

视频压缩的前世今生

本文仅从工程和实施角度帮助用户理解视频压缩技术，不涉及过深的算法原理实现。

1. 视频压缩概述

视频是由连续的图像帧序列组成，图像连续播放呈现出连续动态影像即为视频。所谓视频压缩即视频编码技术，是指通过压缩技术将原始的视频格式转化为另一种数字视频格式，其目的在于减少文件空间占用。比如摄像头经过采集后直接生成的原始视频，不考虑压缩的话，其视频数据量非常庞大，在网络传输和数据存储的过程中，带宽资源和存储设备无法支撑如此巨大的数据量，所以在进行传输或者存储前，需要将原始视频数据进行视频压缩处理。



2. 视频编解码器

视频编解码器是指一个能够针对数字视频文件或者视频流进行压缩和解压缩的程序或者设备。由于编解码实现路径和方式的不同，视频编解码器也分为硬件编解码器和软件编解码器。目前主流的视频编解码器分为 5 个系列：H. 26X 系列（H. 264/HEVC/VVC）、AVS 系列（AVS1/AVS2/AVS3）、SVAC 系列（SVAC1.0/SVAC2.0）、VP8/VP9 系列以及最新的 AV1。其中 AVS 系列和 SVAC 系列均是我国具有独立知识产权的视频编解码标准。评价编解码器优劣的主要评价标准在于：压缩比的高低、编解码速度的快慢、编解码功耗以及解码兼容性。



3. 视频压缩的本质

视频可以被压缩是因为前后帧图像数据中存在大量的信息冗余，主要可以从以下四类进行探讨：

空间冗余：同一图像帧内，相近像素之间的差别很小，具有较高的相似度；

时间冗余：视频中相邻的前后图像帧之间，其中发生变化的像素占整张图像像素的比例极其微小，存储运动关系，该特点在监控场景中尤为常见；

编码冗余：不同像素的出现概率不同，出现概率高的像素分配较少的字节，概率低的分配较多的字节；

视觉冗余：人眼视觉系统对不同像素颜色敏感度不同，编码时可以丢弃这些冗余的信息，而几乎不影响人眼视觉效果。

4. 视频画质评价

谈到视频压缩，就免不了要谈到压缩后的视频画质。在监控领域常见问题就是压缩后视频画质和原始视频是否有区别，能否作为证据进行溯源，是否会影响AI智能分析精度等；在互联网领域大家就更关注压缩后视频画质是否会出现马赛克等而影响用户对视频美观程度和流畅度评价。

视频编码画质评价标准主要可以分为主观评价和客观评价，其中主观评价标准是通过人眼直观观察进行评分，是最为接近的用户对视频画质的整体反馈。但由于主观评价的完成需要耗费大量人力和时间，并且受人为的影响因素较大，这里暂时部深入进行讨论。



本文将从客观评价标准简要阐述下视频编码画质的评价标准，主要介绍目前主流的全参考视频质量评价模型的两种视频评价标准:PSNR（峰值信噪比）、MS-SSIM（结构相似度）。

PSNR：目前最主流的画质评价方法，被公安部相关机构和互联网企业引用最多的评价方式；PSNR 取值 $20\sim 40$ 之间，数值越大，画质越好。

MS-SSIM：通过感知结构信息来评价画质，其结果更为接近肉眼感知效果；MS-SSIM 取值 $0\sim 1$ 之间，数值越大，画质越好。

5. 有损压缩&无损压缩

我们知道视频是由连续的静止图像构成，可以在相对很短的时间内传递大量的信息。视频压缩与图像压缩类似，在视频编码压缩的同时，除去基础的图像压缩之外，需要加入的运动特性压缩的考虑，才能达到较为满意的压缩效果。

有损压缩：解码还原出来的图像/视频/语音和原始文件存在差异，允许一定程度失真。我们在市面上看到的绝大部分音视频文件都是经过有损压缩，由于有损压缩往往压缩的部分是人类感官注意不到或者感知频率较低的部分，有损压缩被广泛的用于工业界相关程序。

无损压缩：解码还原出来的图像/视频/语音和原始文件完全一致。无损压缩的比例很低，只应用于少部分的图像领域，不太适用于语音和视频领域。主要应用场景：部分尖端医学图像（DCM）、遥感图像（TiFF）等。



6. 视频编码标准&视频封装格式

视频编码标准：在国际和国内组织的独立和联合开发中，产生了很多重要的视频编解码标准。视频编码标准的出现，完整地改变了由于各国的电视制式不一致，因此不能直接互通的混乱现象。最早地视频是 1988 年诞生地 H. 261, 此后延续的出现了 H. 263、H. 264/AVC、H. 264/HEVC、H. 266/VVC。其中 H. 264 和 H. 265 成为现阶段互联网视频和监控安防等视频领域应用最广泛地标准。此外还有 AVS 系列（AVS1/AVS2/AVS3）、SVAC 系列（SVAC1.0/SVAC2.0）、VP8/VP9 系列以及最新的 AV1 等视频编码标准。

视频文件格式：很多初入音视频领域的新手很容易搞混视频文件格式（MP4/AVI/RMVB 等）和视频编码标准的区别。我们常见的 AVI/MOV/FLV 其实是一种视频封装格式，是一种视频信息存储的方式，一般由相应的公司自行研发出来，部分视频封装格式带有各自公司的私有的播放器才能播放。我们互联网领域常见的类似 AVI 和 WMV 是就是由微软推出的，以及 RMVB 是由 Real Networks 公司推出。这些封装格式相对应用广泛，绝大部分播放器都是实现自由视频播放。监控领域巨头海康威视的 PS、PES 和大华的 DAV 的封装格式相关体系封闭，往往需要该公司自己的播放器才能实现较为完整流畅的视频播放体验。



7. 分辨率/像素&视频画质

分辨率/像素：视频分辨率是指视频宽高的像素数值，单位 Px。像素为视频宽高像素值的乘积，数值上等于等于宽 X 高。我们常见的 720P、1080P、4K 就是指分辨率 1280×720、1920×1080、3840×2160，分别对应 100 万、200 万，800 万像素值。

视频画质：视频画质是指视频压缩后图像和压缩前的原始图像的失真程度的描述，失真程度越低，画质越高。视频画质的评价受主观和客观评价的影响很大，目前常见的评价画质标准有 PSNR（峰值信噪比），MS-SSIM（层级相似度）、MOS（平均意见分）、MSE（均方误差）等。

8. 硬件编解码和软件编解码

软件编解码：通过 CPU 方式来实现视频的编码和解码过程；

硬件编解码：通过显卡 GPU 或者其他编解码芯片（海思、TI、瑞芯微、全志等编解码芯片）来实现视频编解码过程；

优劣势分析：软件解码相对而言，视频解码过程完整，画质清晰，能够解码全部的视频格式，但由于通过 CPU 进行工作，一方面进一步会消耗所剩不多的 CPU 性能，同时也会产生较大的功耗和造成机器损耗，性价比低。硬件编解码往往相对视频播放流畅，但受限于芯片能力的影响，解码出来的视频格式受限。由于其相对较低设备功耗和高性价比，成为目前最为主流的视频编解码的应用，尤为突出的是显卡的广受欢迎便是因其强大的视频硬件编解码能力和相关图形处理能力。



秒杀节

最高9750元
600W 静音显卡电源

下单赠送显卡支架

¥11699.00

微星 (MSI) 万图师 RTX3070/3080/3090 万图师魔龙 电竞电竞游戏设计独立显卡
76条评价
攀升DIY电脑旗舰店

对比 关注



索泰 ZOTAC

¥18999.00

索泰(ZOTAC)RTX3080PGF OC 显卡/N卡/台式机/游戏/电竞/网课/绘图/设计/独立显卡
5000+条评价
索泰京东自营旗舰店

对比 关注



技嘉 GIGABYTE

江浙沪属日达 同城闪送

¥18299.00

技嘉 RTX3080 显卡 GAMING OC 魔鹰/猎鹰/雪崩白色10G及精英台式机/电竞游戏显卡
35条评价
格慕数码专营店

对比 关注



GIGABYTE 技嘉笔记本旗舰店

¥20999.00

【技嘉旗舰店】技嘉RTX3080 电竞游戏显卡 吃鸡独显 赛博朋克2077游戏显卡
39条评价
技嘉笔记本旗舰店

对比 关注



msi

预约中

¥15999.00

微星 (MSI) 超龙 GeForce RTX 3080 SUPRIM X 10G 超旗舰 电竞游戏设计智能卡
5万+条评价
微星京东自营旗舰店

对比 关注

广告