

# x265 视频压缩教程综合版

欢迎阅读! 若有什么不会的可以加群 [691892901](#). 本教程很难, 入门先看 [x264 视频压缩教程综合版](#). 但现在就要压视频就去拿[急用版教程](#)哦(·ω·)ゞ 用 ctrl+f, 让电脑帮你找内容((((\*. \_.)

## ffmpeg, VapourSynth, avs2yuv 传递参数

```
ffmpeg -i <源> -an -f yuv4mpegpipe -strict unofficial - | x265 --y4m - --output
ffmpeg -i <源> -an -f rawvideo - | x265.exe --input-res <分辨率> --fps <整/小/分数> - --output
-f 格式, -an 关音频, -strict unofficial 关格式限制, --y4m 对应"YUV for MPEG", 两个"--"是 Unix pipe 串流
VSpipe 源.vpy --y4m - | x265.exe - --y4m --output
avs2yuv 源.avs -csp<色> -depth<深> - | x265.exe --input-res <宽 x 高> --fps <整/小/分数> - --output
avs2pipemod 源.avs -y4mp | x265.exe --y4m - --output
```

**ffmpeg 查特定色度采样** `ffmpeg -pix_fmts | findstr <或 grep 关键字>`

**检查/选择色深, 版本, 编译** `x265.exe -V, -D 8/10/12 调整色深`

**多字体+艺术体+上下标.ass 字幕渲染** `ffmpeg -filter_complex "ass='F\:/字幕.ass'" 滤镜`

**中途正常停止压制, 封装现有帧为视频** 输入 Ctrl+C, x265.exe 自带功能

**Bash 报错自动导出+命令窗里显示** `x265.exe [参数] 2>&1 | tee C:\x265 报错.txt`

**8bit 还是 10bit 色深** x265 缩小了精度误差, 8-10 差距比 x264 小, 但 10bit 仍可减轻色带

## 目录

分块..... 1	关键帧..... 7	模式决策..... 16	-CABAC..... 19
变换·(附录-傅里叶变换)· 2	参考帧..... 8	率失真优化 RDO 控制... 17	二进制算数编码..... 19
帧间-动态搜索..... 3	帧内编码..... 8	峰值信噪比 PEAK SIGNAL-TO-NOISE RATIO/PSNR..... 18	SEI 维稳优化消息..... 20
帧间-子像素运动补偿.... 5	量化-码率质量控制模式· 9	环路滤波-去块滤镜.... 18	线程节点控制..... 21
加权预测 WEIGHTED PREDICTION · 5	2PASS 转场优化..... 13	环路滤波-取样迁就偏移 19	色彩空间转换, VUI/HDR 信息, 黑边跳过..... 22
帧间-时域架网搜索..... 6	率失真优化-码率质量控制调优..... 15	带偏移 BO..... 19	IO (INPUT-OUTPUT, 输入输出)..... 23
溯块向量搜索..... 6	自适应量化..... 15	熵编码/残差编码/文本压缩	
GOP 结构建立, 参数集... 7			

## 分块

hevc 中, 帧下结构按面积大小列为帧→瓦 tile→条带 slice→条带段 ss→ctu→cu 单元. cu 和 cb 是 ctu 经动静态隔离, 即动态搜索 motion estimation 与运动补偿 motion compensation 所得的结果. 其中 U 即 unit, 指 YCbCr 三位一体, 而 B/block 则单指其一. ipcm-cu 代表跳过 me-mc, 直达环路滤波的 intra pulse code modulation cu (帧内编码的 pcm 波形 cu), 因为"块"本来就是一串像素值的波形, 只是通过每个"单元"自带的元数据"换行"到了二维而已. intra 代表帧内, 当提及帧内编码, 多指 GOP 建立后, 用于为 B, b, P 帧提供参考源的 IDR 和 I 帧所需要用到的完整单图, 但也包括参考帧内所含的 I 块



**PU - 预测单元** prediction unit 是编码完, 可以用作参考的块. 支持 cu 上对称 rectangle, 非对称 asymmetric partition 划分, 以更好的隔离动静态. 亮度与色度上的分裂法可以不同, 小至 4×4 像素



图: pu 的 4 种对称 rectangular 和 4 种不对称 asymmetric 划分

**TU - 变换单元** transformation unit 的划分与 cu 而非 pu 同步, 实现变换和量化 (´ ▽ `)/



**AU - 存取单元** access unit 是解码端用于启动播放的块, 一般为 IDR-AU

**--ctu**<64/32/16, 默认 64>指定编码树单元最大大小的参数. CTU 越大, 有损压缩效率越高+平面涂抹越高+速度越慢. 一般建议保持默认, 但考虑到动画的大平面建议辅以低量化. 考虑画质优先时建议设<32>, 当分辨率特别小时建议设<16>且调整下面的参数(´~\*)/**--min-cu-size**<32/16, 默认 8>限制最小\ cu 大小, 简化计算步骤, 因为使往后步骤 pu, tu 的划分也会更大. 用多一点码率换取编码速度的参数. 建议日常环境使用 16 或快速编码环境使用 32

**--rect**<开关默认关, 受 limit-modes 限制>pu 对称划分, 用更多编码时间换取码率的参数. 只建议有比较充足时间, 分辨率大于 1440x810 或通篇颗粒的视频用;**--amp**<开关默认关, 须 rect, 受 limit-modes 限制>pu 不对称划分. 用更多编码时间换取码率的参数. 只建议通篇有大量粒子/噪点, 动漫源等分块高收益的视频用

## 变换·傅里叶变换

**一维傅里叶变换 1D-FT:** 给出与原信号波形等高, 从最长的频率周期开始不断缩窄(增加周期)的参考余弦. 在参考余弦波变化的过程中, 记下两条波形吻合度变化的曲线-不同波形周期的振幅, 就得到了频域

信号. 反过来将频域信号所对应的波形加起来就实现了逆变换. 为了将源波形中反相的余弦也考虑在内, 所以计算过程中要取立方转正. "不断缩窄的参考余弦"在低等数学上用  $\cos(1/T \times 2\pi) \{0 \leq T \leq n\}$  表示. 高等数学用  $e^{ix} = \cos(x) + i \cdot \sin(x)$ . 由复数螺的缩窄以实现核磁共振等看中复数兼容性的用途. 见 [desmos 示 1](#), [示 2](#), 及 [3b1b 视频](#)

**二维傅里叶变换 2D-FT:** 宽高上单拆出线来分别进行 1DFT, 通过双求和  $\Sigma\Sigma$  或双积分  $\int\int$  来缩写. 所以在频域中相当于每个像素的变换结果相加或干涉. 亮则振幅大, 远则频率高. FT 强在可编辑性, 是消除光盘扫图等均匀噪声的解

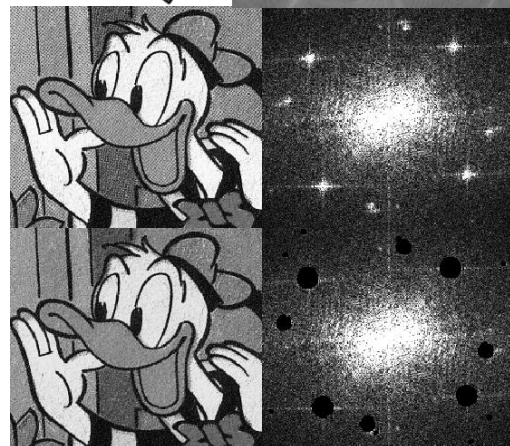
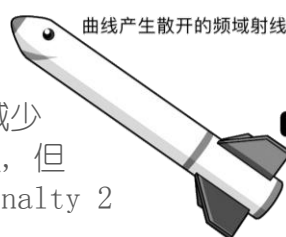
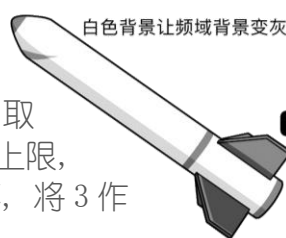
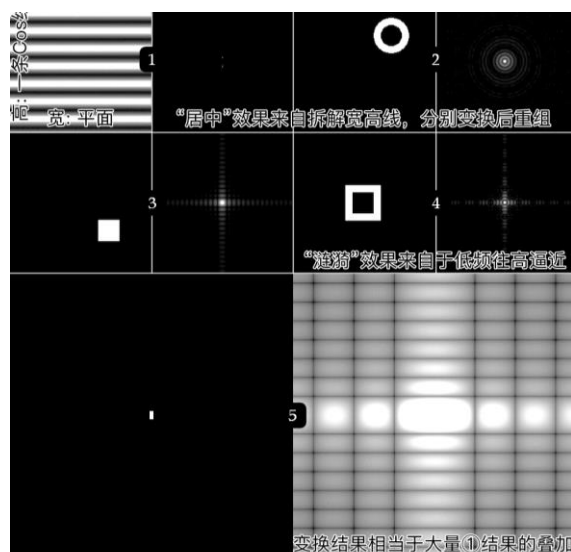
**二维离散余弦变换 2D-DCT:** 预制的二维波形模具, 穷举加减列出每个频段的使用次数, 图像就从空间域转换到频域了. 比 2DFT 快但只有  $8 \times 8$  种波形, 所以删高频效果比 FT 差

**--limit-tu** <整数  $0 \sim 4$  默认关, **tu-intra/inter-depth** 大于 1> 提前退出 tu 分块, 以量化/残差编码质量为代价提速. **tu** 太则易出现量化涂抹, 不利于暂停画质. 1 一般, 画质编码, 取分裂/跳过中花费最小的, 2 以同 ctu 内的首个 tu 分裂次数为上限, 3 快速编码取帧内帧间附近 tu 分裂平均次数为上限, 4 不推荐, 将 3 作为未来 tu 的分裂上限, 相比 0+20%速度

**--rdpenalty** <整数  $0 \sim 2$ , 默认关, **tu-intra-depth=1** 时失效; =2 则  $32 \times 32$  帧内 cu 可用; =3 才支持  $64 \times 64$  帧内 cu> 强制 tu 分块以提高细节保留降低涂抹. 1 提高率失真代价而减少  $32 \times 32$  tu, 或设 2 强制  $32 \times 32$  tu 分块. 用途与 limit-tu 相反, 但可理解为 tu 分块的下限, 例如高 limit-tu, 高 crf 时用 rdpenalty 2 避免  $32 \times 32$  tu 造成涂抹太强画面糊掉的结果

**--tu-intra-depth, --tu-inter-depth** <整数  $1 \sim 4$ , 配合 limit-tu, 默认 1> 空间域 tu 分裂次数上限, 默认只在 cu 基础上分裂一次. 决定量化质量所以建议开高. 建议日常编码设在 2, 提升画质设  $3 \sim 4$

**--max-tu-size** < $32/16/8/4$ , 默认 32> 更大的 tu 大小能提高压缩, 但也造成了计算量增加和瑕疵检测能力变差. 码率换时间+画质. 建议不如直接设 ctu, 因为也可减少  $32 \times 32$  tu



## 帧间-动态搜索

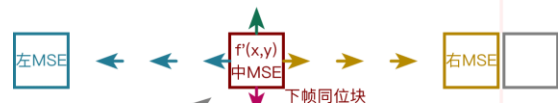
于帧间逐块地找最小失真朝向 dir. of min. distortion, 组成一张张帧间矢量表的计算. 若找出的信息不足, 参考帧与动静态隔离分块的建立就欠缺基础. 图: [Jain&Jain/十字搜索](#), [大小菱 LS-Dia 搜索](#)

## J&J搜索

确立大致方位后逐步细化, 找出唯一块

直径 $D=8$ ,  $\leq \text{merange}$

半径 $r=D \div 2 + 1$



递

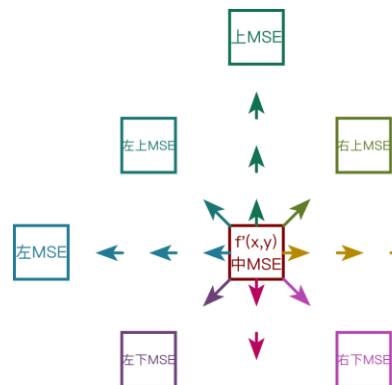
上帧块  
 $f(x, y)$

假设理想动态位移点→

迭代Iteration/位移Displacement 1

- 根据merange限制最大像素数, 为下帧同位块建立J&J搜索范围
- 4x4块和旁边8x8块不一样大, 但矢量可能等长, 故merange不直指块数
- J&J也叫十字搜索, 首次迭代分为上下左右中, 共对比5次MSE
- 半径r长度为 $D \div 2$ , 加上中心1

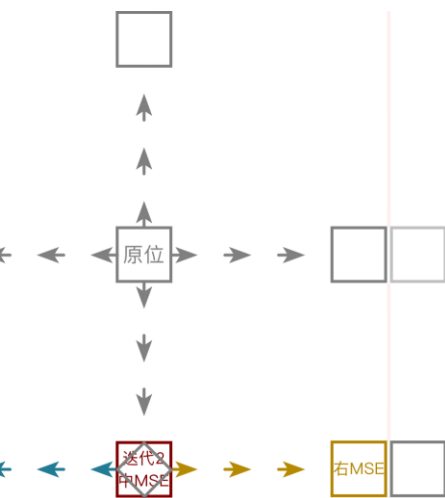
- 若中块与上帧块差异最小, 则下次迭代半径 $r+2$
- 若半径 $r=1$ , 则指定最终迭代, 进行全搜索



迭代/位移

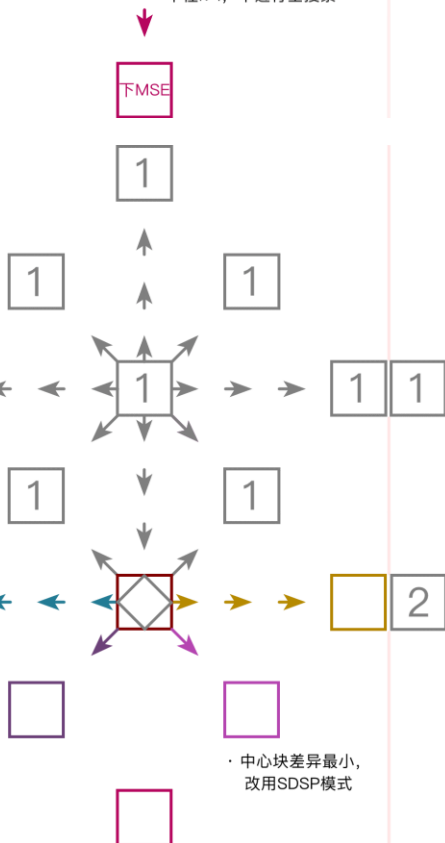
- 根据merange限制最大像素数, 及当前块大小决定直径D(块数)
- 分大菱形Large Diamond ShaPe (LDSP)和小菱形SDSP, x264中只有SDSP
- LDSP迭代到“最小代价MSE或SAD位于中心块”时, 切换SDSP来细化搜索
- x264代码中默认中心块代价最小, 否则边代价更小的块改为中心块以优化性能

- LDSP模式 - 下方块MSE(PSNR)/SAD/MAD与上帧块差异最小
- 若中心块差异最小, 则改用SDSP模式



迭代Iteration/位移Displacement 2

- 中块与上帧块差异最小, 下次迭代时 $r+2$
- 半径 $r>1$ , 不进行全搜索



六边形 hexagonal 搜索将 LDSP 的上下左右斜, 共 8 个外点砍成 6 个, SDSP 细化规则不变, 相比 LS-dia 和 SSSP-LDSP-SDSP(四角星 me star)在小范围里更高效

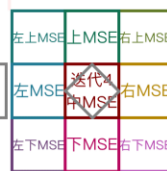


迭代/位移 3

- 中块与上帧块差异最小, 下次迭代时 $r+2$
- 半径 $r>1$ , 不做全搜索



--me<hex~full>搜索方式, 从左到右依次变得复杂, umh 平衡, star 之后收益减. star 四角星搜索, sea 是优化过的 esa 穷举, 浪费性能--



迭代/位移 4

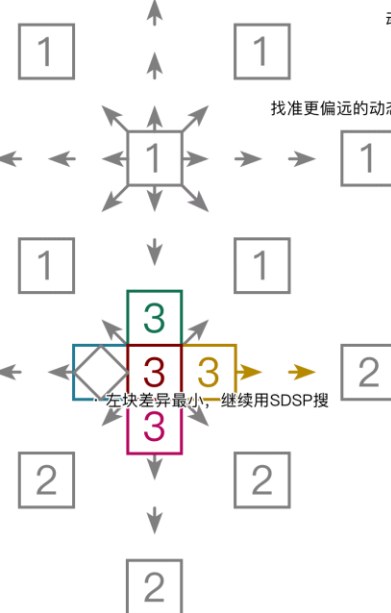
- 半径 $r=1$ , 做全搜索
- 结论: 下移4块

Ref: youtube.be/jRrs9OxLTYM

merange<整数搜索范围, 据动搜算法选>简单说 hex 选 16, umh-star 选>=32; 一般推荐 me umh merange 48. 精致一些需要 DIY 测量距离. 太大会同时降低画质和压缩率, 因为找不到更好的, 找到也是错的), 所以 48 左右顶天, 同时建议用 8 的倍数

LS-Dia大小菱搜索 (LDSP+SDSP)

动态矢量位移比十字搜索更偏僻  
迭代次数却能做到与之相等  
证明了虽然搜索点位越多越慢  
但减少迭代次数就是效率增加  
找准更偏远的动态矢量才能应对高分辨率的压缩

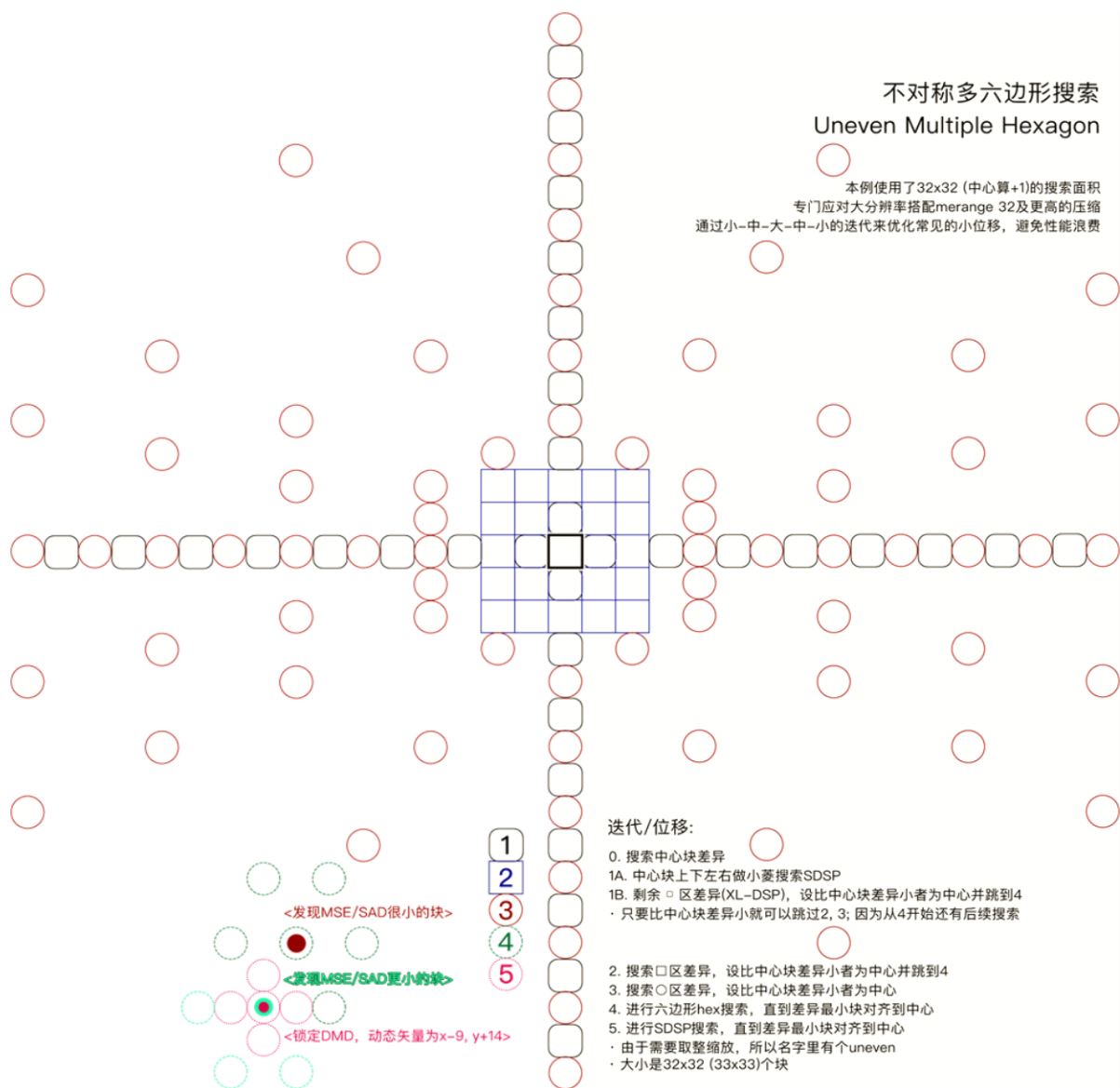


--analyze-src-pics<开关, 默认关>允许动态搜索  
查找片源帧, 耗时增加压缩



· 共下移4块, 左移1块





## 帧间-子像素运动补偿

motion compensation 将动搜所得信息做块-帧插值, 让帧间连贯起来. 防止畸变相对复杂的动态信息让块脱离参考压缩. 冲激响应滤镜 imp. response filter 对超阈值的输入模拟信号出 1, 否则出 0 的滤镜. 冲激~响应与音符~波形的关系所同. hevc 标准要求使用 7-tap 精度(1/4 像素), avc 要求 6-tap. 影响模式决策和率失真优化. SAD, SATD 计算见 x264 教程完整版

--subme<整数默认 2, 范围 1~7, 24fps=4, 48fps=5, 60fps=6, +=7>根据片源的帧率判断. 分四范围. 由于动漫片源制于 24~30fps, 因此可调低; 但同是动漫源 60fps 的虚拟主播则异. 由于性能损耗大, 所以不建议一直开满. 由于 x264 中 rdo 选项直接塞进了 subme, 所以相比 x265 偏高

推荐范围	值	HPel 迭代	HPel 搜索方向	QPel 迭代	QPel 搜索方向	算法
30fps 搭配 rdo	<3>	2 次	4	1	4	SATD
48fps 搭配 rdo	<4>	2 次	4	2	4	SATD
60fps 搭配 rdo	<5>	1 次	8	1	8	SATD
90fps 搭配 rdo	<6>	2 次	8	1	8	SATD
144fps 搭配 rdo	<7>	2 次	8	2	8	SATD

## 加权预测 weighted prediction

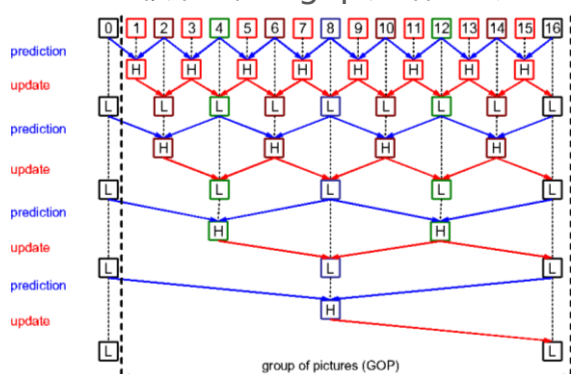


x264 首发, 修复了少数淡入淡出过程中部分 pu 误参考, 亮度变化失去同步的瑕疵. 分为 P, B 条带用的 **显加权 explicit WP**<编码器直接从原画和编码过的参考帧做差>与 B 条带用的**隐加权 implicit WP**<用参考帧的距离插值>插值计算在帧内编码板块有说明. x265 默认开 weightp

--weightb<开关, 默认关>启用 B 条带的显, 隐加权预测. 条带所在 SPS 中可见 P, B 加权开关状态, 及显加权模式下解码器须知的权重. 光线变化和淡入淡出在低成本/旧动漫中少见. 略提高压缩和画质

## 帧间-时域架网搜索

motion-compensated temporal filter 基于快而简单的提升式小波变换 lifting-scheme. 只需前后信息通过 1-2 之差来预测 2-3 的区别, 出的残差更新到高频分区, 没预测错的更新到低频分区. 帧间就是照奇偶奇帧推演, 从 gop 开始到末尾的 0-1-2, 2-3-4, 4-5-6 帧组成第零层奇偶奇网络, 通过分离的低



频 L 带和高频段残差 H 带是第一层网络. 一层的 L 带再重组为 0-1-2, 2-3-4 等计算奇-偶-奇推演, 得到更远的动静态区间, 即 LL 带, LH 带的第二层网络. 该结构做到了迭代 n 次即分离  $2^n$  帧的动静态, 以及所有的预测与残差, 故不像传统动态搜索一样受分辨率大小的缩放性 scalability 限制. 是 svc 的核心算法. 非常依赖转场计算

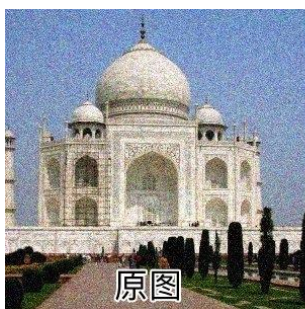
**一维小波变换 1D-WT** 让短波像拉链一样划过一维信号, 此外和 1DFT 一样得到频域信号-热度图的变换算法. 解决了傅里叶变换只有空-频域, 把一整个信号变换而不能描述变化过程的问题, 还可以像拉链一样更换波形. 但缺点是分辨率低, 可编辑性差. 加上 mctf 对预测和残差的追踪后就可以分析视频了

--mcstf<开关, 要求 32GB 内存? 会关闭多线程>mctf 帧上应用空域双边滤镜, 是迭代到最后的 L 帧?

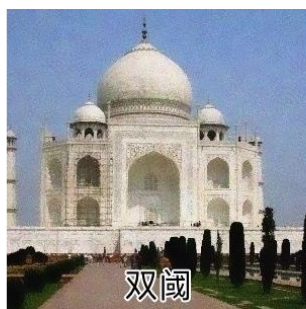
**高斯模糊**是设定范围阈值-模糊波形圆周大小来决定强度的滤镜. 比方块模糊柔和, 在媒体和设计上常用. **双阈滤镜** bilateral filter 加了最小纹理强度的判定, 只模糊振幅超阈的像素. 所以瞎起名叫双阈

## 溯块向量搜索

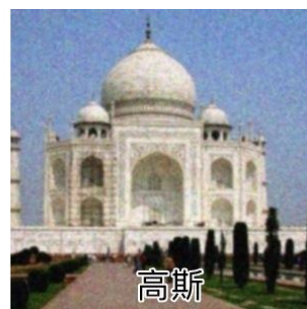
与帧内编码并行, 给动态搜索提供溯块向量(cu 帧内/帧间朝向, 大小)的步骤. 由于移动的物件会



原图



双阈



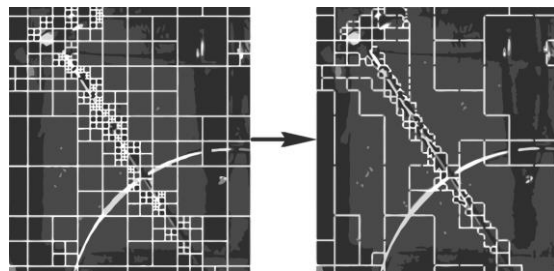
高斯

跨越多个 pu. 所以将涉及同物件的块匹配到一起就能冗余一大批 pu 的动态向量了. hevc 与 avc 一样用 ref 参数于时域上划区, 逐 pu 创建 List0 和 List1 左右两排参考列表. 差别是 hevc 在 avc direct auto 的基础上升级了**高级向量预测 AMVP**; 并合搜索 merge mode 两种方案. 其中 adv. motion vector prediction 的任务是找出向量信息, 分以下几步:

1. 在帧内看当前 pu 左下的邻 pu, 优先匹配向量往帧内指的邻 pu

2. 参考那些向量往它帧指的临 pu；并等比缩放，对齐到邻 pu 已按帧间差异对齐好的向量
3. 若以上步骤没找到参选向量，就把同样的步骤于当前 pu 右上角进行一次
4. 若应了如早批 pu 刚开始算，找不到参选向量的情况下就直接从时域搜索：照帧间参考图像变化的内容差异做缩放调整，从右下角的相邻 pu 找参选
5. 若仍不可用，就找当前 pu 中心位置的其它同位 pu。若最后没凑不齐两个参选向量，代入  $v=0,0$

merge 简单粗暴，从时空域凑五参两被。漏算 pu 边缘且不顾 pu 当前向量以提速，所以可看做是打下手的



--ref<整数  $1 \sim 16$ ，推荐  $\text{fps} \div 100 + 3.4$ >向量溯块前后帧数半径，一图流设 1。要在能溯全所有块的情况下降低参考面积，所以一般设 3 就不管了

--max-merge<整数  $0 \sim 5$ ，默认 2>重设 merge mode 被选数量。用更多时间换取质量的参数。建议高压编码设<4>，其它可设<2, 3>(+\_+)

--early-skip<开关默认关，暂无建议>先查  $2n \times 2n$  merge 被选，找不到就关 AMVP

## GOP 结构建立, 参数集

给视频帧分段并最终整合成 gop 内树叉状的参考结构后，将其中的关键帧递给下一步帧内编码。一来冗余，二来防止参考错误蔓延，照顾丢包人士，二来搭建 NALU 为基础传输 ss 的网络串流架构

1. 按 IDR 帧间隔(keyint)分区，同时 scenecut 分配额外关键帧
2. 按 open/closed-gop 标记 gop 间预设，同时 gop 内的帧拆为条带 slice
3. 条带一样要拆开来以降低解码错误的影响，叫做条带段或 ss

视频参数集 video parameter set→(分枝-播放时间戳，显加权与其它特定解码要求)序列参数集

sequence parameter set→(分枝-解码信息)图参数集 picture parameter set→(分枝-ctu 以上最小单位)条带段 slice segment

--opt-qp-pps<开关，默认关>据上个 GOP 改动当前 PPS 中默认的 qp 值.--opt-ref-list-length-pps<开关，默认关>据前 GOP 改当前 ref 值，而且是前后帧独立改动。<暂停推荐：播放器默认 PPS 跨 GOP 不变，造成解码兼容性错误。应该用`hev1`而非`hvc1`封装进 ISO-BMFF?>

--repeat-headers<开关，已关>在流未封装的情况下提供 SPS，PPS 等信息，正常播放 h.265 源码

--scenecut<整数>设  $x264/5$  设 I 帧的两帧差距触发转场阈值

--hist-scenecut<开关，默认关，推荐开>亮度平面边缘+颜色直方图 SAD 阈值触发转场。在  $x265v3.5+69$  后在彩色视频中超越了近 20%，尤其是 HDR 源比 scenecut 降低了正误判(设 I 帧)和负误判(不设 I 帧)，因此除黑白视频外推荐。12bit 源下会导致编码中途失败，目前由于没有配套的压缩逻辑，所以提高转场精度对于压缩和提速的效果不明。(hist-threshold 于  $x265 v3.5+69$  被移除)

## 关键帧

idr 刷新解码帧 instant decoder refresh

- gop 间划界分段，令解码器清缓存的完整图片的 I 帧，清缓存是为了防治参考/内存错误(¬\_¬) /

### cra 净任意访问 clean rand. access

- open-gop 间划界，带帧内参考，gop 内帧间参考可越界的 I 帧，一般直接叫 cra 帧

### dra 脏任意访问 dirty rand. access

- 一组含 i 块，全解码才重建出 i 帧的 P 帧。压缩更高但比 i 帧更易出错。需要低 min-keyint

### bla 断链访问帧 broken link access

- open-gop 间划界，访问并加载出异分辨率，帧率视频流用的特殊 cra 帧、(¬\_¬) /

## 参考帧

rap/随机访问点 random access point "访问"代表播出画面前读数据的过程; "任意"代表拖进度条, 打开直播, 使进度条上任意一点都要能解出视频的需求, 增加码率提升体验

### rasl 任访略前导, radl 任仿解前导 random access skipping/decoding lead

1. 正常播过来没它们事, 但进度条落在 cra 附近(缺参考)时指定解码/略过的前导帧。防止拖进度条让 gop 崩坏

--no-open-gop<开关, 默认关, 建议长 gop 用>不用 cra/bla, 增加码率增加兼容

--radl<整数默认 0, 小于连续 B 帧, 建议 2~3>原理见上

--min-keyint<整数>指定最小 IDR 帧间隔。防止编码器在 closed-gop 里将两个 IDR 帧挨太近, 导致 P 和 B 帧参考距离受限而设计的。两种选择, 给出画质一样

- 设 5 或更高, 省了设立一些 IDR 帧拖慢速度。快速编码/直播环境直接设=keyint ~(>\_<~)
- 设 1 来增加 IDR 帧, 一帧被判做转场本来就意味着前后溯块的价值不高。而 P/B 帧内可以放置 I 宏块, x264 会倾向插 P/B 帧。好处是进度条落点在激烈的动作场面更密集, 画质编码用

--keyint<整数>指定最大的 IDR 帧间隔, 单位为帧。由于 min-keyint 有设立 IDR 帧的能力, 建议照不精确索引下拖动进度条的偏移延迟 vs 码率设置。--keyint -1 即 infinite。在长度短到不需要拖动进度条, 或者用户一定不会拖动进度条的视频可以使用以降低码率(●\_●)

--fades<开关, 默认关>找流中的虚实渐变 fade-in, 给小到帧间条带(slice, 一组横向 ctu), 整个帧间范围改用 I 条带, 并根据渐变后最亮的帧重设码率控制历史记录, 解决转场致模糊的问题

--pbratio<浮点, 默 1.3>P, B 帧的 qp 值待遇差(如 B 块至少是 P 块 1.3 倍 qp)。B 帧双向参考能从更多帧中找到参考源, qp 更高也能通过参考来达到相同画质。真人录像片源中保持默认即可。动漫片源中 B 帧的出现几率增多, 导致很多 B 帧因找不到合适的参考信息损失画质。所以编码画质的动漫时要通过降低 B 帧的 qp 值分配来保护其画质, 通常使用<1.2>或更激进

--bframes<整数 0~16>最多可连续插入的 B 帧数量。<3~5>快, <8>电影平衡, <12 左右>正常, 若播放设备配置偏高的话可放心的设在<13 左右>, 要注意 bframes>8+keyint>250 会大增内存占用。--b-

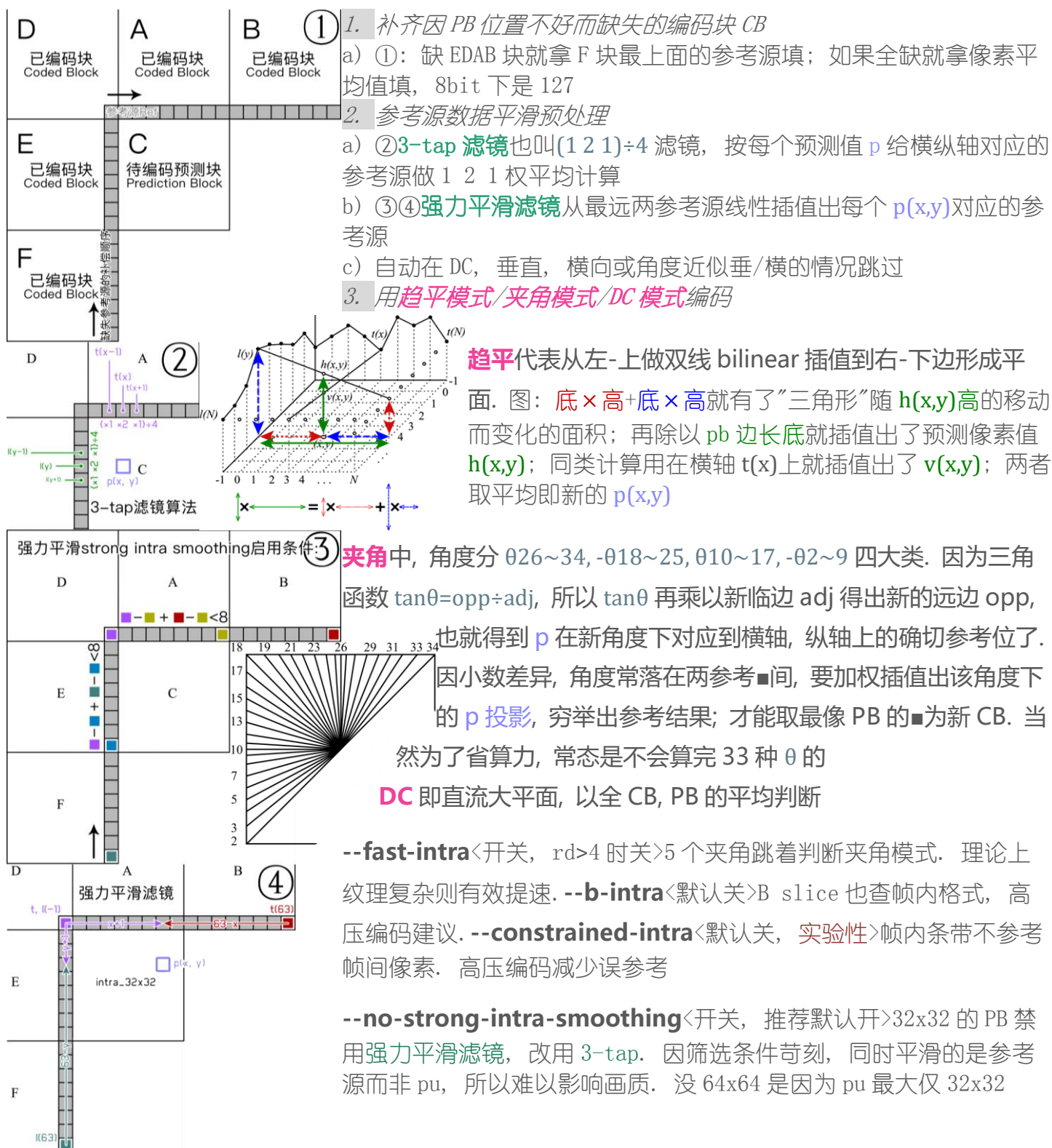
adapt<整数默认 1 快, 推荐 2 精>B 帧适应性算法

--bframe-bias<整数-90~100, 推荐默认>设 B 帧阈值偏移, 增大适合压缩放 B 帧收益大的片源, 如桌面录屏和低成本动漫。调整此值不会带来性能优化

## 帧内编码



组成参考源+参考帧的结构后, 对参考源做类 jpg, png 的单图压缩



## 量化-码率质量控制模式

对数  $\log(x)$  是为方便乘除法交换律, 即路程=速度×时间关系实现到指数关系上, 构成  $9=3^2$ ,  $3=2^{\sqrt{9}}$ ,  $2=\log_3(9)$  对应  $6=2 \times 3$ ,  $3=6 \div 2$ ,  $2=6 \div 3$  关系的人造算法. 编码器用  $\log$  算 qp 值, 因为 qp 是对数, 对应着指数变化的量化强度, crf18~23 外的码率画质变化中能体会到

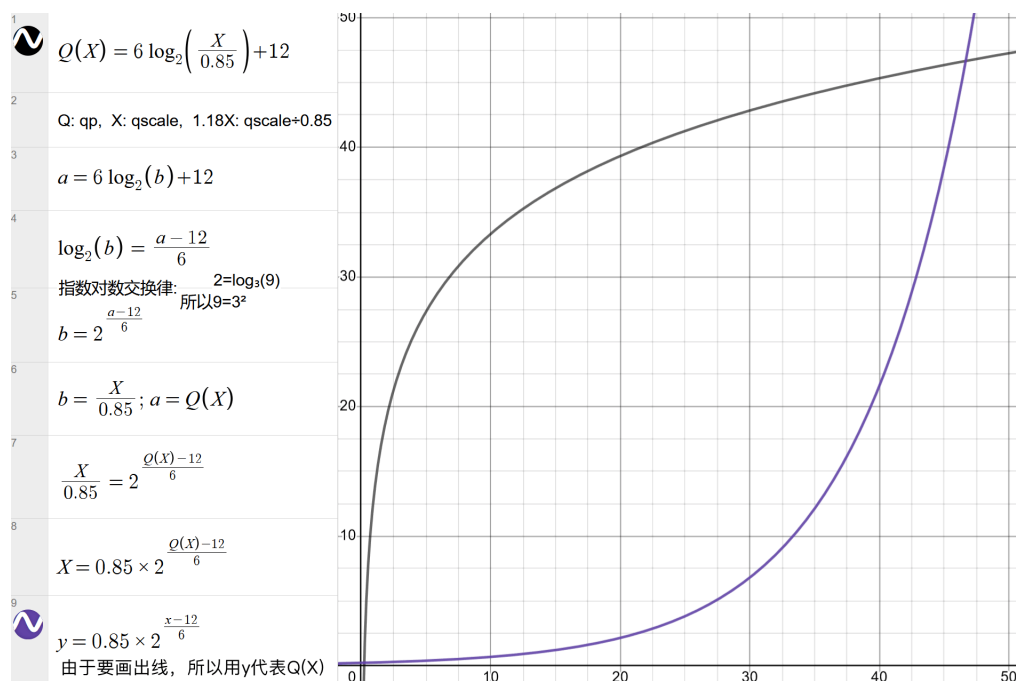


图: 紫 qp-qScale, 黑 qScale-qp 的对数映射

由于当前帧未编码, 所以只能用之前的编码帧预测复杂度, 因此叫推演复杂度/模糊复杂度. 而量化是有损压缩, 所以复杂度由当前量化值下, 已编码帧的失真程度推演, crf/abr 高低决定复杂度而反直觉. 所以 cplxBlur 的精确度一般取决于帧率. 虽然基本上不准, 但 qp 波动变小, 码率变化更稳定

**cplxSum:**  $0.5 \text{ cplxSum}[\text{上帧}] + \text{SATD}[\text{上帧}]$  <迭代值, 初始=半数宏块 ÷ CTU, 逻辑是 SATD[当前]=SATD[旧]+SATD[当前]>查不出为什么要 +2

**cplxCount:**  $0.5 \text{ cplxCount}[\text{上帧}] + 1$  <迭代值, 初始 0, 当前帧数计数, 用于逐帧对 cplxBlur 加权> ÷ 2 是与 cplxSum 同步, 加权是为了推演越往后参考冗余越好, 压缩越强的规律

**cplxBlur:**  $\text{cplxSum} \div \text{cplxCount}$  <迭代值, 据帧所处位置推演加权的新 SATD>接近 100%则说明当前帧复杂度应当往高推演, 画面复杂度处上升趋势. 可扭转 cplxCount 默认的下趋势

**qScale:**  $85\% (2^{((qp - 12) \div 6)})$  <迭代值, 已编码旧帧 qp 转 qScale, 方便其它参数修改更新>

**ABR\_rate\_factor:**  $\text{target\_rate\_window} \div \text{cplxSum}$  <迭代+初始值, abr 下的 qScale 转 qp>

**ABR\_newqScale:**  $\text{qscale} \times \text{overflow} \div \text{ABR\_rate\_factor}$  <迭代结果, 据 abr 控制更新一遍 qScale>

**cplxBase:**  $\text{ctu\_count} \times (\text{bframes} ? 120:80)$  <恒定值, crf 模式默认的复杂度. 若设 bframes 则 CTU 或宏块数量 × 120, 否则 × 80. 注: a?b:c 是 C++中 if a:b else:c 的缩写>

**CRF\_rate\_factor:**  $\langle \text{多参数值} \rangle^{(1-\text{qcomp})}$  是为和 CRF\_qScale 对齐, 因为运算仅在 cplxBase, cplxBlur 上而非整个算式中发生. 最终经 cutree, b

$$\text{cplxBase}^{(1-\text{qcomp})} \div (\text{qScale} \times (\text{crf} + \text{cutree} \& \text{bframe\_offset}))$$

帧偏移再乘进 qScale 后得到实现质量-码率控制的值

## CRF\_newqScale:

〈据 crf 控制更新一遍 qScale〉

$$\text{cplxBlur}^{(1-\text{qcomp})} \div \text{CRF\_rate\_factor}$$

**newqp:**  $6\log_2(\text{qscale\_new} \div 85\%) + 12$  〈目标值〉qScale 经调整后算出当前帧 qp

最终, 实现了帧内画面简单, qp 值(压缩)高; 反过来越复杂 qp 越低的压缩理念. 然而这种质量判断只有两帧, 不够宏观(abr 模式更严重), 所以还有率失真优化量化的必要

## CRF 上层模式

**--crf**〈浮点范围 0~69, 默认 23〉据“cplxBlur, cutree, B 帧偏移”给每帧分配各自 qp 的固定目标质量模式, 或简称质量呼应码率模式, 统称 crf. 素材级画质设在 16~18, 收藏~高压画质设在 19~20.5, YouTube 是 24. 由于动画和录像的内容差距, 动画比录像要给低点. 理论上 crf 高=量化损失多, 率失真优化也就越慢; 但测试出来是 crf+2, 4k4:4:4 12bit 会快~0.5fps

**--qpmin**〈整数, 范围 0~51〉最小量化值. 由于画质和优质参考帧呈正比, 所以仅在高压环境建议设 14~18. **--qpmax**〈同上〉在要用到颜色键, 颜色替换等需要清晰物件边缘的滤镜时, 可以设 --qpmax 26 防止录屏时物件的边缘被压缩的太厉害, 其他情况永远不如关 mbtree (\*~▽~)

**--rc-grain**〈开关, tune grain 时开启〉通过 cplxBlur 抑制 qp 判断被噪声带偏, 胶片颗粒片源用

**--cplxblur**〈浮点 0~100, 默认 20〉第-1 帧不存在, 无法算出第 0 帧的 cplxBlur 所以直接指定

**--qcomp**〈浮点范围 0.5~1, 一般建议默认 0.6〉cplxBlur 迭代值每次能迭代范围的曲线缩放. 越小则复杂度迭代越符合实际状况, crf, mb-cutree, bframes 越有用, 搭配高 crf 能使直播环境可防止码率突增. 越大则 crf, mb-cutree, bframes 越没用, 越接近 cqp. 曲线缩放原理见 [desmos 互动示例](#)

**--rc-lookahead**〈整数, 范围 1~250〉指定 cutree 的检索帧数, 通常设在帧率的 2.5~3 倍, 若通篇的画面场景非常混乱则可以设在帧率的 4 到 5 倍通常在 180 之后开始增加计算负担

**--no-cutree**〈开关〉关闭少见 CTU 量化增强偏移. 只有近无损, 可能 crf 小于 17 才用的到

**\*rceq**〈仅 x264, 字串, 默认  $\text{cplxBlur}^{(1-\text{qComp})}$ 〉可以少算一步 qcomp, cplxBase 还需 qcomp, 但不写 qcomp 就是推荐的 0.6, 所以不用写也行

## CQP 上层模式

**--qp**〈整数, 范围 0~69〉恒定量化. 每 ±6 可以将输出的文件大小减倍/翻倍. 直接指定 qp 会关 crf, 影响其后的模式决策, 综合画质下降或码率暴涨, 所以除非 yuv4:4:4 情况下有既定目的, 都不建议

**--ipratio**〈浮点, 默认 1.4〉P 帧相比 IDR/i, 以及 **--pbratio**〈浮点, 默认 1.3〉B/b 帧相比 P 帧的偏移. 如要求 IDR/I 帧 qp17, P 帧 qp20, B/b 帧 p22 时填写 --qp/crf 17 --ipratio 1.1765 --pbratio 1.1. ratio 即比率, 如 I 帧 qp10, 那么 ipratio 1.4 就代表  $10 \times 1.4 = 14$ , 就得到 P 帧 qp14 了

## ABR 上层模式

编码器自行判断量化程度, 尝试压缩到用户定义的平均码率 `average bitrate` 上, 速度最快

**--bitrate**<整数 kbps>平均码率. 若视频易压缩且码率给高, 就会得到码率更低的片子; 反过来低了会照顾画质强行提高量化, 使码率达标. 如果给太低则会得到码率不达标, 同时画质\*\*的片子. 平均码率模式, 除 2pass 分隔, 一般推流用的“码率选项”就是这个参数, 速度快但同时妥协了压缩. 因此算力够的直播建议用 `crf~vbr` 模式, 码率>画质, 但画质也抓的压片用 `1pass-crf+2pass-abr`

## VBR 下层模式

**--vbv-buFSIZE**<整数 kbps, 小于 `maxrate`>编码器解出原画的每秒最大码率缓存.  $\text{buFSIZE} \div \text{maxrate} =$  编码与播放时解出每 gop 原画帧数的缓冲秒数, 值的大小关联编完 GOP 平均大小. 编码器用到是因为模式决策要解码出每个压缩步骤中的内容与原画作对比用

**--vbv-maxrate**<整数 kbps, buFSIZE 的 x 倍>峰值红线. 防止多个>buFSIZE GOP 连续累积, 给出缓帧启用高压的参数. 对画质的影响越小越好. 当入缓帧较小时, 出缓帧就算超 maxrate 也会因缓存有空而不被压缩. 所以有四种状态, 需经验判断 GOP 大小(“▽”)

- 大:  $\text{GOPsize} = \text{buFSIZE} = 2 \times \text{maxrate}$ , 超限后等缓存满再压, 避开多数涨落, 适合限平均率的串流
- 小:  $\text{GOPsize} = \text{buFSIZE} = 1 \times \text{maxrate}$ , 超码率限制后直接压, 避开部分涨落, 适合限峰值的串流
- 超:  $\text{GOPsize} < \text{buFSIZE} = 1 \sim 2 \times \text{maxrate}$ , 超码率限制后直接压, 但因视频小/crf 大所以没起作用
- 欠:  $\text{GOPsize} > \text{buFSIZE} = 1 \sim 2 \times \text{maxrate}$ , 超码率限制后直接压, 但因视频大/crf 小所以全都糊掉
- 由于 gop 多样, 4 种状态常会出现在同一视频中.  $\text{buf}^{\sim}\text{max}$  实际控制了这些状态的出现概率

**--crf-max**<整数>防止 vbv 把 crf 拉太高, 可能适合商用视频但会导致码率失控; **--crf-min**<整数>用途不明, 可能是反留白习惯所致, 目前 --qpmin 足以[-\_-]卜

## SBRC-CRF 上层模式:

分段式率控制 `segment based rate control` 是为视频平台实现 DASH, M3U8 串流用的功能

**--sbrC**<启用分段式率控制, 要求 `min-keyint=keyint`, `no-open-gop`>由于提高了初始 crf 值的利用率, 所以建议搭配 `--cplxblur=crf` 使用

**SBRC-CRF-VBR 双层模式:**SBRC-CRF 基础上增加两个 VBV 参数, 同时在新分段上重设 VBV 缓存

**SBRC-ABR 上层模式:**最适合直播推流, 就是把 CRF 模式下的参数换成 bitrate

## 2pass-ABR 双层模式

先用 crf 模式分析整个视频总结可压缩信息, 后根据 abr 模式的码率限制统一分配量化值. 有 pass 2 给特别高的平均码率, 输出最小损失的最小体积近无损模式, 以及 pass2 给码率硬限的全局整体压缩模式



**--pass 1** <导出 stats>; **--pass 2** <导入 stats>; **--stats** <文件名>默认在 x265 所在目录下导出/入的 qp 值逐帧分配文件，一般不用设

**--slow-firstpass**<开关>pass1 里不用 fast-intra no-rect no-amp early-skip ref 1 max-merge 1 me dia subme 2 rd 2，也可以手动覆盖掉

## Analysis-2pass-ABR 模式

在普通 2pass 基础上让 pass1 的帧内帧间分析结果 pass 到 pass2，减少计算量

**--analysis-save, --analysis-load**<"文件名">指定导入/出 analysis 信息文件的路径，文件名

**--analysis-save-reuse-level, --analysis-load-reuse-level**<整数 1~10，默认 5>指定 analysis-save 和 load 的信息量，配合 pass1 的动态搜索，帧内搜索，参考帧等参数。建议 8/9

- <1>储存 lookahead
- <2==4>+同时储存帧内/帧间向量格式+参考
- <5==6>+rect/amp 分块
- <7>+8x8cu 分块优化
- <8==9>+完整 8x8cu 分块信息
- <10>+所有 cu 分析信息( ^.^ )/

**--dynamic-refine**<开关，已关闭>自动调整 refine-inter，x265 官方建议搭配 refine-intra 4 使用，相比手动设定提高了压缩率

**--refine-inter**<整数 0~3，默认 0>限制帧间块的向量格式，取决于 pass1 分析结果是否可信

- <0>完全遵从 pass1 的分块深度和向量格式
- <1>分析所有 pass2 中与 pass1 相同分块的向量格式，除 2pass 中比 1pass 更大的分块
- <2>一旦找出最佳的动态向量格式就应用于全部的块，2Nx2N 块的 rect/amp 分块全部遵从 pass1，仅对 merge 和 2Nx2N 划分的块的动态向量信息进行分析
- <3>保持使用 pass1 的分块程度，但搜索向量格式

**--refine-intra**<整数 0~4，默认 0>限制帧内块的向量格式，取决于 pass1 分析结果是否可信

- <0~2>同上，<3>保持使用 pass1 的分块程度，但优化动态向量；<4>=pass1 丢弃不用

**--refine-mv**<1~3>优化分辨率变化情况下 pass2 的最优动态向量，1 仅搜索动态向量周围的动态，2 增加搜索 AMVP 的顶级候选块，3 再搜索更多 AMVP 候选 ( ° - ° ; ) ノ`

**--scale-factor**<开关，要求 analysis-reuse-level 10>若 1pass 和 2pass 视频的分辨率不一致，就使用这个参数

**--refine-mv-type avc** 读取 api 调用的动态信息，目前支持 avc 大小，使用 analyse-reuse 模块就用这个参数+avc (原文解释的太模糊，且未测试)

**--refine-ctu-distortion**<0/1>0 储存/1 读取 ctu 失真(内容变化)信息，找出 pass2 中可避的失真

## 2pass 转场优化(内容已落后，待更新)

**--scenecut-aware-qp**<整数, 默认关, 2 仅转后, 1 仅转前, 推荐 3 前后降低, 仅 pass2 用>转场前/后拉低默认 5 qp 以增加画质. 原理是转场本身就缺参考源, 所以提高已有参考源的画质

**--masking-strength**<逗号分隔整数>于 sct-awr-qp 基础上定制 qp 偏移量. 建议根据低~高成本动漫, 真人录像三种情况定制参数值. scenecut-aware-qp 的三种方向决定了 masking-strength 的三种方向. 所谓的非参考帧就是参考参考帧的帧, 包括 B, b, P 三种帧... 大概

- sct-awr-qp=1 时写作<转前毫秒(推 500)>, <参考±qp>, <非参±qp>
- sct-awr-qp=2 时写作<转后毫秒(荐 500)>, <参考±qp>, <非参±qp>
- sct-awr-qp=3 时写作<转前毫秒>, <参考±qp>, <非参±qp>, <转后毫秒>, <参考±qp>, <非参±qp>
- scenecut-window, max-qp-delta, qe-delta-ref, qp-delta-nonref<被 x265 v3.5 移除>

**--analysis-reuse-file**<文件名, 默认 x265\_analysis.dat>若使用了 2pass-ABR 调优, 则导入 multi-pass-opt-analysis/distortion 信息的路径, 文件名

## Analysis-Npass 间调优

在 Analysis-pass1~2 之间加一步优化计算. 实现比普通 2pass 更精细的码率控制, 1~N 也行

**--multi-pass-opt-analysis**<开关, 默认生成 x265\_analysis.dat>储存/导入每个 CTU 的参考帧/分块/向量等信息. 将信息优化, 细化并省去多余计算. 需关闭 pme/pmode/analysis-save|load

**--multi-pass-opt-distortion**<开关, 进一步分析 qp>根据失真(编码前后画面差). 需关闭 pme/pmode/analysis-save|load

**--multi-pass-opt-rps**<开关, 已关>将 pass1 常用的率参数集保存在序列参数集 SPS 里以加速

## Analysis-pass2-ABR 天梯

**--abr-ladder**<实验性的苹果 TN2224/据说整出一堆 bug, 文件名.txt>编码器内部实现 analysis 模式 2pass abr 多规格压制输出. 方便平台布置多分辨率版本用. 可以把不变参数写进 pass1+2, 变化的写进 txt. 格式为"[压制名:analysis-load-reuse-level:analysis-load] <参数 1+输出文件名>"

```
x265.exe --abr-ladder 1440p8000_2160p11000_2160p16000.txt --fps 59.94 --input-depth 8 --input-csp i420 --min-keyint 60 --keyint 60 --no-open-gop --cutree
```

```
1440p8kb_2160p11kb_2160p16kb.txt {
```

```
[1440p:8:Anld 存档1] --input 视频.yuv --input-res 2560x1440 --bitrate 8000 --ssim --psnr --csv 9.csv --csv-log-level 2 --output 1.hevc --scale-factor 2
```

```
[2160p1:0:nil] --input 视频.yuv --input-res 3840x2160 --bitrate 11000 --ssim --psnr --csv 10.csv --csv-log-level 2 --output 2.hevc --scale-factor 2
```

```
[2160p2:10:Anld 存档3] --input 视频.yuv --input-res 3840x2160 --bitrate 16000 --ssim --psnr --csv 11.csv --csv-log-level 2 --output 3.hevc --scale-factor 0 } analysis-load 填 nil(不是 nul)代表略过
```

## 近无损压缩, 真无损压缩上层模式

**--lossless**<开关>跳过分块, 动/帧/参搜索, 量/自适应量化等影响画质的步骤, 保留率失真优化以增强

参考性能. 直接输出体积非常大的原画, 相比锁定量化方法, 这样能满足影业/科研用, 而非个人和一般媒体所需, **真无损导出有几率因为参考质量提升, 会比近无损小**

**--tskip**<开关, 已关>不在 tu 上使用 DCT 变换  $\sim (\cdot \cdot \cdot -)$

**--cu-lossless**<开关, 已关>将“给 cu 使用无损量化 (qp 4)”作为率失真优化的结果选项之一, 只要码率管够 (符合  $\lambda = R/D$ ) 就不量化. 用更多码率换取原画相似度, 无损源能提高参考冗余

## 率失真优化-宏观码率分配调优

**--rdoq-level**<整数, 范围 0~2>率失真优化控制量化的分块深度. 0=关; 1=不分 tu; 2=4x4, 慢

**--psy-rdoq**<浮点 0~50, 默认 0 关>心理视觉优化偏移率失真优化的程度, 提高能量 J 以改变 rdoq 的用途, 使其更不愿消除系数, 避免模式决策遇到差选项. 类似 crqpoffs. 1080p 设<2.3~2.8>给中低成本动漫, <3~4.8>给电影-高成本动漫, <7~12+psy-rd 3, tskip, tskip-fast, ipratio 1.2, no-sao>留噪. 保留的细节面积越小设得越高

- 常用: psy-rdoq 和 psy-rd 功能冲突, 所以保留 rdoq-level 1, 关 psy-rdoq, 开 psy-rd
- 高码: 有颗粒的情况下同时用低强度的 psy-rdoq 和 psy-rd, rdoq-level 2
- 少用: 目前 x265 psy-rd 还没写 cpu 指令集(慢, 待跟进), 所以关 psy-rd, 开 psy-rdoq

**--nr-intra, --nr-inter**<整数 0~2000, 默认 0 关>基于运动补偿 mc 给变换后-量化前降噪变换结果的参数. 其中帧内 intra 不如 VapourSynth 里的高端降噪滤镜. 但帧间/时域上降噪的 nr-inter 和 x264 的 hqdn3d 类似, 可以拉近参考源和参考帧的差距/残差, 实现在 rc-grain 上进一步稳定 qp 计算, 且在噪点源中相比模糊掉纹理更容易破坏噪点, 结果类似双边滤镜. [1920x1080 下最高不建议超过 250](#)

## 自适应量化

根据源图像的复杂度来判断 qp 值分配的计算, 防止 x265 往细节分配太多码率而造成平面的质量亏损. 对防止图像变得模糊有一定作用 (〰~〰);

**--aq-mode**<范围 0~4, 0 关>aq 只在码率不足以还原原画时启动, 建议<1>标准(简单平面); <默认 2>+启用 variance 调整 aq-strength, 适合录像; <3>+欠码时码率多给暗场些; <4>+让不足以还原原画情况的码率多给边缘些(高锐多线条多暗场少平面)

**--hevc-aq**<开关, “aq-mode 5”>以 1/4-tile 而非边缘适应量化. 据测试 [1](#), [2](#) 比 aq 4 快且适合动漫, 而 aq 4 更适合录播(?) **目前学术方-官方-第三方信息较为割裂, 所以暂无适解**

**--aq-strength**<浮点>自适应量化强度. 据 VCB-s 建议, 动漫的值太高则浪费码率. 动漫中 aq-mode / strength 给<1 对 0.8>, <2 应 0.9>, <3 和 0.7>较为合理, 在真人录像上可以再增加 0.1~0.2, 画面越混乱就给的越高, 在 aq-mode 2 或更高下可以更保守的设置此参数

**--aq-motion**<开关, **实验性**>根据动态信息微调自适应量化的效果 mode 和强度 strength

**--qg-size**<64/32/16/8, 需  $\geq \text{min-cu-size}$ >最小支持自适应量化的 cu. 默认 64 可换取更多速度. 高画

质/平衡都建议设在 32~16. 用途不明的<最浅, 最深>格式能自定义范围, 如 32, 16 代表仅这两种 cu

**--cbqpoffs, --crqpoffs**<整数>调整蓝, 红色平面相比亮度平面的 qp 值差异, 负值降低量化. 若当前版本 x265 的算法把色度平面的量化变高, 可以用这两个参数补偿回来. 由于编码器一直不擅长处理红色, 而人眼又对红光敏感可能因为祖先晚上生火所以为了画质建议比 cb 面设更低( $\Delta -3$  左右)的值

## x265 jpsdr Mod:

**--aq-auto**<多开关, 默认 0 关>. **--aq-mode 5**<由 aq-auto 控制>4 的基础上启用以下参数的支持

值	逐帧 aq	延迟逐帧 aq (hysteresis)	HDR 兼容	aq-mode 5
1	✓			
2, 3	✓	✓		
4	✓		✓	
8	✓			✓
6	✓	✓	✓	
10	✓	✓		✓
12	✓		✓	✓
14	✓	✓	✓	✓

**--aq-fast-edge**<开关>aq-mode4, 5 下的边缘检测跳过高斯模糊过滤, 不适合脏片源

**--aq-bias-strength**<浮点, 默认 1.00>aq-mode3, 5 下 aq-strength 偏给暗场的程度

**--aq-strength-edge**<浮点 0~3.00, 默认=aq-strength>aq-mode4, 5 下针对纹理而非平面的 aq-strength. **--aq-bias-strength-edge**<浮点, 默认=aq-bias-strength>aq-mode5 下 aq-strength-edge 偏给暗场的程度

## 模式决策

**mode decision** 整合搜来的信息, 宏观上重新定制分块, 参考, 跳过, 编码, 量化方案. 因为选码率最小的压缩方案常常不平衡

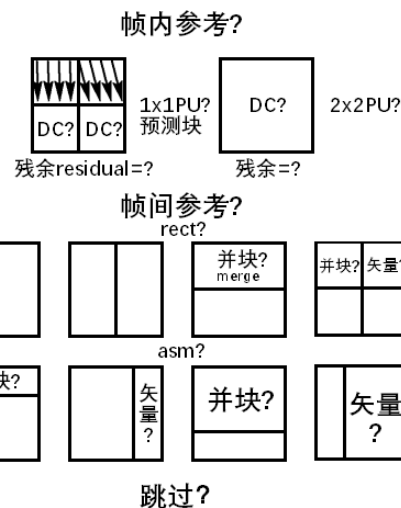
**--rd**<1/2/3/5, 默认 3>率失真优化参与 md 的程度, 越大越慢. <1>优化帧内参考, 并块/跳过决策 <2>+分块决策 <3>+帧间决策 <5>+向量/帧间方向预测决策>建议快速用 1, 2; 高压使用 3, 有时间用 5. 5 比 3 慢上一倍, 原画画质一般还难以看出提升

**--limit-modes**<开关>用附近的 4 个子 CU 以判断用 merge 还是 AMVP, 会大幅减少 rect/amp 的效果, 明显提速. 会增大或减少体积, 微降画质但难以察觉

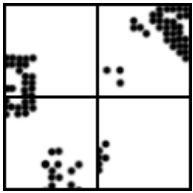
**--limit-refs**<整数 0~3, 默认 3>限制分块用信息可参考性. <0 不限>压缩高且慢; <1>用 cu 分裂后的信息+差异信息描述自身(推荐); <2>据单个 cb 的差异信息建立 pu; <3=1+2> ㄟ(ㄟ\_ㄟ)

**--rskip**<整数 0~2>前 cu 分块被跳过时, 判断后 cu 接着搜索分块还是提前退出的参数. 画面越接近全屏, 低成本动漫就用得越多

- <0>继续分析. 适合信噪比差, 噪声高的源. 原画很干净则不如<1>
- <1>rd0~4 下据临 cu 是否细分而定; rd5~6 下看附近 2Nx2N cu 分块难度而定, 高压和一般情况推荐
- <2>直接对比 cu 纹理密度 edge density, 快且不比前者差, 但存在对源的画质要求及客观判断"画质"能力的要求







--rskip-edge-threshold<0~100 默认 5: 趋向于分块, 需 rskip 大于 1>用 sobel 算法获取 cu 纹理密度, 除以块所占面积的百分比值. 纹理密度>阈值=分块. 量化强度越高越关键, 与 x264 deadzone 略像. 8×8 或 16×16 块下默认 5%即含 3 或 12 个系数就分

- <像素风>据像素变大的程度决定. 如画面除以 2 回到 1x1 像素大小, 则使默认值乘以 2 以提速
- <抗涂抹>rskip 2 rskip-edge-threshold 3, 即“有一点不平就分块”. 比 rskip 0 快, 用于已知要保留

雪景等全屏大量动态信息的源的情况下, 节省传统分块计算时间用. 可以在不添噪点的情况下达成其抗涂抹的目的

--tskip-fast<开关, 已关>跳过 4x4 tu 的变换, 忽略部分系数 coefficients 来加速, CbCr-tu 也取决于 Y 块是否被跳过. 在全屏小细节的视频中有显著加速效果. 建议除高压以外的任何环境使用

## 率失真优化 RDO 控制

**率失真优化** rate distortion optimization 据多个码率下测得的失真程度点, 挑出低于率失真曲线的值, 再根据码率/crf 做**模式决策**. x264/5 用拉格朗日代价函数  $J = D + \lambda \cdot R$ . 即“开销 = 失真+ $\lambda$ ·码率”. x265 中失真 distortion 用 MSE 判断

$$(x265) \text{ Mean Squared Error} = 1 \div \text{宽高} \sum_{x=0 \rightarrow T_x} \sum_{y=0 \rightarrow T_y} |f(x,y) - f'(x,y)|^2$$

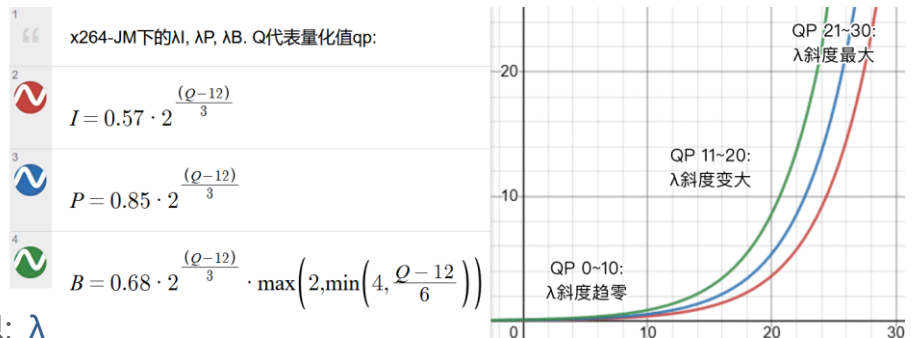
$$(x264) \text{ Sum of Squared Err} = \sum_{x=0 \rightarrow T_x} \sum_{y=0 \rightarrow T_y} |f(x,y) - f'(x,y)|^2$$

$$(x264 \text{ fgo}) \text{ NSSE} = \sum_{x=0 \rightarrow T_x} \sum_{y=0 \rightarrow T_y} |[N(x,y) - N'(x,y)] \cdot \text{fgo}| + |f(x,y) - f'(x,y)|^2$$

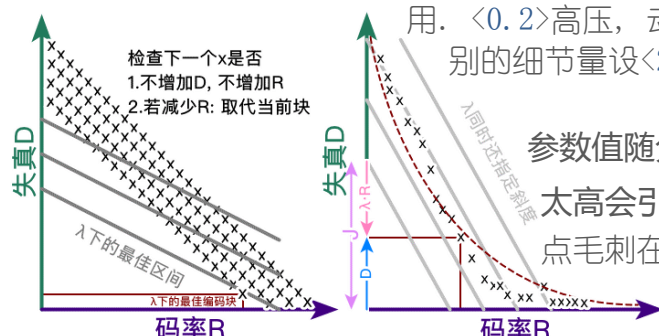
- $\sum_{x=0 \rightarrow T_x}$  代表块宽度求和范围,  $f()$  和  $f'()$  分别代表参考块和参考源
- $\sum_{y=0 \rightarrow T_y}$  代表块高度求和范围,  $x, y$  代表块中的像素坐标,  $||$  求绝对值, 否则求和时正像素值差异会减去负

**拉格朗日值 $\lambda$** 从 qp 值得出, 即 crf, abr 指定的率失真斜率区间. qp 越大斜度越小.  $\lambda=0$  则代价=失真, 给多少码画面都一样. 所以压缩开最大.  $\lambda$ 趋 0 则代价趋失真, 即压缩一截下去不会影响多少画质, 稍微给点码意思意思;  $\lambda$

远大于 0 则代价>失真, 提升画质收益大于压缩率降低, 应该减少压缩保画质



--psy-rd<浮点 0~50 默认 2, 需 rd3, 默认 0 关, 和 x264 不同>心理视觉优化影响率失真优化的程度, 增加量化块的能量, 抗拒帧内搜索, 使模式决策 mode decision 遇不到差选项. 注意搭配 psy-rdoq 使用. <0.2>高压, 动漫据纹理设<0.5~2>. 录像设<1.5~2.5>, 星空与 4k+级别的细节量设<2.8>或更高



参数值随分辨率大小变化. 注意噪声和细节都是高频信息, 所以开太高会引入画面问题.图: 复杂度对真人录像的重要性, 但这些点点毛刺在低成本/大平面动漫里就很难看了

--rd-refine<开关, 建议开, 需 rd 5, 有瑕疵 bug>率失真优化分析完成帧内搜索 cu 的最佳量化和分块结果, 耗时换压缩率和画质. x264 中直接嵌入 subme 8 中, 还多一个最优动态向量分析. 有斜面上随机出现块失真的 bug, 若图中失真可接受则可开:



--dynamic-rd<整数, 范围 0~4>给 VBV 限码的画面调高率失真优化以止损. 1~4 对应 VBV 限码的画面的 rd 搜索面积倍数, 越大越慢

--splitrd-skip<开关, 已关>启用以在“所有当前 CU 分割致失真程度之总和”大于“任意同帧 CU 分割致失真程度之总和”时, 不跟随当前 CU 分割之结果来独立计算 rd 值以加速

--qp-adaptation-range<浮点 1~6, 默认 1>心理视觉优化调 qp 的最大范围, 信息不足且没有负值: 高则提高而不是降低 qp?

## 峰值信噪比 peak signal-to-noise ratio/PSNR

信号电压声音强度都用分贝表示, 因为分贝通过信号值越大/越小, 电平增长越弱的对数线  $y=\log(x)$  显示. 功率转 dB 的算式即

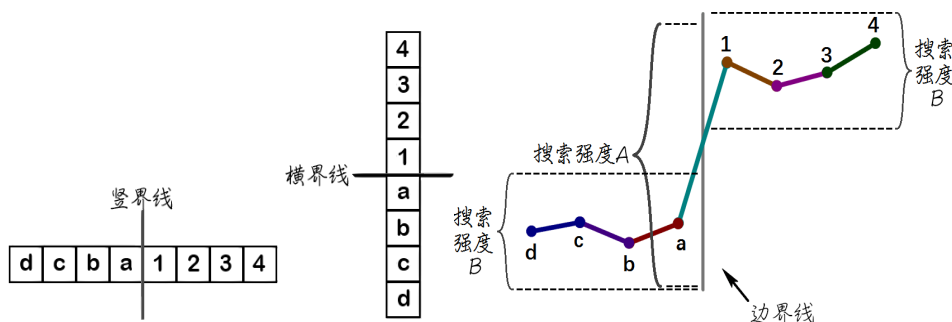
$PSNR=10\log_{10}(\text{最大}\div\text{当前})$ . 最大 $\div$ 当前即最大像素值 $\div$ MSE; 8bit 下  $255\div$ MSE; 10bit 下则  $1023\div$ MSE



## 环路滤波-去块滤镜

修复高量化  $qp>26$  时宏块间出现明显横纵割痕瑕疵的平滑滤镜. 编码器内去块相比于外部滤镜能得知压缩待遇信息(两个相邻块的量化, 参考待遇差异是否过大)从而避免误判原画纹路. 码率跟不上就一定会出现块失真, 所以除直播关掉以加速外, 任何时候都应该用; 但去块手段目前仍是平滑滤镜, 因此要降低强度才适用于高码视频, 动漫, 素材录屏等锐利画面

边界强度 boundary strength(去块力度判断): 取最小 8x8 块间的界线举例. (不是 4x4)



- 平滑 4: a 与 1 皆为帧内块, 且边界位于 CTU/宏块间, 最强滤镜值
- 平滑 3: a 或 1 皆为帧内块, 但边界不在 CTU/宏块间
- 平滑 2: a 与 1 皆非帧内块, 含一参考源/已编码系子
- 平滑 1: a 与 1 皆非帧内块, 皆无参考源/已编码系子, 溯异帧或动态向量相异
- 平滑 0: a 与 1 皆非帧内块, 皆无参考源/已编码系子, 溯同帧或动态向量相同, 滤镜关

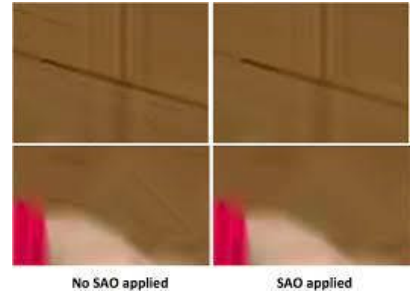
--deblock<平滑强度:搜索精度, 默认 1:0, 推荐 0:0, -1:-1, -2:-1>两值于原有强度上增减

- 平滑 $\geq 1$ 时用以压缩,  $<0\sim 1$ 时略微降低锐度, 适合串流

- 平滑<-2~-1>适合锐利视频源, 4k 电影, 游戏录屏. 提高码率且会出现块失真
- 平滑<-3~-2>适合高码, 高锐动画源和高画质的桌面录屏. 高码率, 增块失真, 但高码动漫观感还是比 1 好
- 搜索<大于 2>易误判, <小于-1>会遗漏, 建议保持<0~-1>, 除非 qp>26 时设<1>

## 环路滤波-取样迁就偏移

sample adaptive offset 滤镜. 逐 CTB 划分. **界偏移 eo** 缓解纹理边缘因"高频波形的遮盖因强量化或去块丢失"的问题. 可能分以下几步. 算法略像帧内搜索的趋平滤镜, 但是是强制的. 适合修复纹理量化出振铃的损失. 关闭则遮不住平面的蠕虫失真



- CTB 内据分块结果建立一批像素大小的 3x3 采样域
- 由 3x3 排列规律找出三个像素排成一排的横 0, 纵 1, 左倾 2 和右倾 3 四种可能
- 中心像素的值同时小于旁像素, 同时大于旁像素, 等于任意一边像素的条件, 否则视作非边界
- 确保 3x3 间的边界连续性, 实现边界验证和性能优化(?)

**带偏移 bo** 是一种对比源+补偿编码差异的平面 CTB 滤镜. 方法是 1.划分 32 条色深带来均分当前色深下的像素值, 2.分为 24~31, 0~7 的明带, 暗带; 以及 8~23 的中间带. 限制最大只能补偿 4 条相连的色深带, 以确保 CTB 中色深差异不会大到触发 eo, 同时涵盖足够大的斜面渐变, 然后率失真优化找出所谓的补偿值: 每条色深带的偏移值. 因此适合补偿平面, 斜面和曲面渐变 CU, 适合修复平面量化的损失

- 共有 0=无, 1=横 E0, 2=纵 E0, 3=左倾 E0, 4=右倾 E0, 5=中间带, 6=明暗带, 共 7 种补偿方案
- 共有 0=无, 1=Y, 2=Y+Cr, 3=Y+Cb, 4=YCbCr, 5=Cr, 6=Cb, 7=CbCr 种平面补偿开关

**参数融合 merge** 将相邻两个 CTB 的 sao 信息(补偿方案, 平面补偿开关等)根据参数决定直接用上/左块, 还是对比像素趋势更接近哪个. 和选择 bo, eo 具体的偏移值一样由率失真优化决定

**--no-sao**<关闭--sao, 默认开 sao>由于针对的是强量化环境, 所以高画质源+crf<17 的情况下可以关

**--sao-non-deblock**<开关>启用后, 未经由 deblock 分析的内容会被 sao 分析🔴🔴

**--no-sao-non-deblock**<默认>sao 分析跳过视频右边和下边边界( /)u(\ )

**--limit-sao**<开关, 默认关>对一些计算采用提前退出策略, 不是改善画质的, 但 crf≈18, cutree 和 bframes 16 下可以开, 以保留一定影响

**--selective-sao**<0~4, 默认 0>从条带(横向一组 CTU)角度调整 sao 参数, 1 启用 I 条带 sao, 2 增加 P 条带, 3 增加 B 条带, 4 所有条带. 可看作新的 sao 控制方式, 或搭配 limit-sao 的新方法

## 熵编码/残差编码/文本压缩-CABAC

游程编码将降维后的块/条带丢给熵编码(率失真优化要解码检查每道压缩, 所以要经历多次熵编码).

x264/5 中用 context adapt. binary arithmetic coding. 相比于 cavlc 和霍夫曼编码, cabac 的压缩率更好.**二进制算数编码** bac 在 0~1 的模具中, 用**匹配符(阵列)**, **出字概率(次数)**及**待压内容**压缩. 例如

$x=\{y, u, \text{式}\}$  三种符; 出字 **yuyu** 式中的概率  $P=.25, .25, .5$  故照下图规律, 经 8 层细分后给出二进制格子里符合条件的小数压缩:





**--idr-recovery-sei**<开关>sei 里写进 idr 帧，串流时防止整个 gop 都找不到参考帧而崩坏

**--frame-dup**<开关默认关，必须开 vbv 和 hrd，有 bug>将 2~3 面近似的连续帧换成同一帧

**--dup-threshold**<整数 1~99，默认 70>相似度判定值，默认达 70%重复就判为相似

## 线程节点控制

参自 [pugetsystems.com](https://pugetsystems.com): Cinebench, 虚幻 5 上 5800X 比 5800X3D 快近 10%反映了持续计算下 5800X 频率更高的优势; Lightrooms 等单图处理上 5800X3D 跑过了 5800X, Photoshop 上 5900X, X3D, X 同时打平(优化?); 剪辑视频并非持续高负载, 因此 AfterFX, Premiere, DaVinci 上 X3D 和 X 打成平手. 所以选择时首先看生产力软件的优化, 其次看任务类型. 3D 缓存处理器比同型号更慢的原因是目前视频编码对内存带宽需求不大, 更高频率才是优势. 因此压制的视频越长性价比越低. 但中间若有一堆内存读写间歇, 3D 缓存处理器就能用低延迟读写赢回速度

**--pools**<整数/加减符,,, 默认+, +, +, +>x264 中--threads 的升级版. 如--pools +, -, -, -表明 pc 有 4 个节点, 仅占用第一个. +代表全部处理器线程. 这样能防止多处理器系统上跑一个 x265 时, 所有处理器访问第一个节点的内存而造成延迟等待. 应该是跑和节点一样多的 x265, 每个节点各自运行. 单 cpu 系统直接作--threads 用, 如--pools 8 指该 pc 有 1 个节点, 占用该节点上处理器的 8 个线程

不要企图设置大于实际线程数的 pools/threads 提速. 会因为处理器随机并发的特性从任务数量上冲淡参考帧建立等要之前的步骤算完才能开始的时间窗口. 否则编码器只能跳过参考压缩, 造成处理器占用降低, 码率增加以及压制变慢的副作用

TR1000~2000 系处理器是用多个节点拼出来的, 所以单处理器的内部要按多个节点分开算, 特例是 2990WX, 2970WX, 核心组 1 和 3 没有内存控制器, 0 和 2 有内存控制器, 所以 1, 3 不能用

**--pmode**<开关>使用平行帧内搜索, 目前出现了难以应付噪点, 会造成画质下降, 码率提高的问题

**--asm**<avx512>avx512 was a mistake - Intel engineer

## 多线程 vs 多参考

用多线程一次编码多帧来占满算力, 还是一次只编一帧, 确保所有参考画面可用的决策. 确保所有帧同时吞吐 $O(\cdot \times \cdot)$ . 虽然 x265 有 tile 这种集合多个分片的并行化. 造成多线程多参考帧困难的原因有:

1. ctu 比宏块大, 相似性降低了
2. 参考前要等环路滤波和率失真优化, 还有已编码信息的依赖, 使得很多参考, 特别是高 ref 设定下来不及找就跳过
3. 参考帧的波前编码 wavefront parallel process (压制/播放的多线程改进版)因一行参考 ctu 的存在而卡死, 重启波前编码等没了多余算力

**--pme**<开关, 默认关. 任意年代>=16 核处理器>使用平行动态搜索 parallel ME. 解决了单核性能不足导致 me 后的编码步骤等待, 占用降低速度变慢的问题. 让旧服务器超多核中低频处理器变得更有编码性价比. 而达到 16 核的处理器也能提速

--frame-threads<整数 0~16~线程数/2, 默认 0 自动>同时压多少帧, 设 1 能让前后整帧可参考, 非 1 就只给 ctu 下方的一行 ctu. 设 1 的代价是 cpu 占用显著降低, 压制减速(-, -)

--lookahead-threads<整数 0~16~线程数÷2, 默认 0(关闭)>分出多少线程专门找参考, 而非与帧编码一同占线程, 可能只有开 frame-threads 1 时手动启用以增加 cpu 占用, pme 和 pmode 同理

## 色彩空间转换, VUI/HDR 信息, 黑边跳过

纯元数据, 写错或忘写也可以改. hdr 应用早期因制定方猜不出元数据主次, 所以制定主 master-display, 次 maxcll. 但现在...hdr 电视只读取"次要"的 maxcll 和 maxfall, 并且忽略"主要"的 master-display. 另外光强/光压的单位是 candela (1cd=1nit)

由于 bt601, bt709, HDR-PQ, HLG 标准所重用的亮度范围不一(偏亮或偏暗), 所以在编码, 心理学优化算法, 编码器参数上其实都要调整适配, 所以出现了适配不当的可能

--master-display<G(x, y)B(, )R(, )WP(, )L(, )>写进 SEI 信息里, 告诉解码端色彩空间/色域信息用, 搞得这么麻烦是因为 hdr 作为新标准不敢确定播放硬件需要什么信息, 所以就把 master-display 写成必须参数了. 绿蓝红 GBR 和白点 WP 指马蹄形色域的三角+白点 4 个位置的值×50000. 光强 L 单位是 candela×10000

SDR 视频的 L 是 1000, 1. 压 HDR 视频前一定要看视频信息再设 L, 见下

- DCI-P3 电影业内: G(13250, 34500)B(7500, 3000)R(34000, 16000)WP(15635, 16450)L(?, 1)
- bt709: G(15000, 30000)B(7500, 3000)R(32000, 16500)WP(15635, 16450)L(?, 1)
- bt2020 超清: G(8500, 39850)B(6550, 2300)R(35400, 14600)WP(15635, 16450)L(?, 1)

RGB 原信息(对照小数格式的视频信息, 然后选择上面对应的参数):

- DCI-P3: G(x0.265, y0.690), B(x0.150, y0.060), R(x0.680, y0.320), WP(x0.3127, y0.329)
- bt709: G(x0.30, y0.60), B(x0.150, y0.060), R(x0.640, y0.330), WP(x0.3127, y0.329)
- bt2020: G(x0.170, y0.797), B(x0.131, y0.046), R(x0.708, y0.292), WP(x0.3127, y0.329)

--max-cll<最大内容光强, 最大平均光强>压 HDR 一定照源视频信息设, 找不到不要用, 例子见图:

```
Bit depth : 10 bits
Bits/(Pixel*Frame) : 0.120
Stream size [速率(像素×帧率)]: 0.251 GiB (84%)
Default : Yes
Forced : No
Color range : ATOM Titan File 3.8.3 (4.8.3.0)
Color primaries : BT.2020
Transfer characteristics : PQ
Matrix coefficients : BT.2020 non-constant
Mastering display color primaries: R: x=0.680000 y=0.320000,
G: x=0.265000 y=0.690000, B: x=0.150000 y=0.060000, White point: x=0.312700 y=0.329000
Maximum Content Light Level: 1000 cd/m2
Maximum Frame-Average Light Level: 640 cd/m2
Control display color: Display P3
Control display height: min: 0.0050 cd/m2, max: 4000 cd/m2
最大内容亮度等级: 1655 cd/m2
最大帧平均亮度等级: 117 cd/m2
```

图: max-cll  
1655, 117/L(40000000, 50)/colorprim bt2020/colormatrix bt2020nc/transfer smpte2084

--hdr10<自动开关>当 master-display, max-cll 启用就在 sei 中指示 hdr10 相关参数, 原本参数名叫 --hdr(和 hdr-opt 一样), 改名是为了指明它能优化新的 hdr10, 而非旧的 hdr

**--hdr10-opt**<开关, 已关>逐块为 10bit bt2020, smpte2084 视频做亮度色度优化, 其它视频无效

**--display-window**<←, ↑, →, ↓>指定黑边宽度以跳过加速编码, 或者用 **--overscan crop** 直接裁掉

**--colorprim**<字符>播放用基色, 指定给和播放器默认所不同的源, 查看视频信息可知: bt470m, bt470bg, smpte170m, smpte240m, film, bt2020, smpte428, smpte431, smpte432. 如图→为 bt. 2020

**--colormatrix**<字符>播放用矩阵格式/系数: fcc, bt470bg, smpte170m, smpte240m, GBR, YCbCr, bt2020nc, bt2020c, smpte2085, chroma-derived-nc, chroma-derived-c, ICtCp, 不支持图↑的 bt2020nc

**--transfer**<字符>传输特质: bt470m, bt470bg, smpte170m, smpte240m, linear, log100, log316, iec61966-2-4, bt1361e, iec61966-2-1, bt2020-10, bt2020-12, smpte2084, smpte428, arib-std-b67, 上图 PQ 即 st. 2084 的标准, 所以参数值为 smpte2084

ffprobe 会将三个信息并一行写: Stream #0:0(und): Video: prores (XQ) (ap4x / 0x78347061), yuv444p12le (tv, bt2020nc/bt2020/smpte2084, progressive)

## IO(input-output, 输入输出)

**--seek**<整数, 默认 0>从第 x 帧开始压缩**--frames**<整数, 默认全部>一共压缩 x 帧

**--output**<字符串, 两边带双引号>例: **--output** "输出文件地址+文件名" "输入文件地址+文件名"

**--input-csp**<i400/i422/i444/nv12/nv16>在输入非默认 i420 视频时需要的参数, rgb 需转换

**--dither**<开关>使用抖动功能以高质量的降低色深(比如 10bit 片源降 8bit), 避免出现斑点和方块

**--allow-non-conformance**<开关>不写入 profile 和 level, 绕过 h. 265 标准的规定, 只要不是按照 h. 265 规定写的命令行参数值就必须使用这个参数 ⚡(↪↪)"⚡"

**--force-flush**<整数 0~2, 默认 0>性能不足录屏用. 当编码不过来就跳过几帧的措施(?):

- <0>等全部帧输入再编码
- <1>不等全部帧输入完就编码
- <2>取决于条带种类, 调整 slicetype 才能用

**--field**<开关>输入分行扫描视频时用, 自动获取分场视频的帧率+优先场, 替代了 **--interlaced** 参数

**--input-res**<宽 × 高>在使用 x265 时必须指定源视频的分辨率, 例如 1920x1080

**--fps**<整数/浮点/分数>在使用 x265 时必须指定源视频的帧率, 小数帧填小数, 勿四舍五入

**--chunk-start, --chunk-end**<开关, no-open-gop>chunk-start 允许跨 GOP 制作数据包(?), 改由 chunk-end 参数将数据包结尾和剩下的视频帧断开(?). 据描述看, 由于数据包接收顺序一定会被打乱, 所以只可参考其之前, 而不可参考之后的内容, 跟 http 的数据包编码协议有关 Σ(↪↪)"Σ"

**--temporal-layers**<开关, 默认关>使 x265 更兼容 svc 标准, 将非参考 b 帧(相当于空信息)分离到另

一层视频流中，解码器可以选择跳过而降低性能损耗，可能会造成兼容性问题

## 下载 附录与操作技巧

[LigH](#)

[jpsdr](#)

[Rigaya](#)

[Patman](#)

[ShortKatz](#)

[DJATOM-aMod](#)

[MeteorRain-yuuki](#)

[.hevc GCC10](#) [单文件 8-10-12bit] 附 x86, Windows XP x86 版 [附 libx265.dll](#)

[.hevc GCC12.2+MSVC\\_Illvm 1928](#) [8-10-12bit] 附 Broadwell 版 [支持 aq-mode 5](#)

[.hevc GCC 9.3](#) [8-10-12bit] 附 x86 版

[.hevc GCC 11+MSVC1925](#) [8-10-12bit]

arm64~64e 加 x86 版 [?] 需 macOS 运行编译命令文件 ?

opt-Intel 架构与 zen1~2 优化 [10bit], opt-znver3 代表 zen3 优化 [10-12bit] [GCC 10.2.1+GCC10.3](#)

[Ismash.mkv/mp4 或 .hevc](#) [能封装, 但传说 lavf 不如 pipe 可靠] [GCC 9.3+ICC 1900+MSVC 1916](#) [8][10][12bit]+[8-10-12bit]

[ffmpeg](#) 多系统兼容, 备用地址 [ottverse.com/ffmpeg-builds](#)

[mpv 播放器](#) 比 Potplayer 好在没有音频滤镜, 不用手动关; 没有颜色偏差, 文件体积小

[x265GuiEx \(Rigaya\)](#) 日本語, auto-setup 安装, [教程点此](#)

[Voukoder; V-Connector](#) 免费 Premiere/Vegas/AE 插件, 直接用 ffmpeg 内置编码器, 不用帧服务器/导无损再压/找破解. 只要下两个压缩包, 放 Plug-Ins\Common 文件夹就行了



### gcc 是什么, 为什么同版同参的编码器速度不同

把源码编成程序的软件即编译器. x265 有 mingw(gcc 套件), 套件版本新旧影响编出程序的效率, msvc 体积更小, 但需要 VCRUNTIME140\_1.dll; icc 需要 libmmd.dll; Clang 需要...?

速度不一样还可能源自内建函数. 函数即等待变量输入的算式. 由于 8bit x265 中有大量开发组手动编写的内建函数, 所以不同编译器给出的程序速度也不等. 而 10bit x265 完全没有手动编写的内建函数, 所以编译器只有优化源码. 同样, 速度测试应以 10bit x265 为基准(¬\_¬")

### rc 指 release candidate

有的 x265 编译的文件名上有 rc, 指已修复所有被提出的问题 且编译器认为 ok 的版本、(・ω・ヾ)

### 杜比视界 dolby vision 不深入研究

有两种 dv 格式, 单视频流和双视频流, 双视频流有视频层和 db 强化层, 强化层可被一般的 hevc 解码



器丢弃, 单视频流就只有私有解码器能播放. 作者认为 dolby vision 还不如光线追踪. 内容和设备生产方要各交各钱合作, 所以用户只能交大钱才能达到标准体验. 而多数用户因剧情而非试听购买设备, 所以更方便的平板电脑, 笔记本电脑还是追剧主力. 另外, 在本教程写出的几年时间里该技术如上述所料, 已经变得没人讨论了

## CMD 操作技巧 color 08

将原本黑景白字改成黑景灰字的单行命令, 降低视疲劳

## cmd 窗口操作技巧%~dp0

"%~"是填充字的命令(不能直接用于 CMD). d/p/0 分别表示 drive 盘/path 路径/当前的第 n 号文件/盘符/路径, 数字范围是 0~9 所以即使输入 "%~dp01.mp4" 也会被理解为命令 dp0 和 1.mp4

这个填充展开后可能是"C:\"+ "...\"+ 1.mp4, 路径取决于当前.bat 所处的位置, 这样只要.bat 和视频在同一目录下就可以省去写路径的功夫了. 若懒得改文件名参数, 可以用%~dpn0, 然后直接重命名这个.bat, n 会将输出的视频, 例子: 文件名=S.bat → 命令=--output %~dpn01.mp4 → 结果=1.mp4 转输出"S.mp4" (/·ω·)/

## cmd for 循环批量压制(确保文件名无重复, 预先分离出音频, 预先将视频套滤镜渲染好)

给出 bat 文件所在目录下完整 pdf 路径+文件名: for %a in (\*.pdf) do echo '%~dp0%a'

批量压 mkv: chcp 65001

@ for %1 in (\*.mkv) do (x265 [参数] --output 'D:\文件夹\%%~n1.mp4' '%~dp0%1' & qaac [参数] -o 'D:\文件夹\%%~n1.aac' '%~dp0%~n1.flac')

## ffmpeg 批量压 mp4, 音频拷到新文件: chcp 65001

@ for %3 in (\*.mp4) do (ffmpeg -i '%3' -c:v copy -i '%~n3.aac' -c:a copy '%~n3.mp4')

chcp 65001 会让 cmd 以 unicode 形式读取, @是不打出输了什么命令进去, %%~n1 是%%1 去掉了文件后缀 o(-\_^)

## LSMASHWorks 崩溃 0xc0000005 可能是内存问题

## Worm effect 瑕疵 原因未知, x265 低码+no-sao 可复现的噪点横向拉伸效果

--preset	superfast	veryfast	faster	fast	medium	slow	slower	Very slow	placebo
ctu	32	64							
最小 cu	8								
连续 B 帧	3	4					8		

