

2021-11-11



LVS 2021 大会相关内容分析

wangwei1237@gmail.com

LVS 2021 整体情况介绍



1. 视频编解码技术



1. 视频编解码技术——编解码标准和应用

各大厂商均开始探索新的视频编解码标准，以提升更复杂视频应用场景下的用户体验

* H266: 腾讯, 阿里, 字节跳动

* 阿里有一个主题演讲专门介绍在 H266 方面的工作。

* 火山引擎的展台展示了H266 编码 8K 视频的效果。

* 国内公司对 H266 的贡献: 华为, 字节跳动, 腾讯, 阿里, 快手, 海康威视, 富士康, 中兴, 清华大学, 鸿颖创新, OPPO

* AV1: 作为一款开源的编解码标准, 从这次 LVS 2021 看, 业界也开始有公司在探索

* 拍乐云基于 AV1 的实时视频系统技术实践

* 欢聚集团 AV1 技术实战

* 部分Windows 机器和 GPU 硬件设备已经宣布开始支持硬件解码能力

目前应用较为广泛的 H264, H265, VP9等讨论的相对较少。同时, 视频应用场景的复杂化对于更高压缩率/更高质量的编解码标准有更迫切的需求。

* 更高的画质要求: 4K/8K, HDR

* 更低的延迟要求

1. 视频编解码技术——编解码场景和架构

通用编解码标准之上的额外的编解码能力加持

* AI 应用能力

- * AI 视频插帧在互动场景中的应用
- * AI 在视频处理、视频内容理解等领域的应用
- * AI 辅助编解码的 ASIC 解决方案

* 基于兴趣区域编码的 ROI 编码

- * 字节跳动

* 基于内容自适应的 CAE 编码

* 增强编解码技术

编码架构从传统的编码架构 (H26x) 开始向基于深度学习的编码架构

- * 阿里
- * 腾讯
- * 字节跳动
- * 快手

1. 视频编解码技术——硬件编解码

硬件编解码越来越重要，视频编解码从 CPU -> GPU -> VPU (FPGA, ASIC)

* GPU 和 VPU 的区别

- * GPU 本质上更多是用于图形渲染，只是因为有可编程的管线阶段所以为并行计算提供了机会
- * GPU 的价格相对较为昂贵，功耗相对较大
- * VPU 为视频编解码专门定制，在算力上具备天然的优势，功耗更低，集成度更大
- * 谷歌自研的 Argos VCU 芯片，对于 YouTube 而言，单纯的硬件成本的节省就可达到亿\$级别。

* 这次 LVS 中，介绍自家 VPU 芯片产品的厂家也比较多，并且这些公司大多有类似快手这样的公司参股

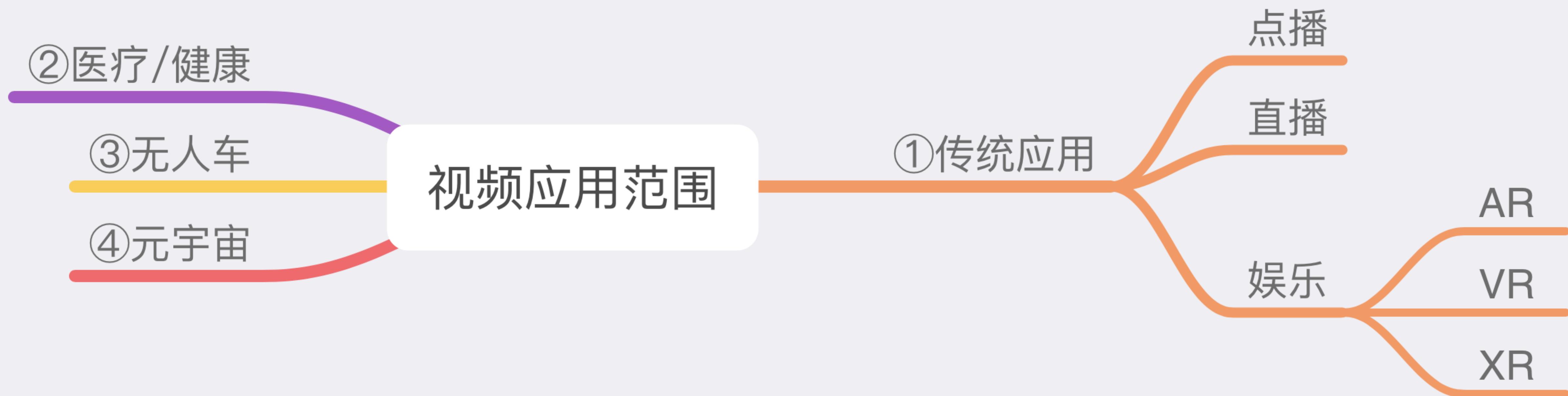
- * NETINT，刚完成数亿元的A+轮融资
- * 涌现
- * 瀚博
- * IBM

硬件编解码的加速无疑能够加速视频的处理流程，进而在节省成本的基础上提供极致体验

- * 涌现的直播云端转码芯片，可以实现实时的一转四编码
- * IBM 的 AI 芯片可以在硬件上实现视频实体识别，视频切割等能力
- * 字节跳动，快手也均在业务上探索异构框架，提升编解码速度，进而提升用户体验
- * 硬件编解码对于创作者而言最直观的体验就是进一步降低提交作品到作品发布之间的延时

虽然有卖产品的嫌疑，但是可以看出趋势；未来，异构计算视频业务的必然趋势。

2. 视频的应用范围



视频应用场景从传统的娱乐场景开始扩展，例如医疗/健康，无人车，元宇宙等。

这次 LVS 新增了一个无人车专场：《软硬结合，加速高级别自动驾驶落地》

比较值得一提的是这次 LVS 的各个专题中，大家也都在不同程度的谈论元宇宙。

3. 视频业务实践



3. 视频业务实践——架构升级推动生产力提升

视频架构的不断演化、更新是视频业务生产力提升的主要手段，没有之一

* 快手的黄琦讲到：视频作为一种资源消耗性的重工业产业，降本增效是核心，效率是核心的竞争力。而快手的视频基础架构从工具集->服务->函数计算引擎的升级是效率提升的主要手段。

* 同样是视频业务，从目前的情况看 Facebook(Meta), Netflix, 快手, 微博等在架构演进上基本思路都是一致的，基本都演进到了函数计算引擎的阶段，并且从大会介绍的效果看，对于效率提升的收益是非常明显的。

* 这里涉及到的研究有：快手黄琦的主题演讲《视频大时代下基础架构的演进趋势》，微博黄阳全的《微博视频处理系统的云原生之路》，字节跳动浩铭的《从体验出发，构建以增长为目标的视频服务体系》

不同的公司，面对相同的问题空间时，即时没有过交流，但是大家最终的解决方案出奇的一致。没有必要非要看别人是怎么做的，搞清楚自己问题的真正所在，比搞清楚别人是怎么做的更重要。

虽然巴别塔没有建成，但是在地球的所有区域，人的四大基本活动：衣食住行也是惊奇的相似。

3. 视频业务实践——场景细分，降本增效

场景细分，为不同的场景找到最适合的优化方案

- * 快手对视频进行生命周期的细分，根据视频的生命周期来平衡带宽、计算、存储三种资源，实现成本的优化。
- * 同样的，字节跳动也会针对不同的场景进行对应的优化，例如朱利安的《抖音背后的RTC是如何追求极致》的分享中针对不同的RTC场景进行对应的优化。例如不同的硬件采用不同的推流方式等。
- * 字节跳动还会针对不同的机型进行对应的优化，进而提升QoE，例如从视频的预加载，预解码，预渲染，硬件异步初始化等。例如，对于适合的场景，直接采用预解码解析出视频第一帧作为封面，从而避免了封面图的加载，对于抖音这种大体量的APP而言，其中节省的成本是不容小视的。
- * 字节跳动还构建了一个由2W+的手机设备构成的设备库，用以支持不同端采用不同的策略，以充分利用不同硬件设备的能力，提神该用户体验。

产品的基本满足和充分好用之间差异非常大。对于不同的场景，不同的设备，要找到该场景下的最优的用户体验。细分场景的优化已经成为大家共同讨论的一个问题。

3. 视频业务实践——基于线上AB实验的QoE提升

本次 LVS 2021 有一个视频 QoE 的专题分享，但是整体看，目前大家的 QoE 提升思路，视频质量评估思路和方法和我们目前的方式是基本一致的，没有特别领先的概念出来。

但是，字节跳动浩铭的《从体验出发——构建以增长为目标的视频服务体系》提出了他们基于AB实验进行QoE提升的思路，以及该思路在抖音的实践。基于AB实验而不是QoS指标来进行优化，并且建立了一套完善的、强大的、面向元无知的埋点体系用于支持AB实验。

- * H265硬解覆盖度80%
- * 平均首帧耗时100ms
- * 成本节省 50% (编码优化+浪费节省+单价优化)

3. 视频业务实践——边缘计算的不断应用

边缘计算越来越重要，这次 LVS 2021，边缘计算是唯一一个一整天都有专题讨论的话题。可见，随着5G、硬件的发展，边缘计算变得越来越重要。

中国移动的郭勐分享了《视频边缘智能服务》也透露出运营商在视频相关的边缘计算中也开始做布局。

视频边缘计算的普及和硬件芯片能力的增加有着密切的关系，例如摄像头采集到数据之后，可以直接利用现场网络中的计算能力（AI分析芯片）进行视频分析进而将结果回传云端，而不需要将视频数据回传到云端进行处理。

另外，私有 IoT 网络的不断扩大，5G 网络本身特性带来的基站数据量的增加，也推动着边缘网络（计算）的趋势。

3. 视频业务实践——软硬结合，提升效果

从卖方市场看当前产品趋势，视频相关硬件的发展趋势越来越好。

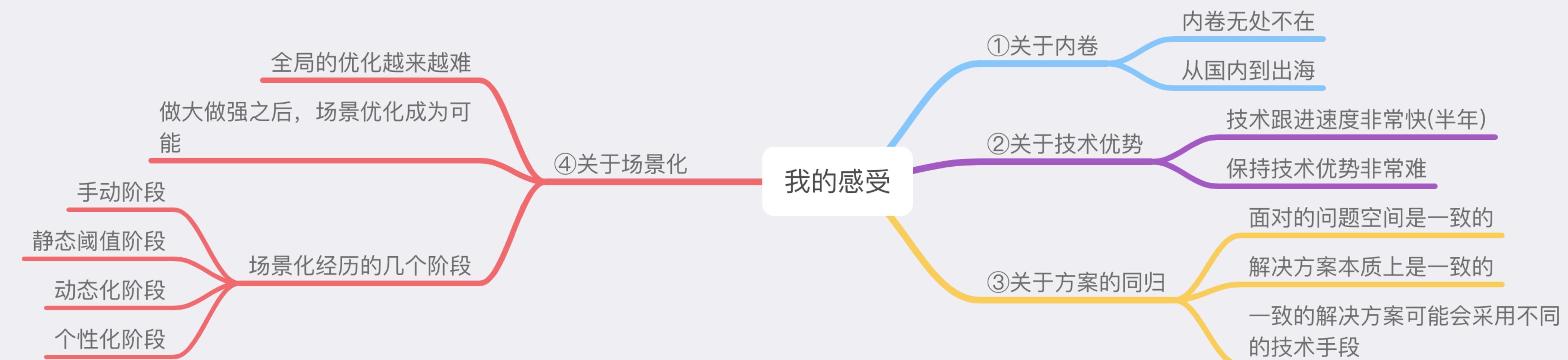
AI推理计算，视频编解码计算等能力已经慢慢从GPU中拆分出来，形成独立的硬件单元。

目前需要均衡的是VPU硬件单元中的AI能力与编解码能力的算力权衡问题，不同的产品对这两者的权衡也不相同。例如IBM的硬件更多的偏向AI能力，而NETINT的硬件更多偏向编解码能力。

快手分享了《基于GPU, FPGA, ASIC等平台对快手的多媒体业务的异构方案探索》，介绍了他们的异构方案在语音应用方面的探索。

尤其是在视频业务，异构计算必然是大势所趋。

4. 我的感受



Thanks!

技术奔腾向前，视频势不可挡

面朝大海，春暖花开

