



北京
2019

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

开源协同下的视频云

— 以开放合作的心态打磨视频产品

赵军@Tencent
2019.8.24



LiveVideoStack
— 音视频技术社区 —

- 在腾讯云负责视频云系统优化
- 老码农，每天坚持读**代码BUG**写**代码BUG**
- FFmpeg maintainer，入选FFmpeg决策委员会
- 开源爱好者，参与过大量开源项目，包括Linux Kernel等
- 坚信简单胜过复杂



主题概述

明眸 音视频编解码/画质增强方案

智眸 智能媒体检索/分析/审核方案

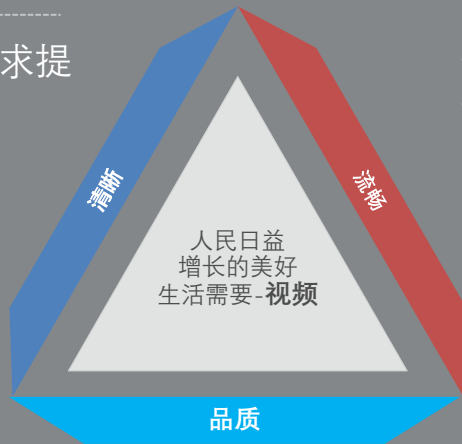
云剪 在线媒体内容生产方案

开源 取自开源，回馈开源

- 编码优化
- 画质增强
- 智能检索
- 智能分析
- 在线编辑
- 无缝发布
- 外部开源
- 内部协同

超高清清晰度

- 体育赛事、游戏等直播清晰度要求提升，带动视频行业要求提升
- 国家政策鼓励 4K、8K 爆发增长



网络带宽约束

- 4G 已经普及，5G 开始试点
- 网速仍是制约视频发展的最大因素之一

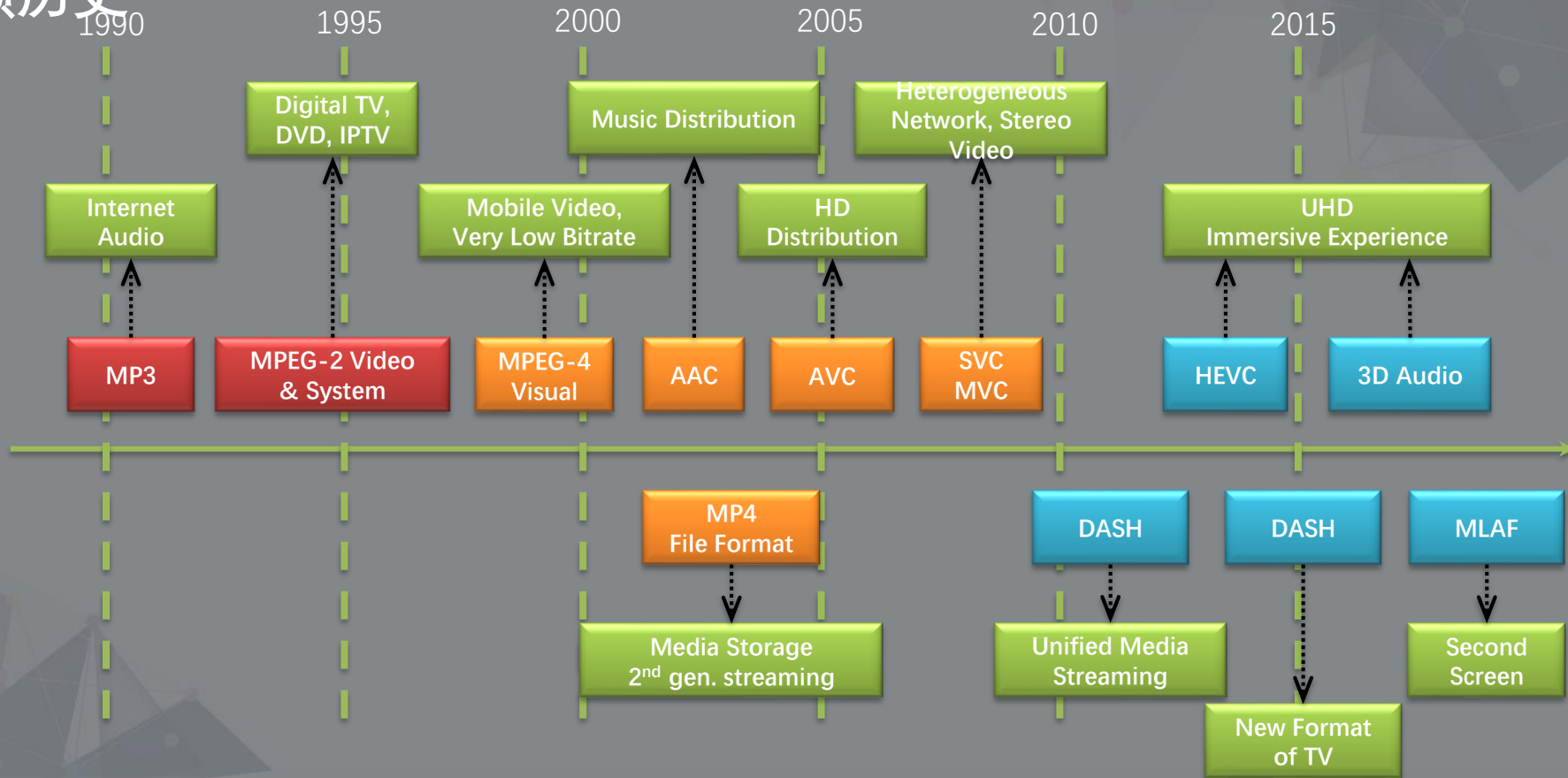
更高追求

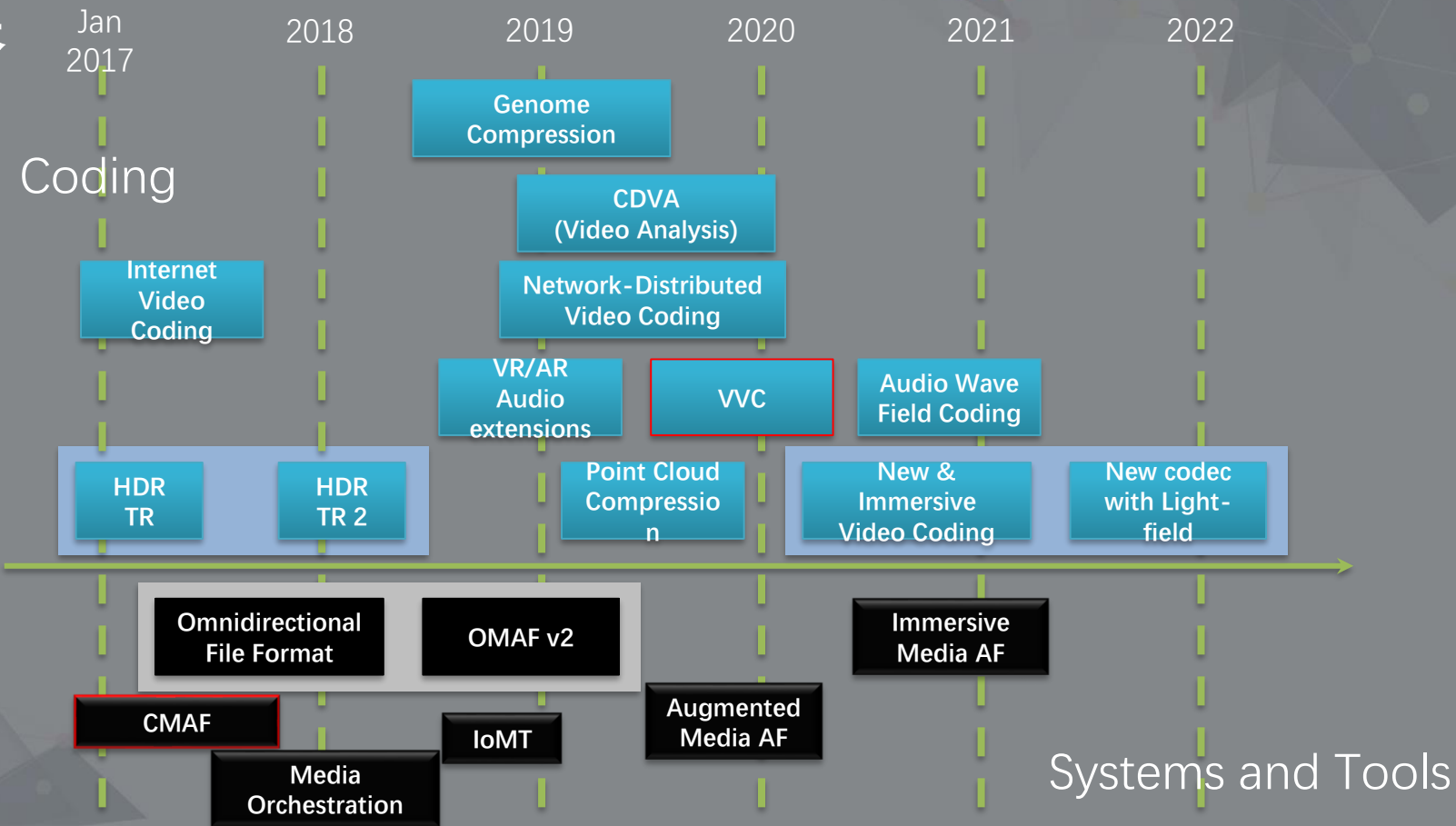
- 设备能力越来越强，需要配合终端实现更好的效果
- 人们已不满足纯摄像头拍摄的效果，追求更多的趣味性、附加能力

明眸 —— 极速高清

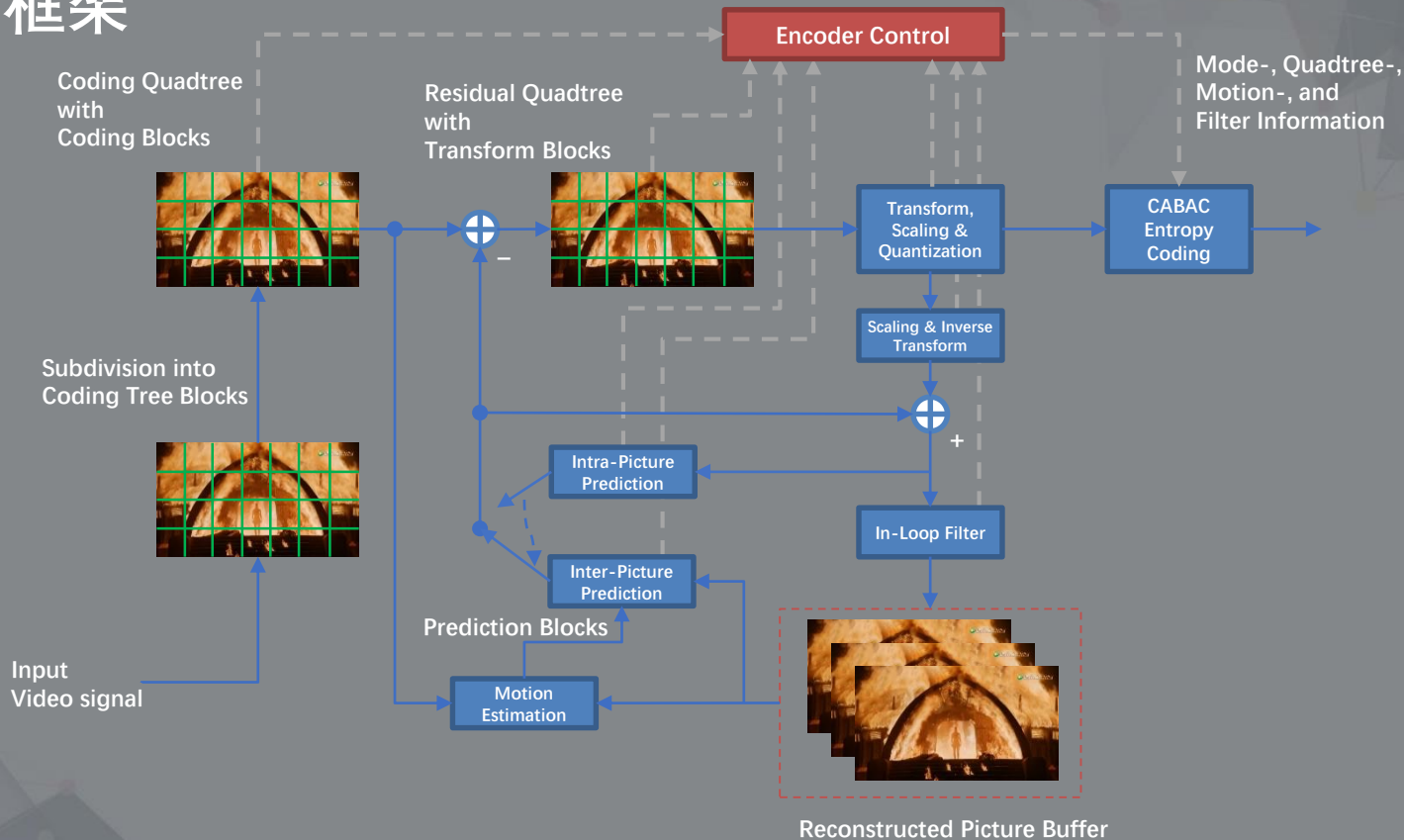
“如果人能从历史中学习，我们将能学到多少东西啊！但是
激忿和党派蒙住了我们的双眼，而经验的光亮就像船的艉灯，
只能在我们背后的波涛上留下一点余辉。

-- *Samuel Taylor Coleridge (Recollections) - Wikipedia*”

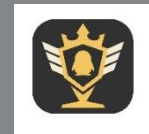
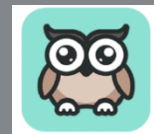
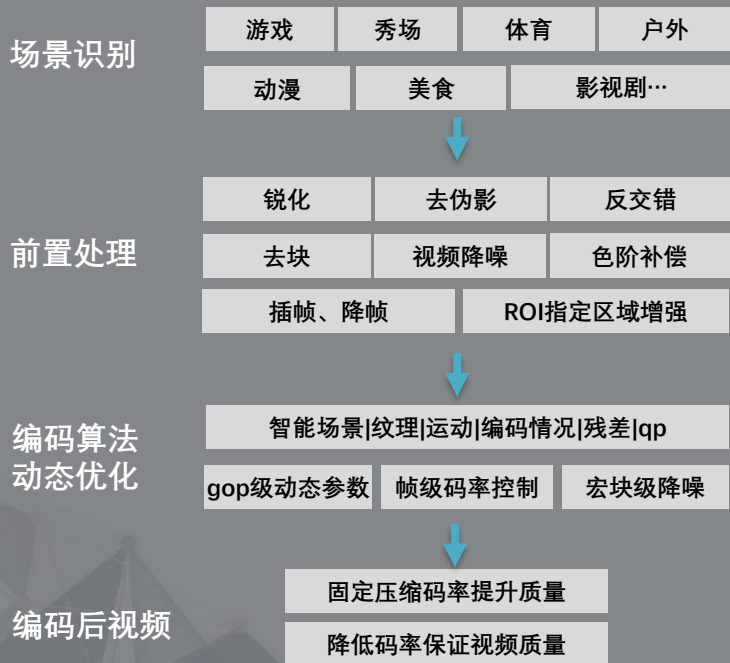




混合编码框架



在不降低视频质量情况下降低视频码率，把编码bits放在重要的地方，提高视频主观画质。



图像质量评估

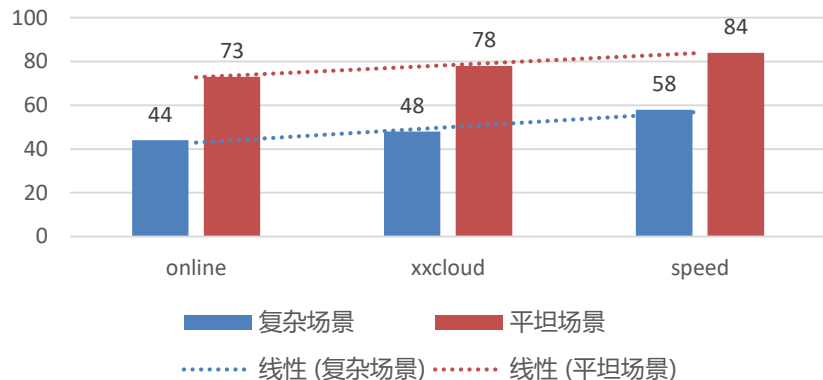
PSNR和SSIM的问题

1. 它只能衡量其与原图的差异也即相对质量，不能计算绝对质量。
2. 孤立地计算每一帧，无法利用视频中相邻帧的相关性和运动信息。
3. PSNR/SSIM值的计算法过于straight-forward，结果跟人眼的主观感受并不总能匹配。

VMAF优点

1. 算法主要关注三个指标：VIF/DLM/TI以及SVM预测
2. 根据人眼感官打分，训练模型，更符合真实人眼感受

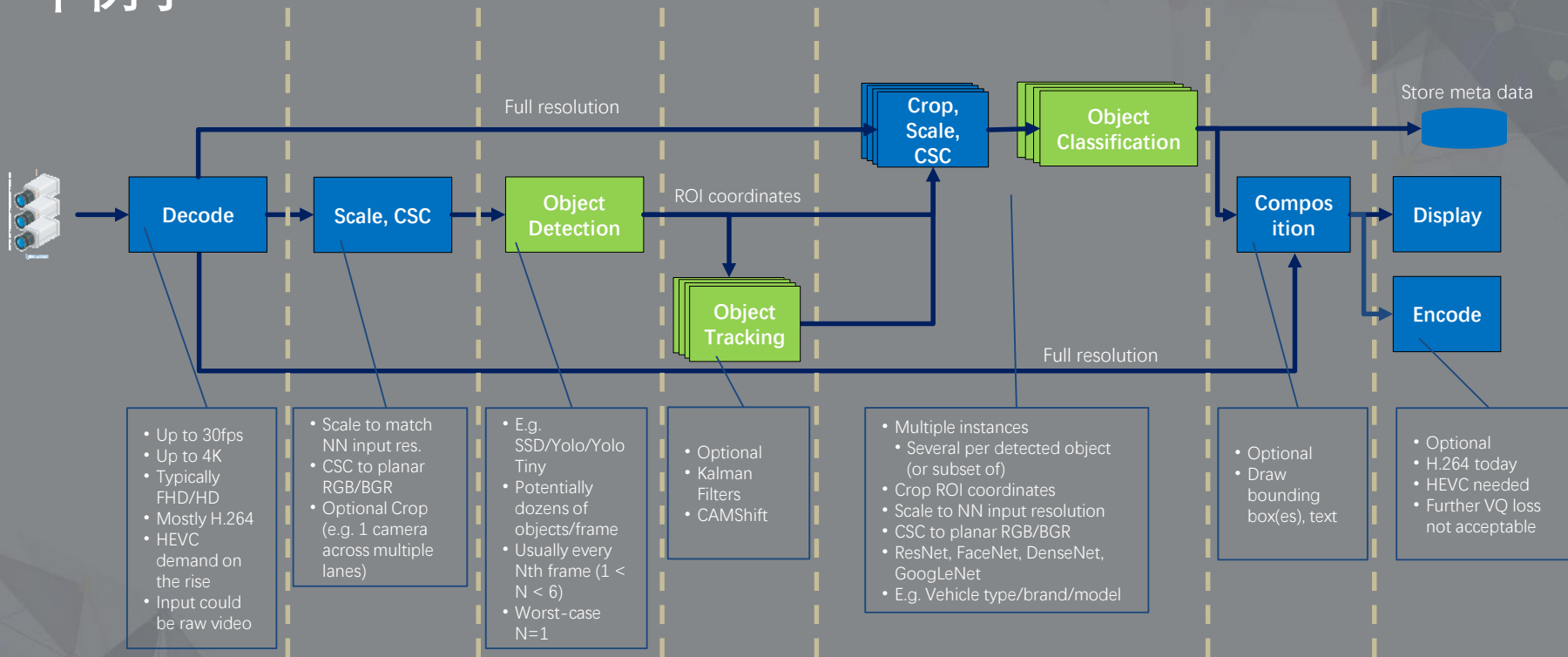
绝地求生8M源转2M码率vmaf打分测评



结果

VMAF评分提升10+, 同VMAF分码率节省30%

智眸 —— AI + 视频



视频AI

人像

声音/文字

图像

拆条/集锦

封面

分类

标签

涉黄

涉暴

涉政

智能识别

智能分析

智能审核



腾讯云

根据需求，灵活组合

泛娱乐直播场景

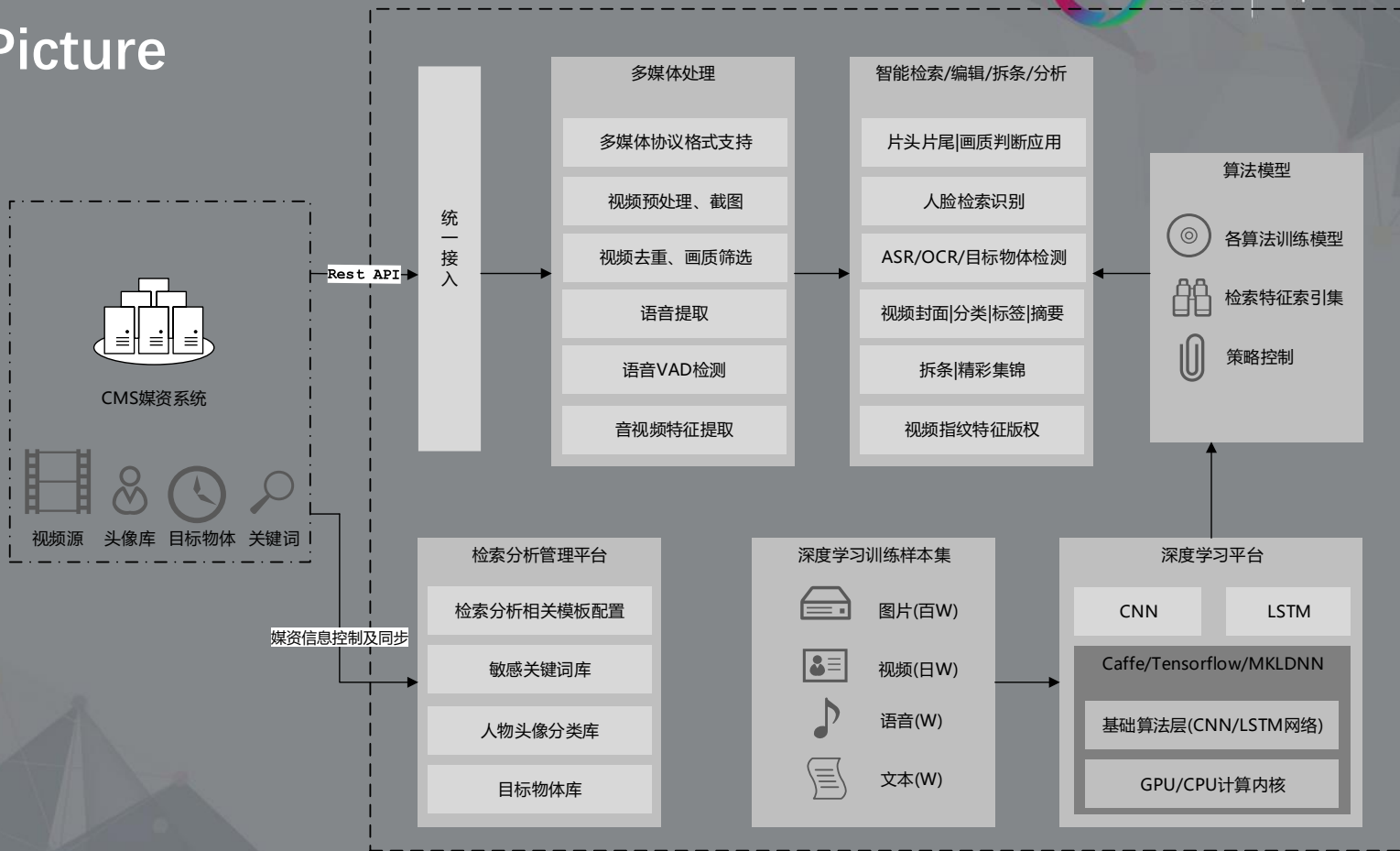
广电媒资管理场景

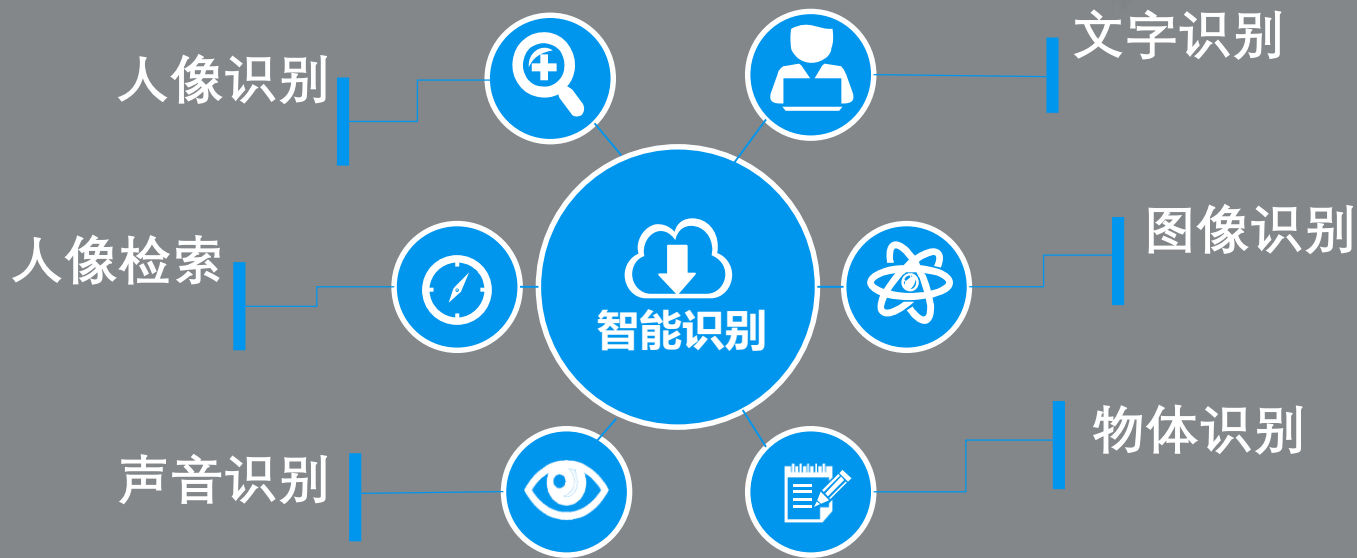
内容编辑场景

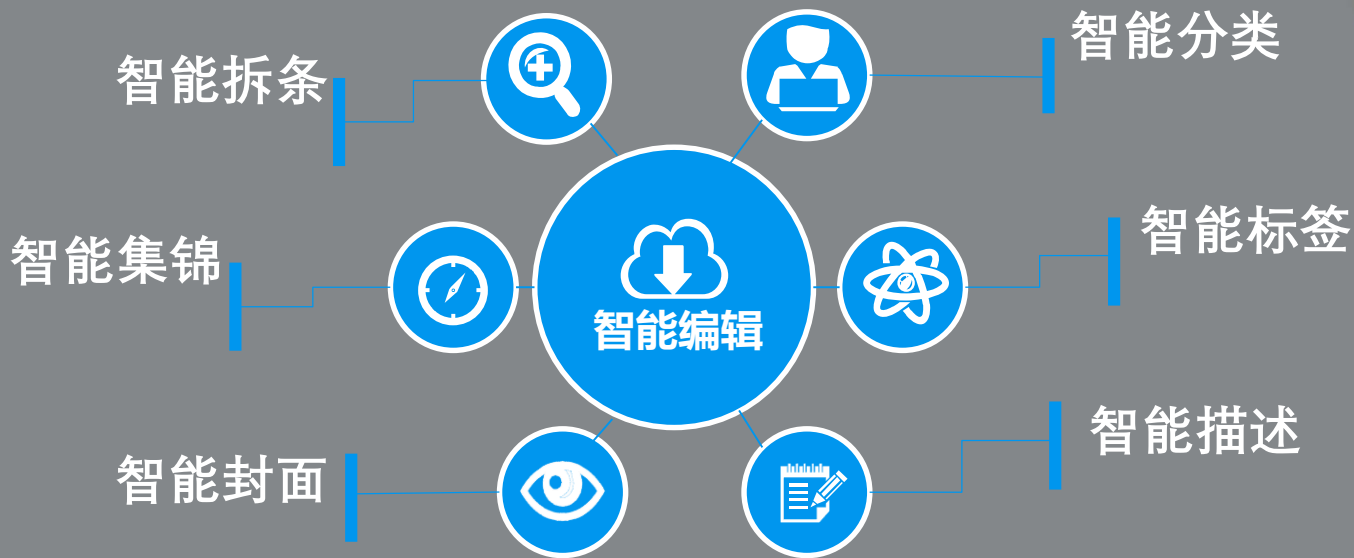
内容审核场景

内容监控场景

Big Picture







云剪 —— 视频生产

上传/汇聚

- 多终端采集
- 高速上传

制作与处理

- 编辑能力
- 模版管理
- 素材管理
- AI智能编辑
- 多平台发布

内容管理

- 媒资管理
- AI标签
- 视频编目
- 结构化分析

传输与分发

- 多协议适配
- 跨网络传输
- 多运营商
- 版权保护

终端播放

- 多屏适配
- 场景数据
- 广告变现
- 视频播放

1. 轻量、模版化、智能、素材库
2. 内容生态：泛娱乐、电竞等行业的PGC对于在线编辑刚需
3. 打通内容媒体全链条，提供完整的端到端方案

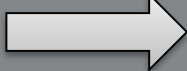
内容创作者/机构

云端剪辑工具

内容分发平台

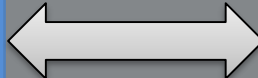
PGC/UPGC
MCN
直播平台
电竞内容

内容创作



**采
编
播
存
发**
本地上传/网络截取
图文声特效
合成/预览/超清
云端存储
多平台一站分发

高效内容输出



分发平台接入

付费

使用视频云基础组件



媒体引擎

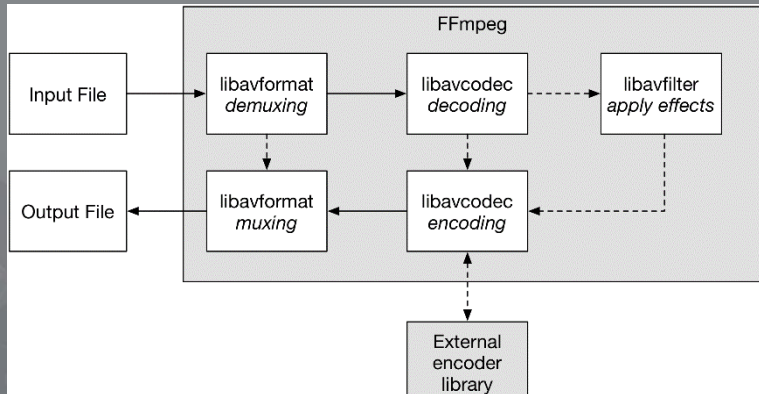
存储/转码/视频处理/AI技术

视频云PaaS能力

开源 —— 开源协同

FFmpeg简介

- **multiplatform** software project (Linux, Mac, Windows, Android, etc...)
- Comprises several **command line tools**: ffmpeg, ffplay, ffprobe, ...
- Comprises **C libraries** to handle multimedia at several levels
- **Free Software / FLOSS**: LGPL/GPL



开源与协同

“Why Software Is Eating The World

--By Marc Andreessen”

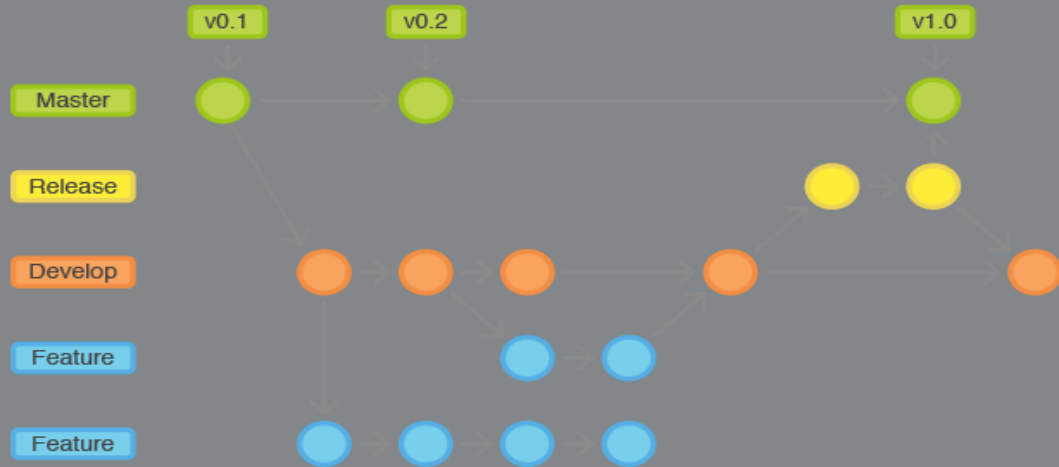
统一版本管理

- 38 个repo的故事
- 避免重复的轮子，经常看到造出来半个轮子，三角形的轮子

及时反馈社区

- Feature, Bug Fix, 性能优化, 文档更新, samples……

打造一个完整的工作流程



接口与框架

“Before I built a wall I'd ask to know What I was walling in or walling out, And whom I was like to give offense. Something there is that doesn't love a wall, That wants it down.”

—Robert Frost, “Mending Wall”

接口与框架上的考虑：

- 使用FFmpeg 屏蔽不同的Os， 不同的硬件平台， 不同的Codec细节
- 使用FFmpeg构建灵活的media pipeline
- 集成Deep Learning 框架到FFmpeg 的AVFilter 模块是个好主意吗？

“If you can’ t measure it, you can’ t improve it.

-- *Peter Drucker*”

性能之痛

“Premature optimization is the root of all evil”

--Donald Knuth”

优化是永恒的话题

- 所有优化的前提是理解**算法与数据流向**
- 用Profiling的数据说话

理解硬件（多核，多线程，SIMD，Cache等等；GPU又是怎么回事，FPGA， SoC呢）

CPU 优化的两大利器

- pthread： 解码（Frame/Slice）， Filter(Slice)， 编码器的性能优化(x264)
- SIMD 到底是什么？

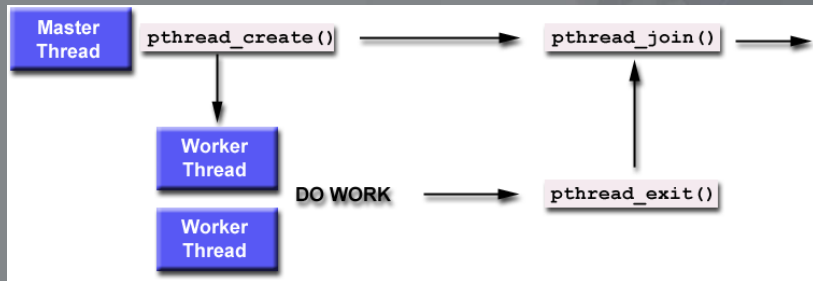
GPU 优化的方式

- 尝试与CUDA对抗的OpenCL
- OpenGL也可以用来做加速
- Vulkan 会一统江湖吗

CPU加速

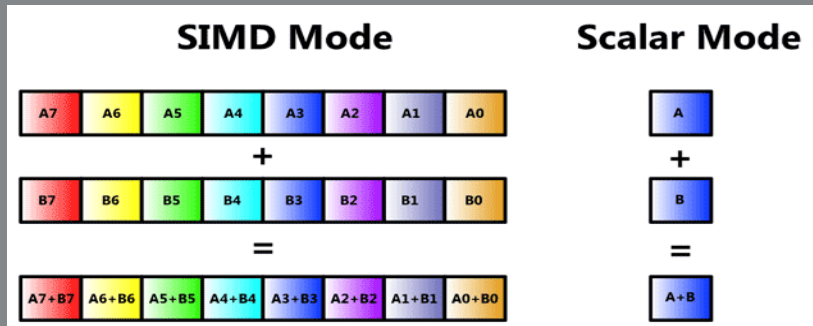
Thread 加速

- 充分释放多核的能力
- AVFilter 的 `AVFILTER_FLAG_SLICE_THREADS` 意味着什么?
- FFmpeg 使用了工作组模式, 对不相关的数据以行(row)为单位划分处理
- 内部某个项目, 线程优化之后, 从7ms 到了2ms



SIMD 加速

- 三种汇编优化形式:
 - intrinsics
 - inline assembly
 - hand-written assembly
- 以nasm/yasm 汇编语法为主 (why?)
- 复用了x264 的汇编优化库 x86inc.asm
- Nlmeans 的 优化是个典型的例子

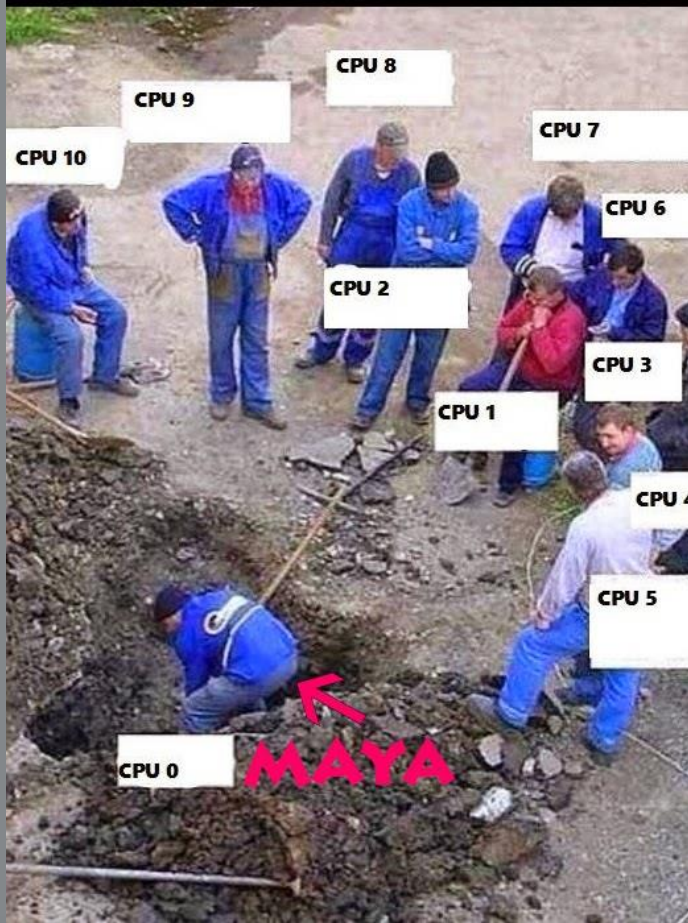


多讲一点线程的问题

我们的窘境：

一核有难，八核围观

It's okay,
I didn't need speed anyway



多讲一点线程的问题 续

考试时间

使用FFmpeg的时候，如何设置我们的线程？

真实的故事

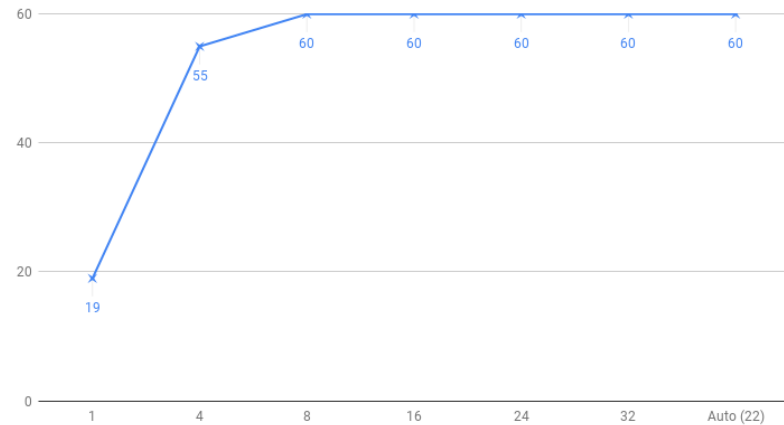
故事一：增加一行代码，一条复杂的Pipeline的执行时间减半

故事二：嘿，哥们，我们拥有1200 个线程在一个进程里面

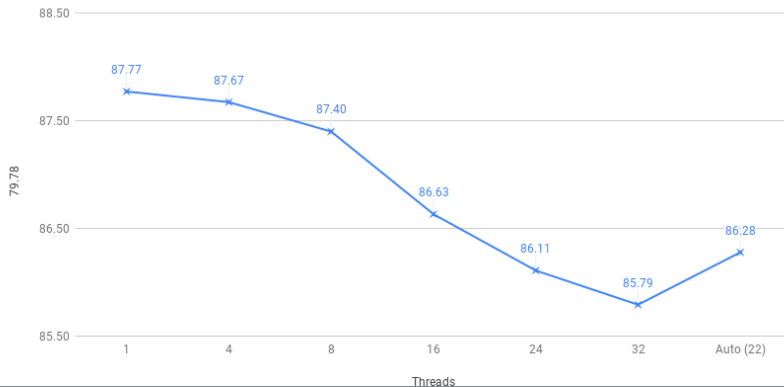
故事三：QA告诉我AV1 encoder crash了，但我始终不能重现这个问题

故事四：什么？线程不只是影响转码性能？

Threads and 720p Capture FPS

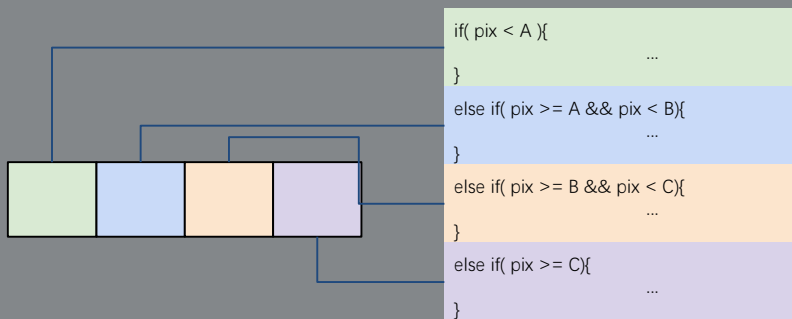
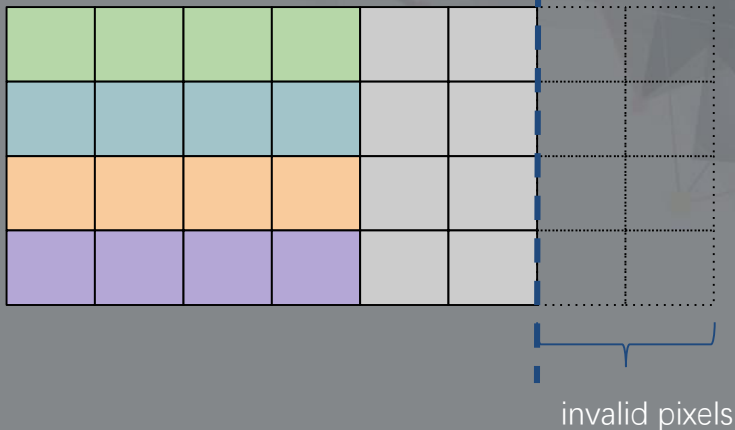


VMAF by Thread Setting



SIMD 优化面临的一些问题

- 发现数据并行算法有时候并非易事
- 不同硬件之间的移植性
 - sse-to-neon
 - neon-to-sse
- 边界处理(SIMD 一般都有内存对齐的需求)
 - Padding, Predication, Fallback
- 分支的处理
 - Predication, Fallback to scalar
- 不是所有Ops都被SIMD指令支持
 - Division, High-level function (eg: math functions)
- Floating-Point
 - Unsupported/cross-device Compatibility



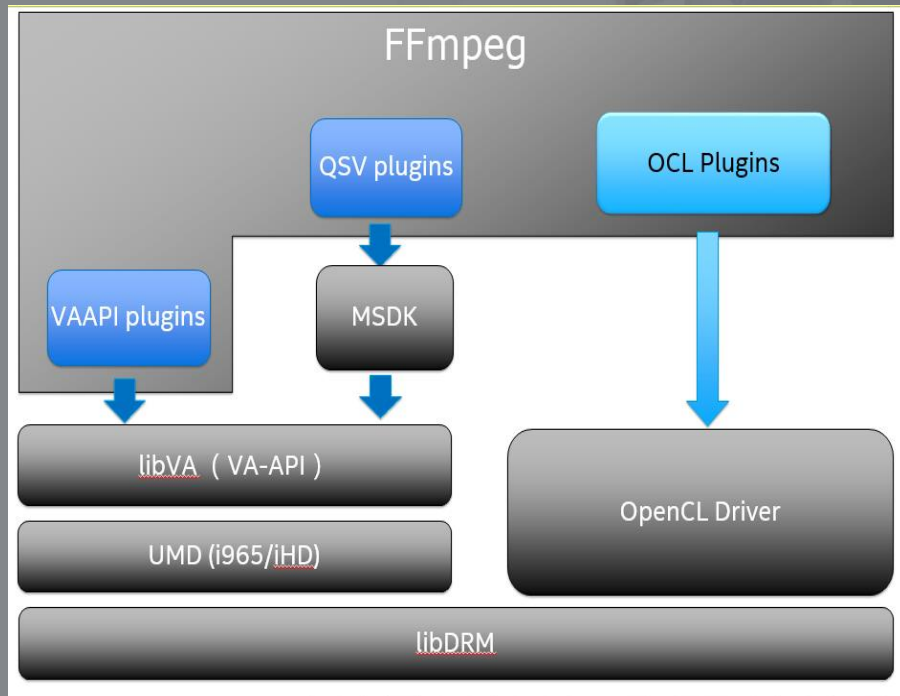
GPU加速

FFmpeg HWAccel with Intel GPU:

- FFmpeg QSV plugins are based on MediaSDK; it has widely accepted by customers.
- VA-API is lower level API; FFmpeg VA-API plugins provides more flexible solution for customers.
- Integrate 3rd-party OCL/OpenCV/Vulkan video processing library to enrich the solution.
- FFmpeg Plugins with HW acceleration will speed up development for different usage

一个有趣的问题:

什么, 我不知道我是否真的使用了硬件加速功能?



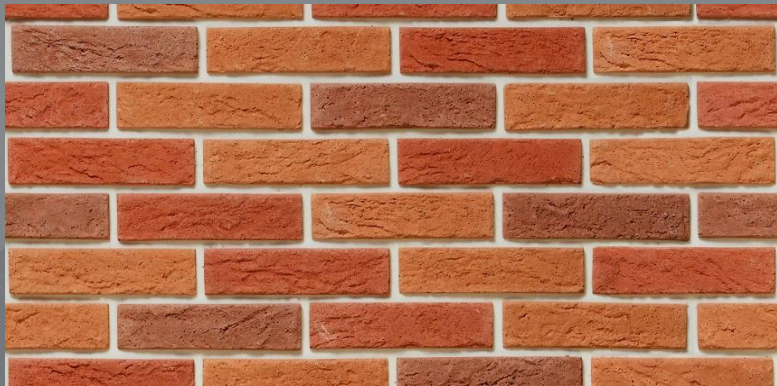
AI 是FFmpeg的一部分吗？

“To be, or not to be, that is the question

-- from Hamlet, spoken by Hamlet”

- 实现一个特定 Inference filters? 还是集成一个重度的inference框架?
- 扩展已有的简单版本的DNN_Inference API?
 - FFmpeg 已经集成了部分DL的功能，主要是Super Resolution filter，Derain，但性能堪忧
存在一个非常简单的DNN API

期望的架构

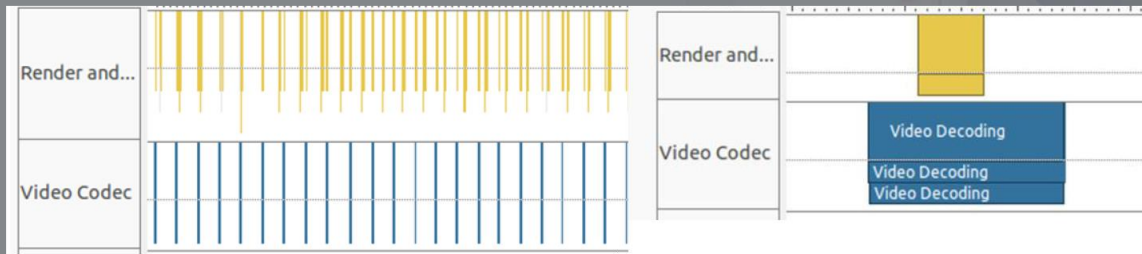
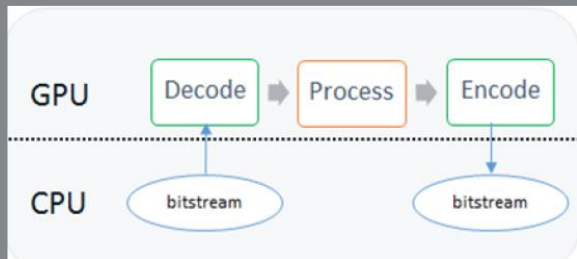


实际的产品



开放性问题

● CPU 与 GPU的调度问题



● CPU 与 GPU的数据交换问题

➤ 数据从 CPU到GPU与数据从GPU到CPU并不对等

- ✓ Mmap
- ✓ SSE4/AVX/...
- ✓ GPU Copy
- ✓ OpenCL SVM

➤ Buffer Sharing

https://www.khronos.org/registry/OpenCL/extensions/intel/cl_intel_va_api_media_sharing.txt

<https://www.freedesktop.org/wiki/Software/Beignet/howto/libva-buffer-sharing-howto/>

Q&A

Backup

动态参数

ref	deblock
chroma	Aq-strength

• • • • •

码率控制

sigmoid	texture ration
motion ration	visual sense

● ● ● ● ● ● ● ● ● ●

● ●

自适应降噪

frame type	block type
qp detail	templ strength

• • • • •

编码参数模板

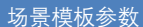
视频场景实时识别，直播场景识别
延时1-3分钟，根据场景识别结果，
不同场景配置不同最优编码参数
模板

动态优化

结合纹理和运动检测因子，在模板参数的基础上，进行参数细化

结果

vmaf 评分提升10+, 同vmaf评分
码率节省30%





高精度

MegaFace竞赛100万规模人脸检索首选识别率83.29%世界第一，同时通过异源数据训练提升各类不同场景下的照片的识别精度。



人脸质量评估

针对非配合场景下，自动判断光线、姿态、角度等，选择最优人脸图片，提高人脸检索精度。



人脸变化识别

腾讯优图自研人脸演变模型，即使随着年龄增长，皮肤、肌肉、骨骼发生明显变化，优图天眼也能正常识别。



多角度识别

腾讯优图自研去遮挡技术，有效减少眼镜、口罩、帽子等遮挡物对人脸识别算法的影响。

语音文字识别



微信智聆

- **海量数据积累**：立足腾讯社交数据大平台，积累了数十万小时标注语音数据，识别率高。
- **算法业界领先**：应用业内最领先语音识别算法，包括LSTM，BLSTM，Deep CNN，LFMMI，CTC 等算法；结合超大规模语言模型，覆盖广泛的垂直领域。
- **支持深度定制**：提供深度定制服务，根据业务需求深度定制具有行业特点的模型，让行业专有名词识别得更精准，更高效。
- **噪声环境识别**：模型鲁棒性佳，能够识别来自嘈杂环境的音频信息，不需要客户进行降噪处理。



腾讯优图

- **准确性高**：腾讯云文字识别 OCR 可自动从图片中定位并识别字段，印刷体的平均准确率可达 90% 以上，手写体的识别平均准确率高达 85% 以上，鲁棒性强。
- **稳定性强**：腾讯云的身份证识别、名片识别、营业执照识别服务已成功应用于微众银行、QQ、广点通等腾讯内部核心业务，接受过海量用户和复杂场景的考验，各项反馈良好。
- **适用性高**：优图自研的 OCR 技术，涵盖了整个证件检测识别框架的所有核心算法，支持横向、竖向拍摄，适应透视畸变、光照不均、部分遮挡的情况，具备非常高的复杂环境可用性。

致谢

- 腾讯云

vultureli; jorryliu; tasyticfan; definesun; chenxinliu; tomajsjiang;
vacingfang; tstan

- 腾讯视频

geminili; peterzhuang; macweng

- FFmpeg社区

悟空; Ruiling Song; Michael Niedermayer; Zhongli



北京
2019

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

Thank you



LiveVideoStack
— 音视频技术社区 —