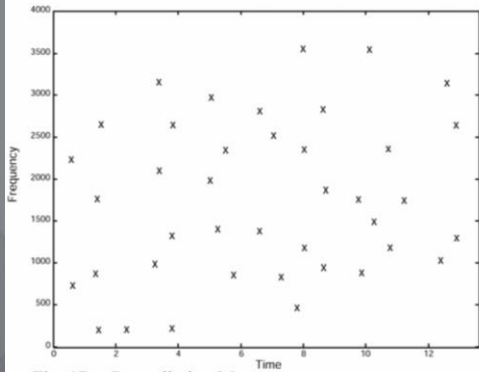


Fig. 1B - Constellation Map

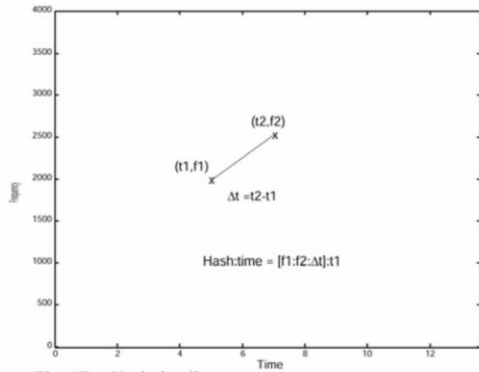


Fig. 1D - Hash details

指纹格式:

[“frequency of the anchor”;
frequency of
the point”; “delta time
between the anchor and the
point”]

量化位数:

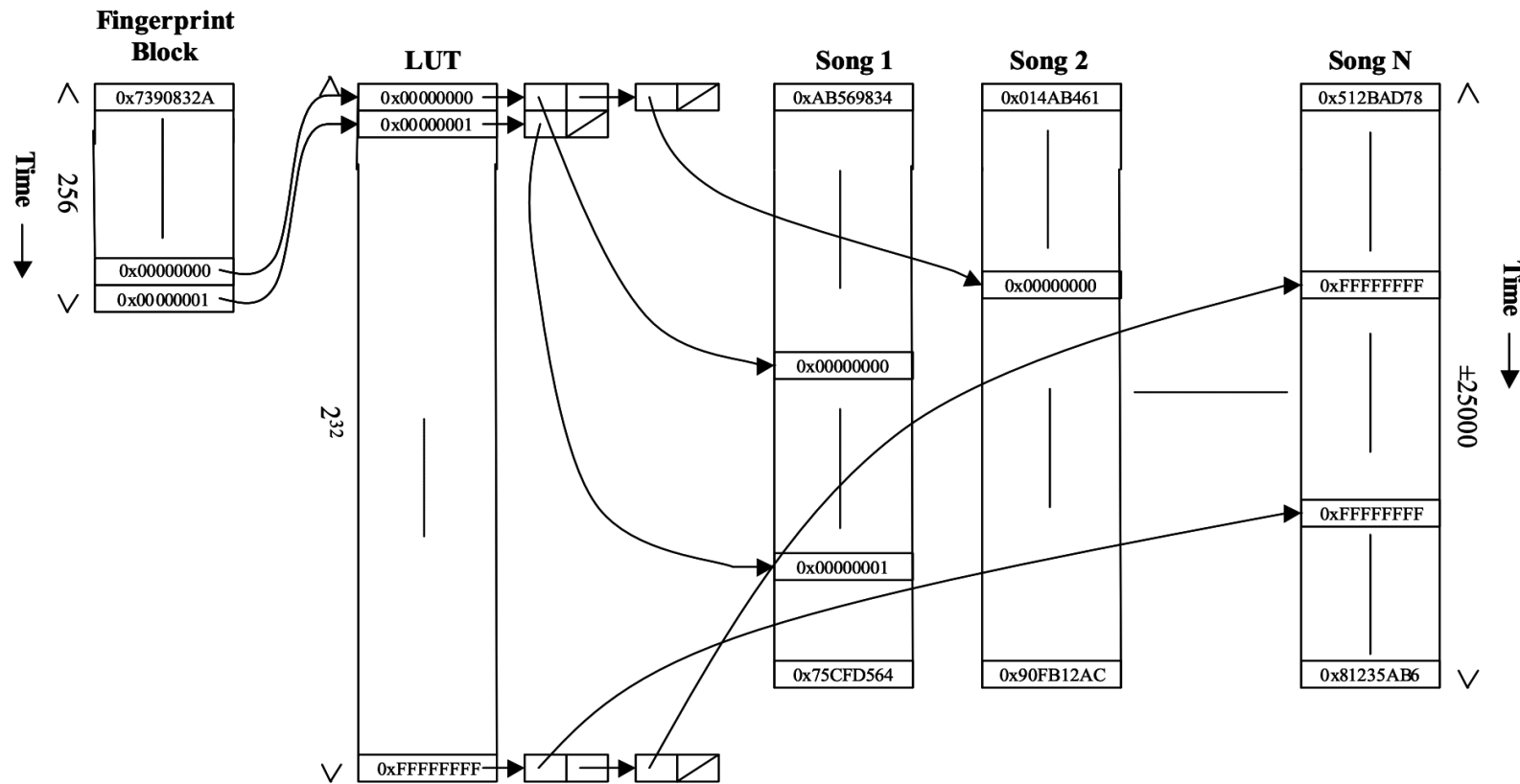
[“frequency of the anchor” 9bits ;
“frequency of the point” 9bits ;
“delta time between the anchor
and the point” 14bits] : 32bits

音频指纹算法介绍



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want



音频指纹算法比较



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

	优点	缺点
Philips	指纹库小 算法简单 系统建设成本低	抗噪性差 变速变调失效 系统适应性、可扩展性低
Shazam	抗噪好 变速变调适用 系统可扩展性强	指纹库大 算法调优复杂 对指纹库质量要求高 对节奏舒缓的歌曲性能较差 系统建设成本高

音频指纹工程化



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

听歌识曲

信息补全

音质升级

直播导流

版权保护

曲库清洗

音乐发现

音纹指纹业务中台

音频指纹数据
库

音乐信息库
metadata

音频指纹库
数据管理

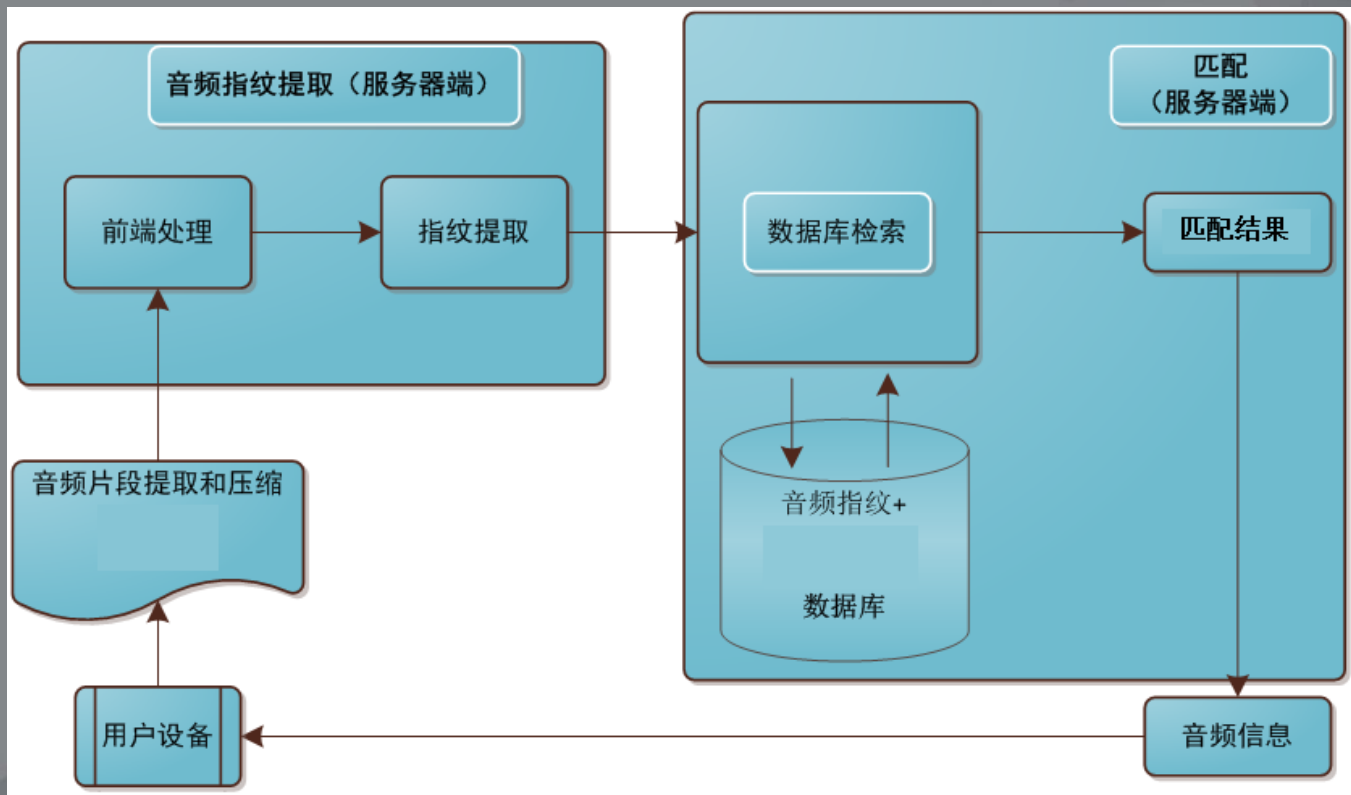
音频指纹库
线上运营

音频指纹数据流程图



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want





北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

应用场景：通用

- **Broadcast Monitoring(广播监控):** 对广播内容进行监控, 例如广告商希望他的广告片能按时播放; 艺人希望播放他的作品得到应有的报酬; 音频内容创作者希望版权保护。
- **电视第二屏互动(手机摇一摇):** 手机APP通过运用音频指纹技术的识别, 仅需几秒钟的录音即可识别当前播放的频道和节目进度, 从而方便应用开发者制作出与节目相关的互动和应用。
- **版权保护:** 全球知名网站Youtube 和Facebook 都是通过音频指纹技术建立的Content ID系统将受版权保护的内容指纹与用户新上传的内容进行匹配, 可快速检索是否含有侵权内容

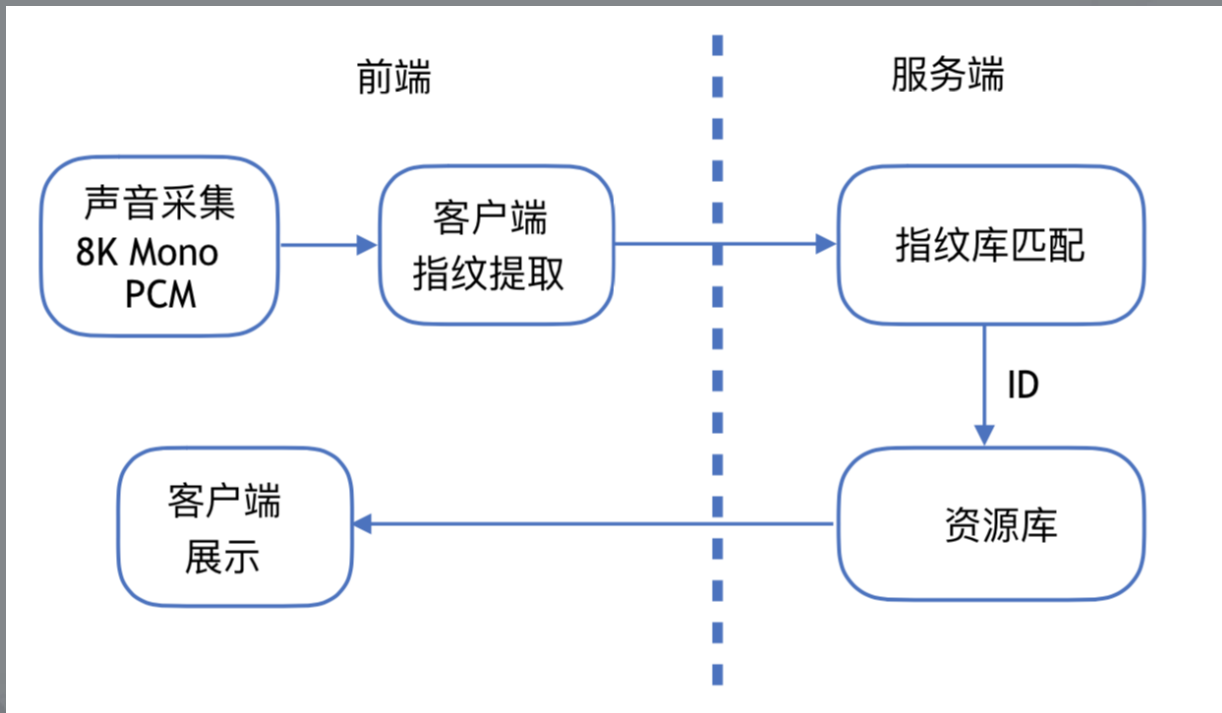


北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

应用场景：流媒体音乐平台

场景1：听歌识曲



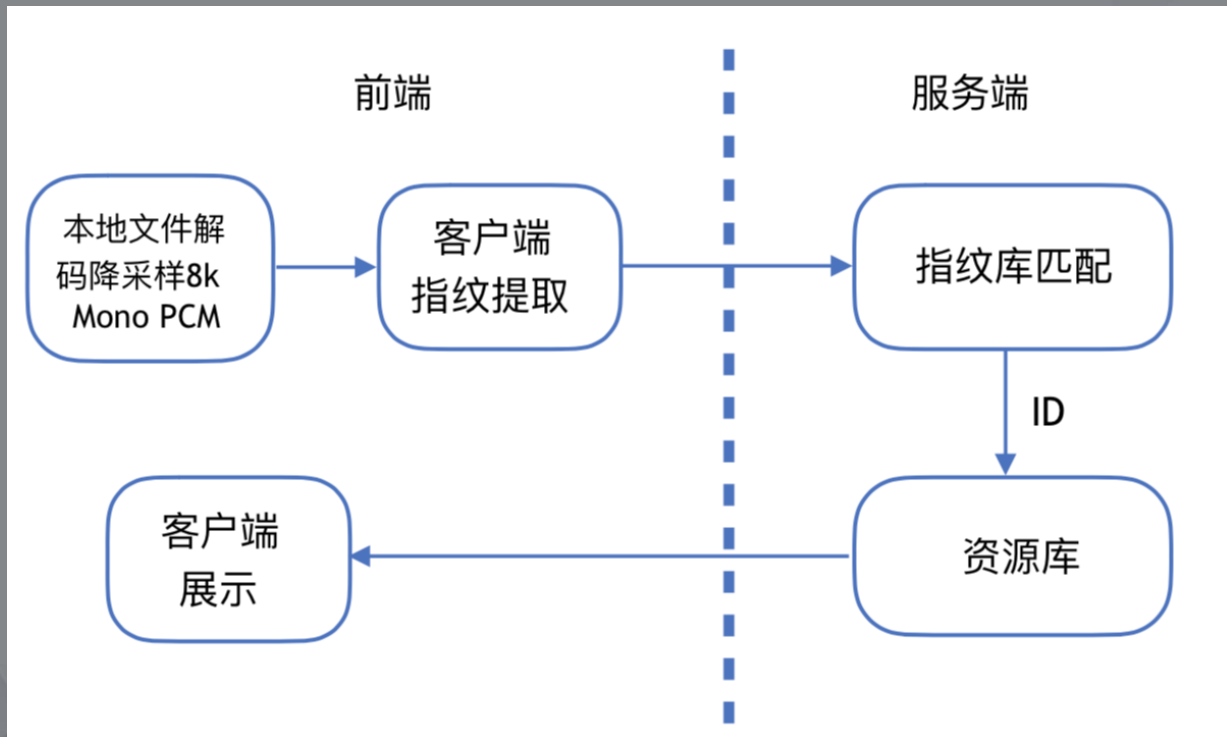
应用场景：流媒体音乐平台



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

场景2：本地歌曲信息补全



有的本地歌曲没有ID3或ID3信息不全，会通过音频指纹来识别，再向服务器去拉取 相关消息来补齐，包括：词、图相关信息，基本流程类似于线上的听歌识曲。

应用场景：流媒体音乐平台

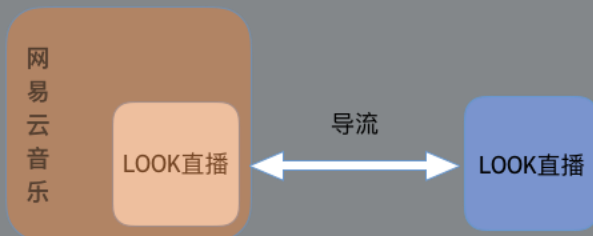


北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

场景3: K歌与直播导流

问题：怎么实现用户受欢迎的导流？



解决步骤：

- 1) LOOK 直播中主播是否在唱歌？
- 2) 唱什么歌？
- 3) 导流策略；

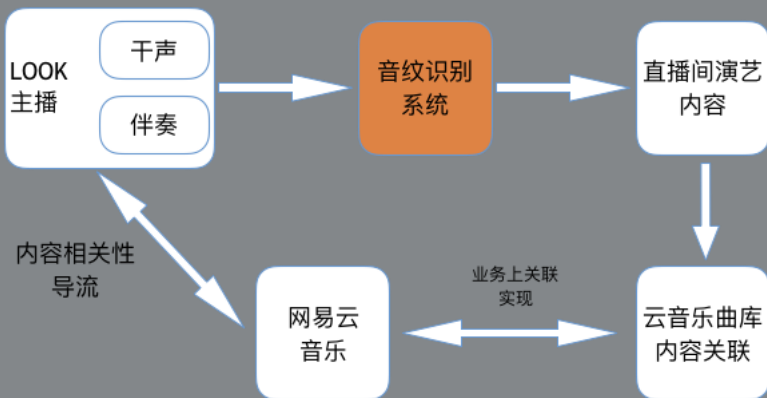
应用场景：流媒体音乐平台



北京
2019

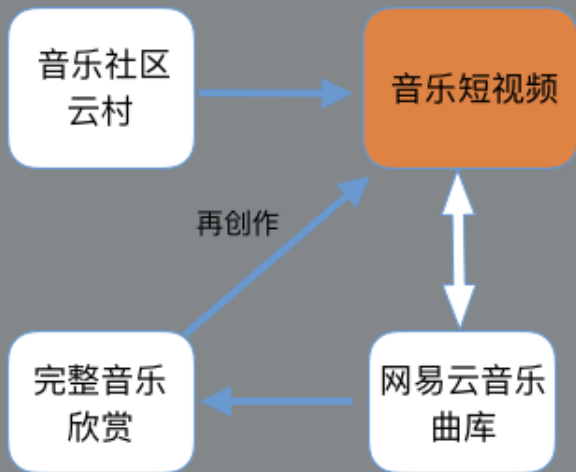
遨游“视”界 做你所想
Explore World, Do What You Want

场景3: K歌与直播导流



应用场景：流媒体音乐平台

场景3: Mlog BGM关联



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want





北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

技术进阶展望

- 无法识别同一音乐的不同版本--- 相似度;
- 提效率降成本;
- 算法鲁棒性;
- 音频指纹业务精细化运营;



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

参考文献

- 1、李伟,李晓强,陈芳,王淞昕. 数字音频指纹技术综述[J]. 小型微型计算机系统,2008,(11):2124-2130. [2017-08-14]
- 2、2002, Jaap Haitsma A Highly Robust Audio Fingerprinting System
- 3、Avery Wang, An Industrial Strength Audio Search Algorithm



北京
2019

遨游“视”界 做你所想
Explore World, Do What You Want

Thank you



出品: LiveVideoStack CSDN
—— 音视频技术社区 ——