



技术开启新“视”界

Technology Bring New Vision

2018.10.19-20 北京丽亭华苑酒店

LiveVideoStackCon²⁰¹⁸ 讲师热身分享会

9月6日19:30 准时开始

《短视频App的架构设计》



展晓凯

全民快乐 研发高级总监

11



关注LiveVideoStack公众号
回复『热身』获得进群方式

LiveVideoStack
— 音视频技术社区 —

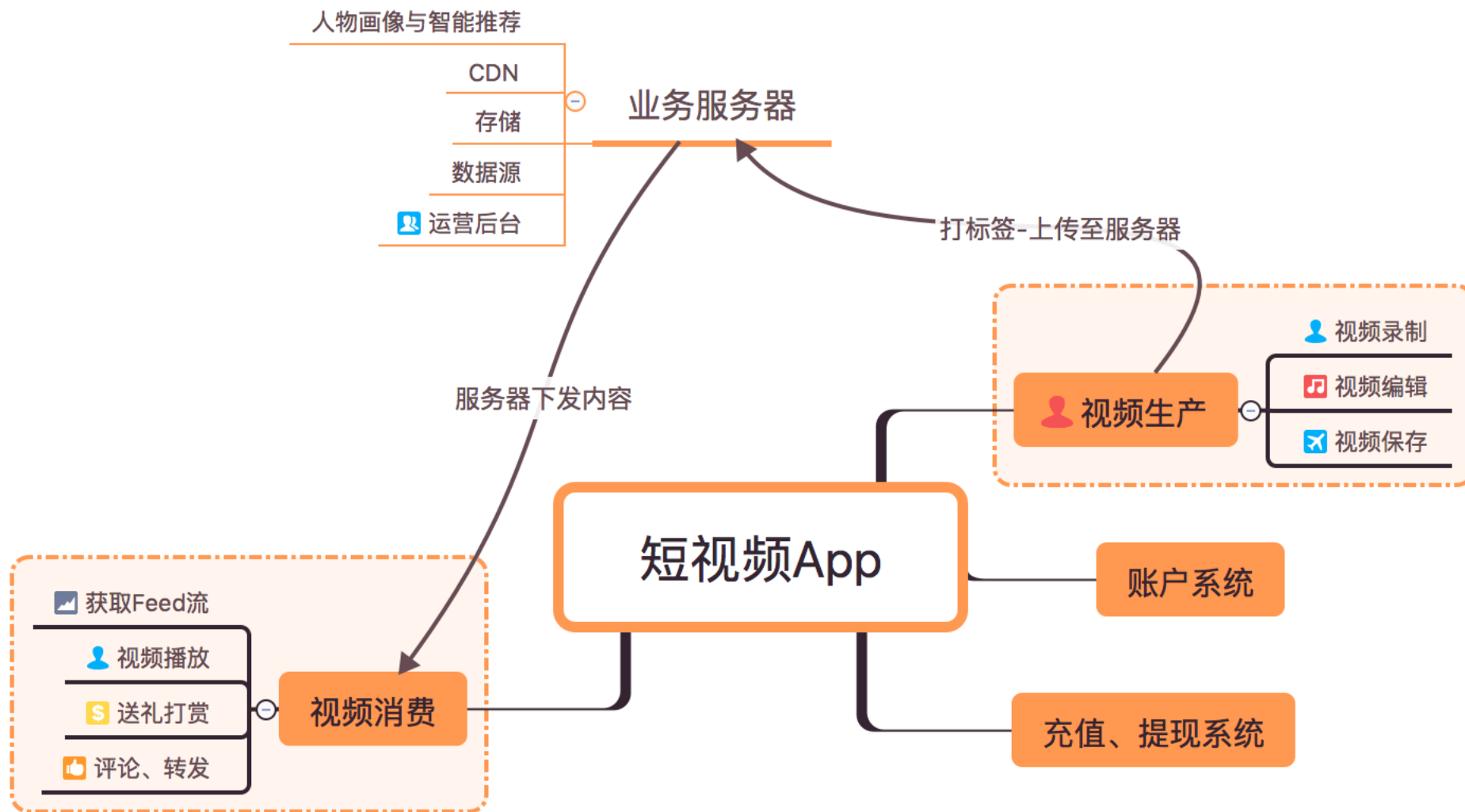
CSDN

The background is a deep blue gradient with a subtle pattern of white dots. Overlaid on the left side are several concentric circles and arcs in a lighter blue color. Some of these arcs have degree markings, such as 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, and 260. There are also some curved arrows and dashed lines, giving it a technical or architectural feel.

短视频APP的架构设计

展晓凯

短视频App的场景



短视频生产

视频录制模块

视频预览

伴奏播放

视频录制/暂停

倍速/半速录制

视频编辑模块

视频播放

视频特效

音频特效

快放/慢放处理

视频保存模块

特效时间Model

特效处理

编解码器

封装格式IO

视频录制模块

视频预览

Android/iOS平台可以使用Camera + OpenGL ES进行视频的预览
拿到纹理ID之后可以进行美颜、贴纸等特效处理

伴奏播放

将本地m4a或者mp3文件进行解码(ffmpeg或者其他解码方案)
伴奏播放的同时可以将解码的原始伴奏写入到磁盘中方便后续使用

视频录制/暂停

开始录制之后将预览的视频帧进行编码并写到文件中
如果用户点击暂停需要停止编码，当用户继续的话就将视频帧在进行编码

倍速/半速录制

需要将伴奏进行慢放或者快放，按照这个节奏进行录制视频，最终再把视频按照正常速度进行播放。

需要将伴奏的PCM进行变速不变调的处理

变速不变调处理Tempo之W-SOLA实现

基本概念(假设处理样本为单声道, 44.1k采样率)

Segment: 处理音频数据的窗口大小, 比如40ms, 换算bufferSize为1764

Overlap: 音频的重叠区域, 比如16ms, 换算bufferSize为712

注意: 保证2倍的overlap要小于segment; 为了效率保证bufferSize是8的倍数

Search: 寻找相似度最大的查找区域, 比如10ms, 换算bufferSize为441

算法实现

把送进来的数据首先保存到一个全局缓冲区里面, 定义为input_buffer
每次以segment进行一个数据单位进行处理

先在Search区域中搜索出和overlap最佳匹配的位置

所谓最佳匹配是指波形的相似程度(每个对应采样点相减的平方之和)

从最佳匹配位置开始和overlap中的数据进行Mix放入到全局的output_buffer中

Mix规则按照overlap进行fade_out, 原始数据进行fade_in的方式进行混合

从input_buffer中拷贝segment-2*overlap个采样到output_buffer中

将segment的最后一段overlap的长度复制到全局overlap中

最关键的一步: 根据factor计算出skip的采样点个数, 在input_buffer中skip

视频编辑模块

视频播放

首先需要有一个标准的视频播放器
播放器可以把

视频特效

增加贴纸

视频特技(灵魂出窍\井格\点击\九宫格等特效)

音频特效

将录制的视频静音或者加上声音特效
加上背景音乐

快放/慢放处理

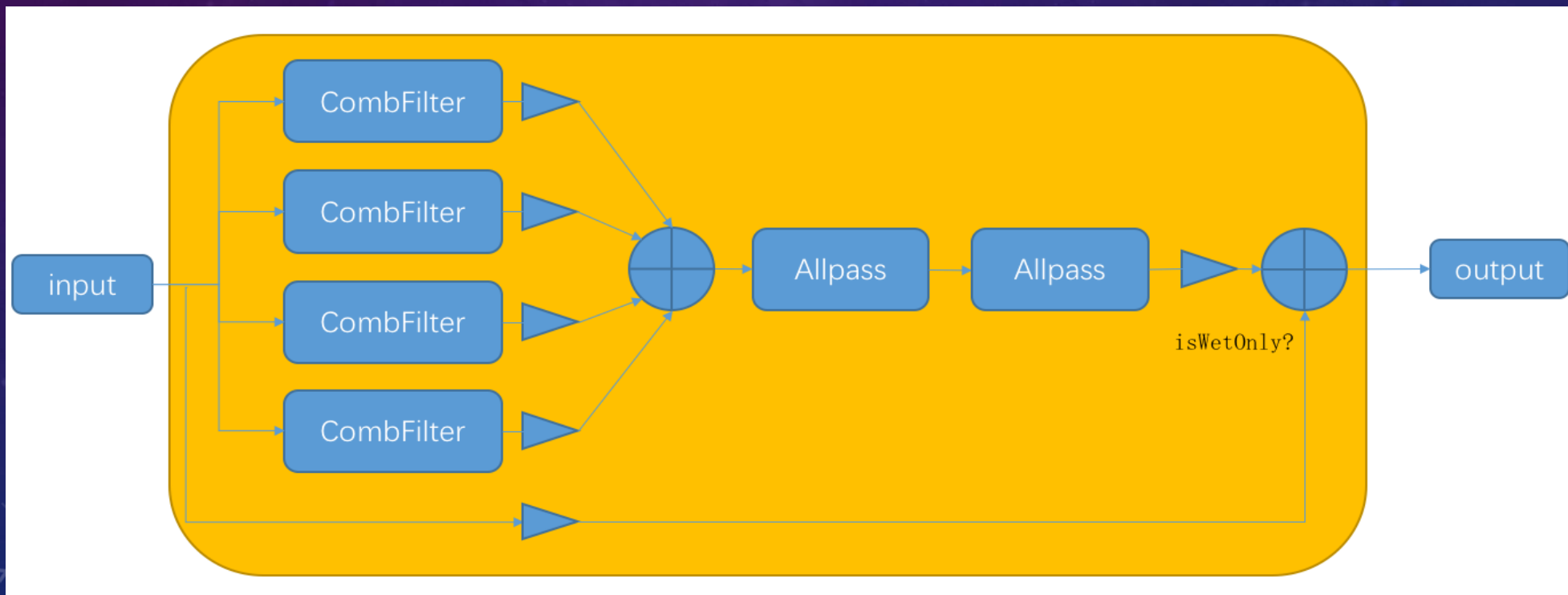
如果用户选择了快放和慢放记录住时间点,然后将视频抽帧或者复制视频帧进行重新生成一个完整15s的视频

音频效果器之混响器的实现

施罗德（Schroeder）混响模型来实现，施罗德（Schroeder）混响模型使用4个并联的梳状滤波器和2个串联的全通滤波器来建立混响模型。

梳状滤波器提供了混响效果中延迟较长的回声；

而延时较短的全通滤波器则起到了增加反射声波密度的作用；



视频保存模块

特效时间Model

视频编辑阶段会把视频特效作用的起始时间和终止时间记录到Model中，在视频保存阶段会在对应的时间段进行特效作用。

特效处理

无论音频特效还是视频特效，按照特效时间Model进行作用对应的特效处理，处理过后的音频PCM或者视频帧送到编码器

编解码器

解码器解码成原始数据(texID与PCM)交给Processor进行处理，最终交给编码器编码为H264与AAC数据

封装格式IO

最终将H264与AAC封装成为mp4，存储到本地文件中

优秀的判断力来源于经验；
但经验来源于错误的判断。

谢谢



技术开启新“视”界

Technology Bring New Vision

2018.10.19-20 北京丽亭华苑酒店

LiveVideoStackCon²⁰¹⁸ 讲师热身分享会

9月13日19:30 准时开始

《微博短视频高并发架构》

● **刘志勇** 新浪微博 视频平台资深架构师

12



关注LiveVideoStack公众号
回复『热身』获得进群方式

LiveVideoStack
— 音视频技术社区 —

CSDN