LiveVideo StackCon 音视频技术大会

技术开启新“视”界
Technology Bring New Vision

手淘H265高效编解码器的实现

篱悠

LiveVideoStack 音视频技术社区

CSDN

1 项目背景

2 H265介绍

3 H265高效编解码器的实现

4 总结和未来展望

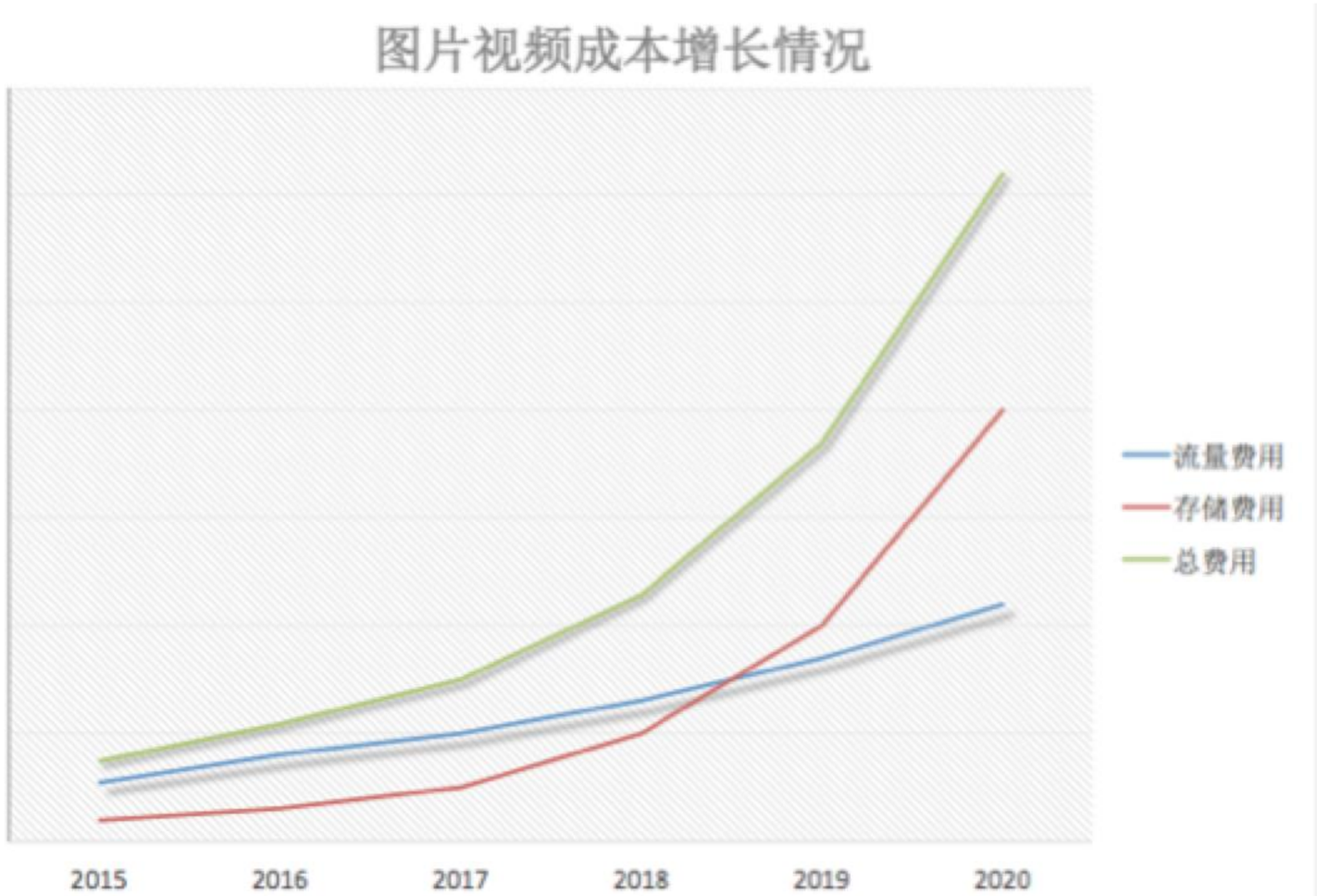
项目背景

手淘图片、直播、短视频、优酷等业务中存在着大量的图片和视频。

- 图片和视频数量不断增长
  指数级

- 图片和视频的清晰度需求不断提升
  360p->480p->720p-> ……
  ->1080p->4k/8k

消耗大量带宽成本和存储成本



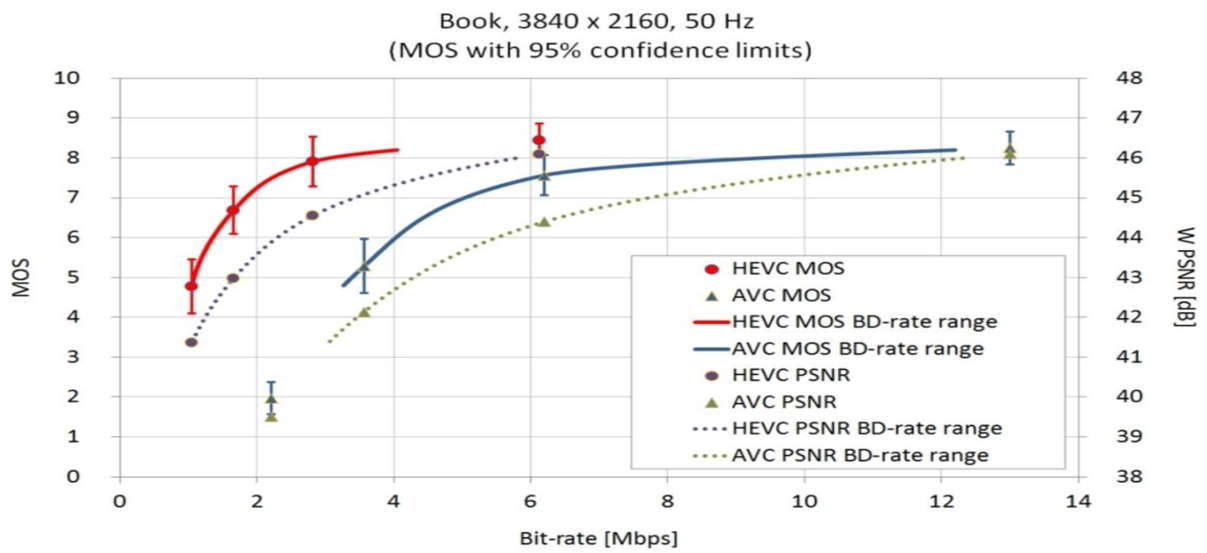
大规模并发推流场景的成本效益

编码解决方案：H265

1. 4K典型测试序列，相同画质下，
   h265比h264可节省50%带宽
2. 针对现场720p直播，可节省30%流量

问题：

1. 端上计算能力受限、功耗要求高
2. 缺少快速、高效的端上编码方案



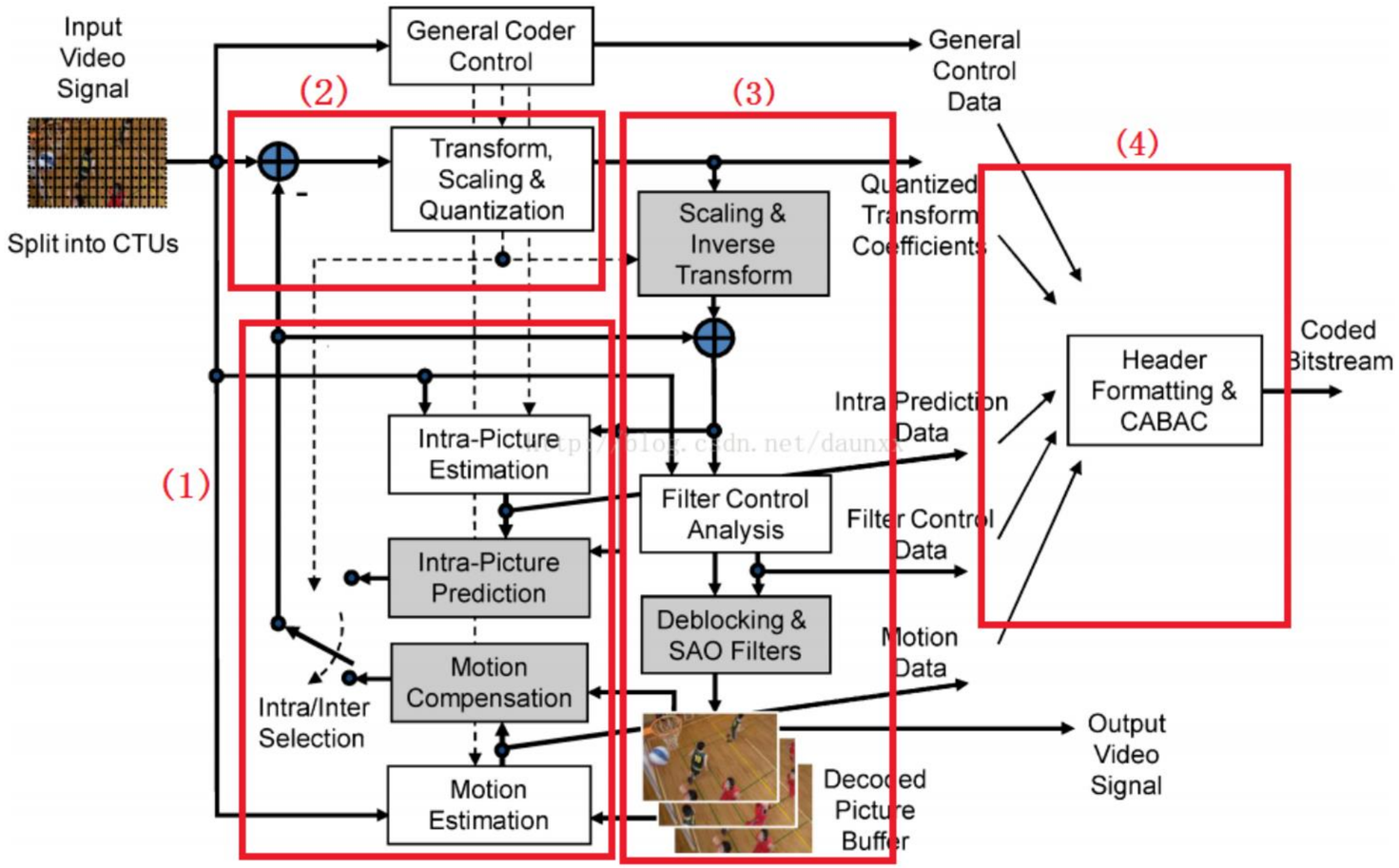
H265编码框架

1. 预测模块
 - 包含帧内预测，帧间预测

2. 变换量化模块
 - 针对原始图像块与预测图像块的差值做DCT和Quant

3. 解码模块
 - 解码后的图像用于下一帧的预测

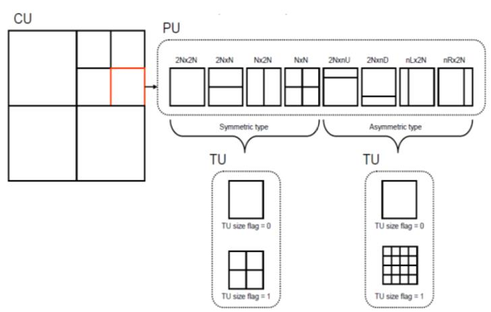
4. 熵编码模块
 - 针对预测信息和残差系数做算数编码，进一步消除编码冗余

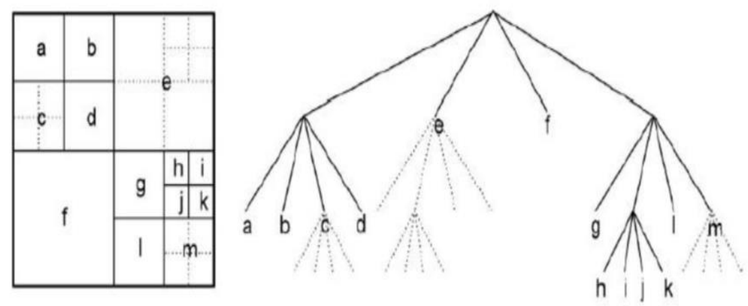


H265技术亮点介绍

1. 灵活的编码结构

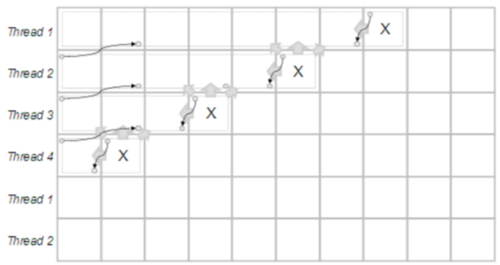
2. 灵活的块大小
 - 针对原始图像块与预测图像块的差值做DCT和Quant

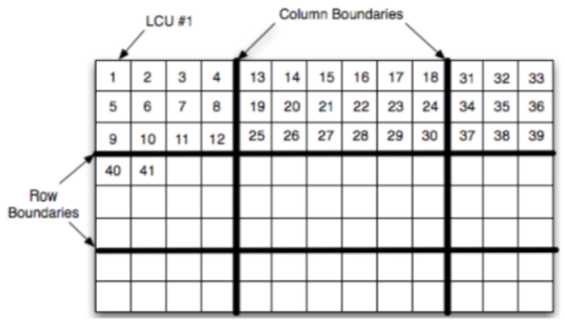




3. SAO
 - 解码后的图像用于下一帧的预测

4. 并行化设计
 - 针对预测信息和残差系数做算数编码，进一步消除编码冗余





H265技术亮点介绍（续）

5. 相对h264的技术改进

在插值，mv预测，帧内预测，变换，去块滤波等方面做了大量改进

	H.264	H.265
MB/CU 大小	4×4 ~ 16×16	4×4 ~ 64×64
亮度插值	Luma-1/2 像素{1,-5,20,20,-5,1} Luma-1/4 像素{1,1}	Luma-1/2 像素{-1,4,-11,40,40,-11,4,-1} Luma-1/4 像素{-1,4,-10,57,19,-7,3,-1} Luma-1/4 像素{-1,3,-7,19,57,-10,4,-1}
MVP 预测方法	空域 MVP 预测	空域+时域 MVP 预测 AMVP\Merge
亮度 Intra 预测	4×4 / 8×8 / 16×16：9/9/4 模式	34 种角度预测 + Planar 预测 DC 预测
色度 Intra 预测	DC, Horizontal, Vertical, Plane	DM, LM, planar, Vertical, Horizontal, DC, diagonal
变换	DCT4×4/8×8	DCT4×4/8×8/16×16/32×32 DST4x4
去块滤波器	4x4 和 8x8 边界 Deblock 滤波	较大的 CU 尺寸，4x4 的边界不进行滤波

表 2　H.264 和 H.265 关键特性对比

H265 VS H264编码效率

HM模型相比JM模型码率节省
— RA模式 save 39%
— LowDelay模式 save 44%
— AI模式 save 25%
平均节省36%

计算复杂度
— 编码增加300%-400%
— 解码增加50%

Video Sequence	HM 4.0 HE Bit Rate Savings for Equal PSNR		
	Random Access	Low Delay	All Intra
Class A	43%		29%
Class B	44%	48%	26%
Class C	34%	41%	23%
Class D	32%	38%	18%
Class E		51%	29%
Average	**39%**	**44%**	**25%**

MSU HEVC Comparison 2017

Fast mode

RDCurve

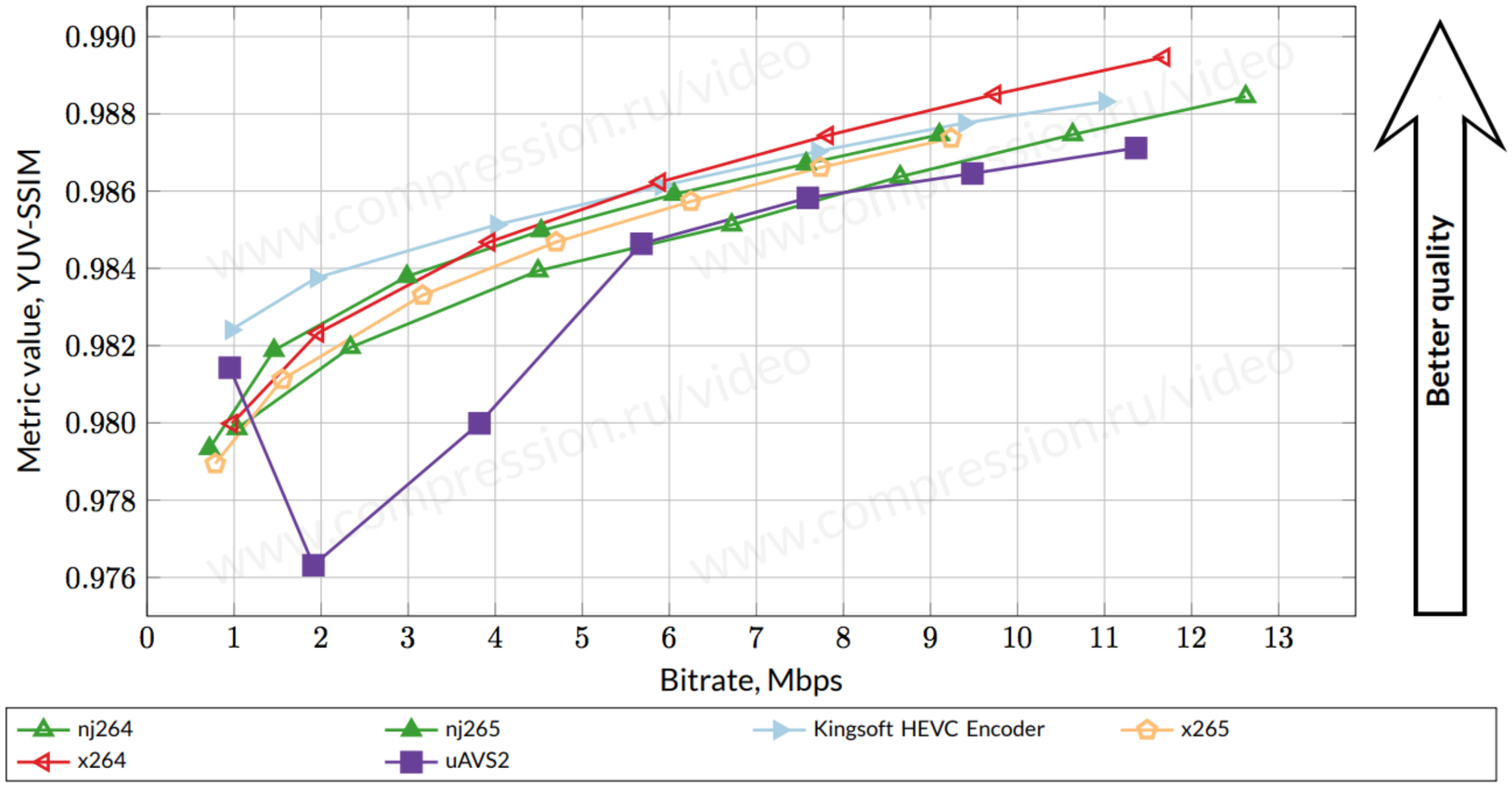


FIGURE 1: Bitrate/quality—use case "Fast Encoding," Color Tune sequence, YUV-SSIM metric

MSU HEVC Comparison 2017

Fast Mode

Speed

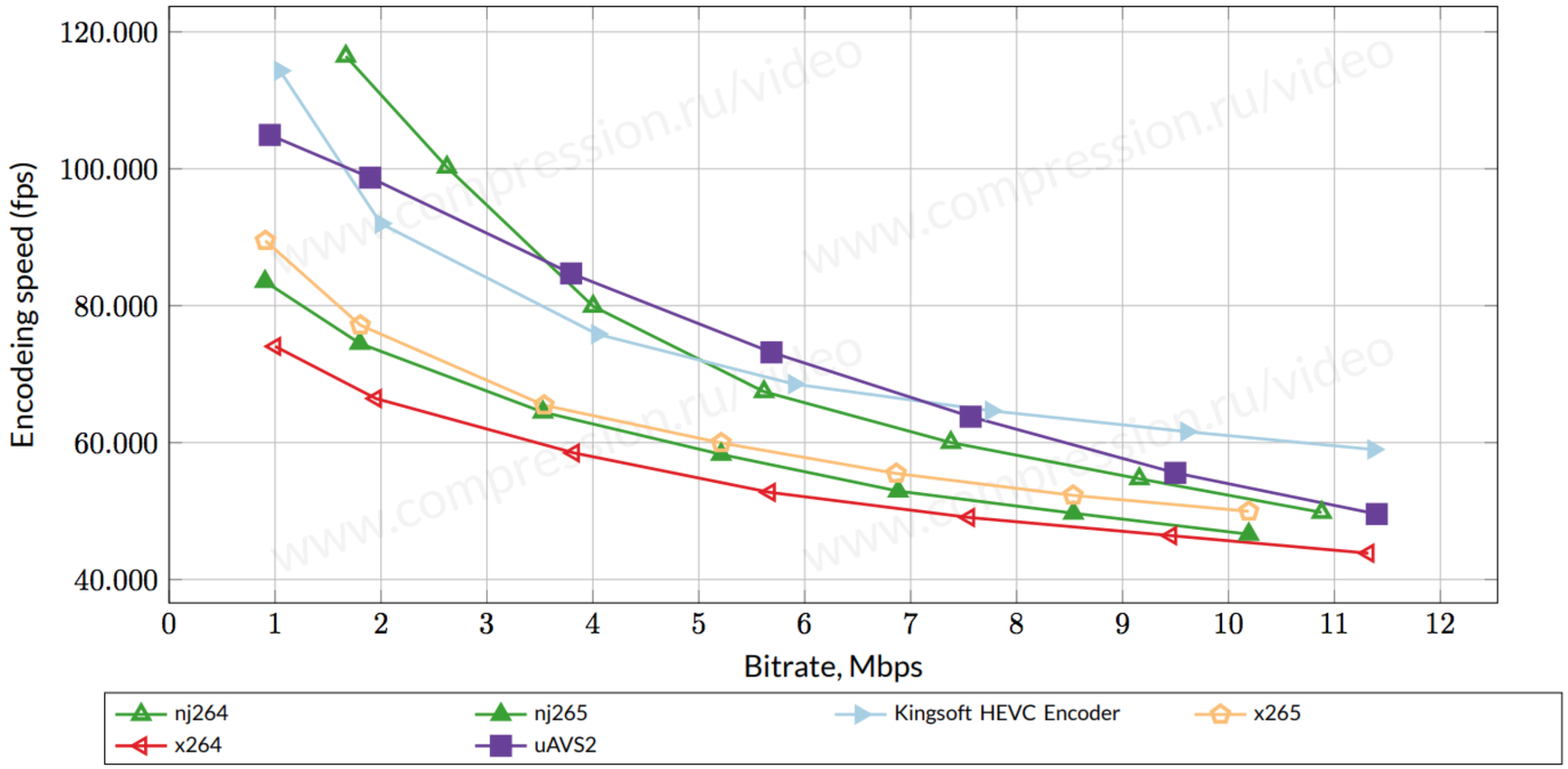


FIGURE 4: Encoding speed—use case "Fast Encoding," Steadicam sequence

高效实现之———RDO 优化

最优编码模式$p0$的搜寻过程可表示为　$p0 = \min(Dp + \lambda R(p))$
其中，p为CU/PU/TU模式的组合方式，$D\ p$ 和$R(p)$ 为此模式下的失真和码率；

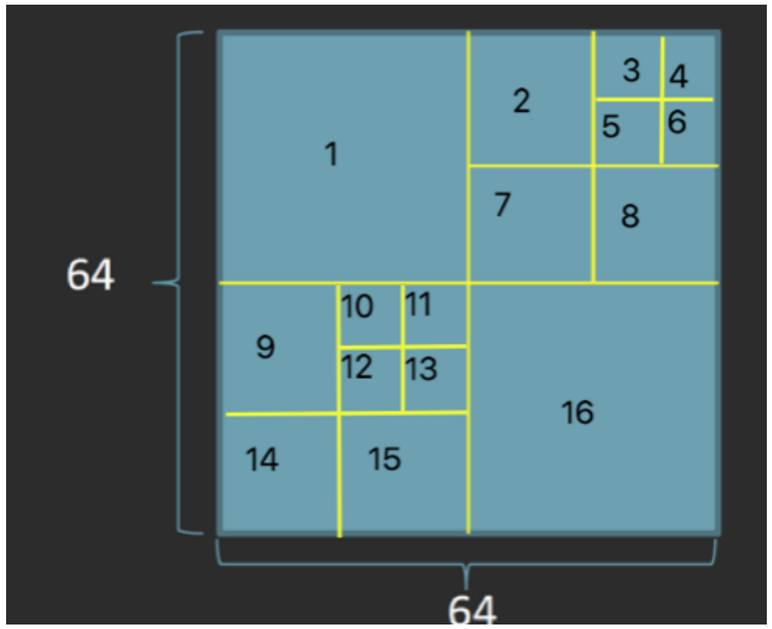
随着编码模式数量增加，率失真计算成为编码过程的计算瓶颈

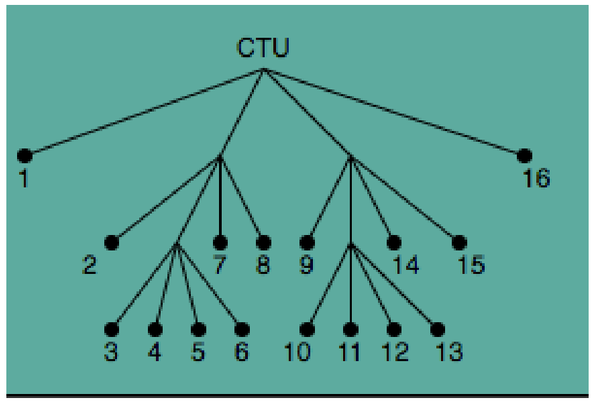
1. 对失真代价的计算精度提出更高要求，h264中基于satd的率失真优化不可行；
2. hevc支持不同CU/PU/TU组合模式，使得可选编码模式数量激增；

解决办法：

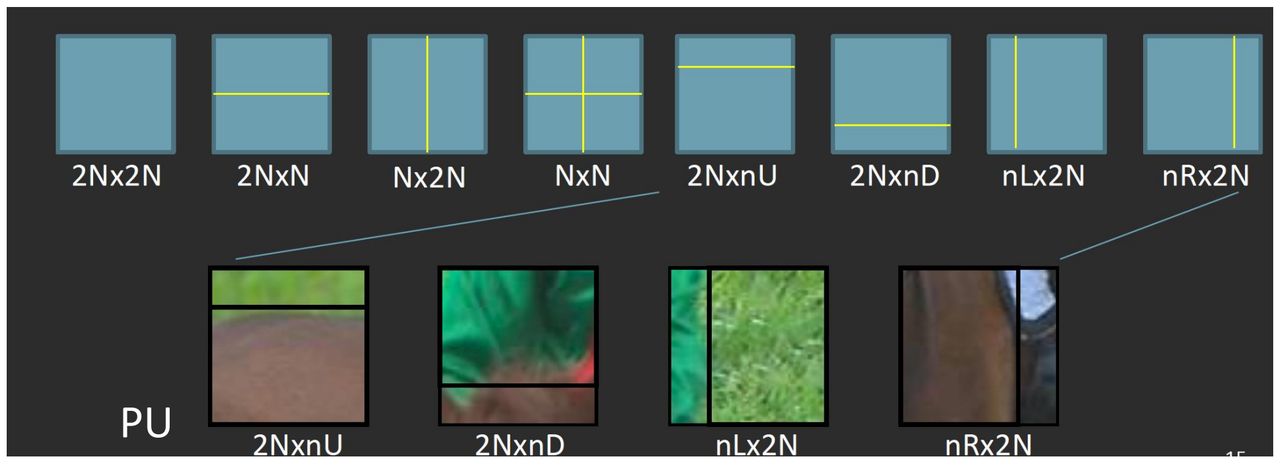
1. 高效预判CU层级；
2. 基于内容的CU遍历提前终止策略；
3. 基于卷积神经网络解决图像分块的非线性问题
4. 提前预判残差AZB块，减少D和R的计算；
5. 量化误差D和残差比特数R的快速计算模型；
6. 基于单调性的ME快速实现算法；
7. 快速帧内预测模式选择（35种模式决策）

HEVC编码之模式划分CTU/CU/PU/TU





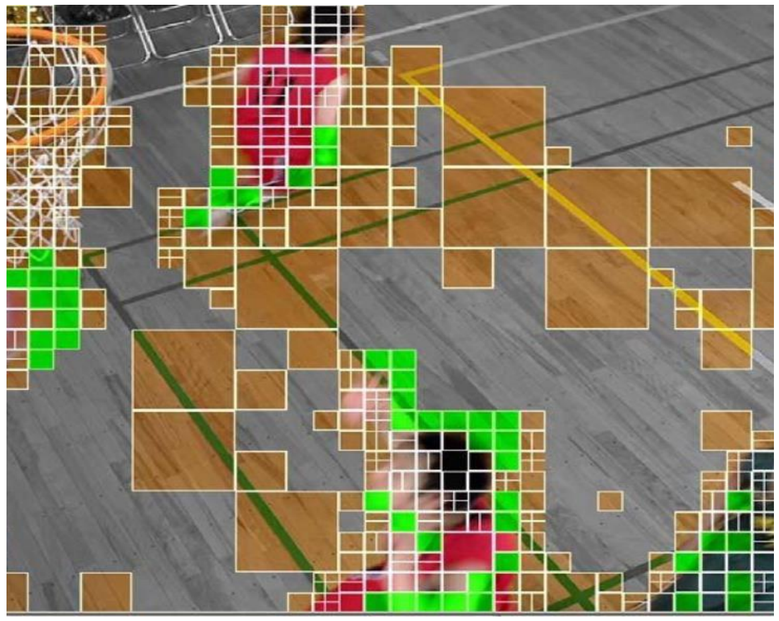
Video Coding Standard	Supported predict block size（PU size)
MPEG-2	16x16
MPEG-4	16x16, 8x8
H264	16x16, 16x8, 8x16, 8x8, 8x4, 4x8 ,4x4
HEVC	64x64, 64x48, 64x32, 64x16, 48x64, 32x64, 16x64, 32x32, 32x24, 32x16, 32x8, 24x32, 16x32, 8x32, 16x16, 16x12, 16x8, 16x4, 12,16, 8x16, 4x16, 8x8 ,8x4, 4x8



HEVC编码之模式划分CTU/CU/PU/TU

CU: 64x64, 32x32, 16x16, 8x8 共计4种；
PU: 64x64, 64x48 ….., 8x8 ,8x4, 4x8 共计24种；
TU: 32x32, 16x16, 8x8, 4x4, 共计4种；

全遍历：192 — 384 种选择！

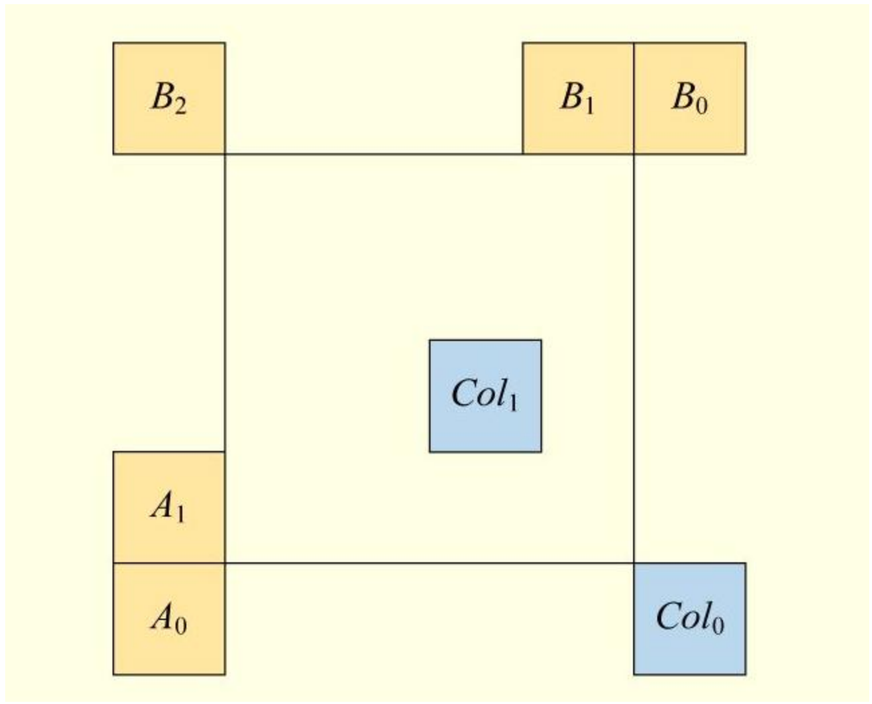


RDO 优化-快速模式决策

深度预估预估：

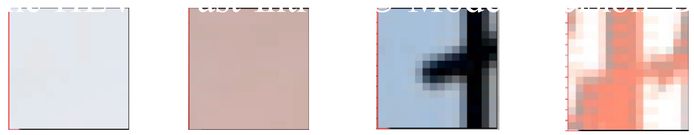
1. 利用时间和空间相关性，从参考块的depth来预估本块的depth范围（ Min_depth, Max_depth)

2. 结合本块的运动和纹理信息，界定depth的精确范围 (Exact_Min_dpeth, Exact_Max_depth)



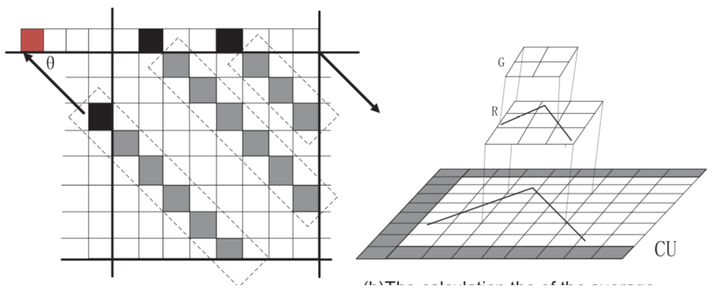
RDO 优化-快速模式决策

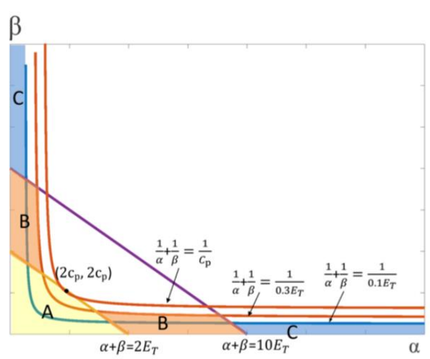
纹理Corner检测



平坦纹理：不再Split

带Corner的纹理，按照强度判决是否继续划分



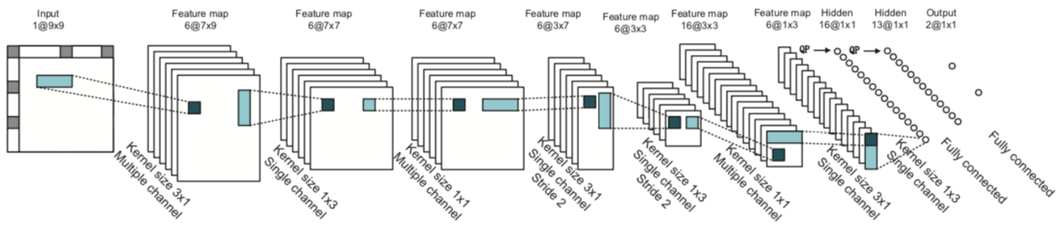


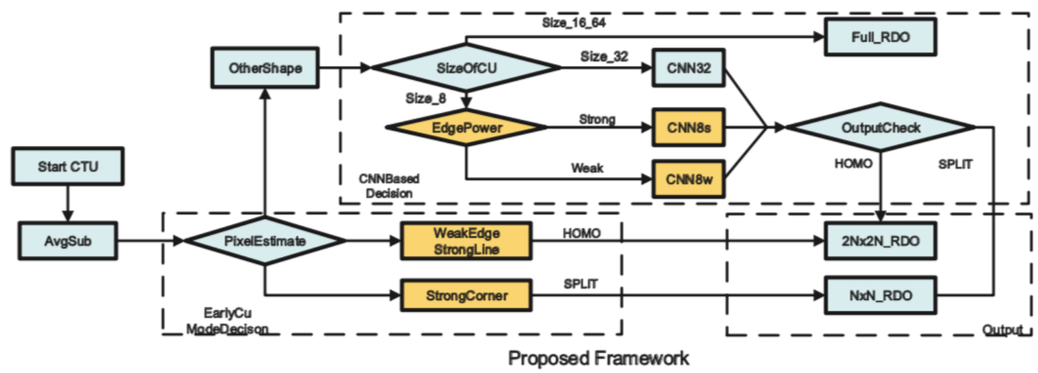
RDO 优化—快速模式决策

CNN分类

虽然通过纹理检测的方法，可以区分 > 70%的情况，

但是还有30%的情况无法用简单的线性方法进行区分，所以需要引入非线性分类方法





Proposed Framework

RDO 优化—快速模式决策

AZB决策

All zero block是量化后系数为全0的块；

1. Skip AZB提前判决
2. 非Skip 的 AZB提前判决

RDO 优化— distortion和bits估计

$p0 = \min(Dp + \lambda R(p))$

1. 其中Dp 的计算方法为：原图和重建图像间的SSE，这需要对编码模式p，完成其预测、变换、量化、反量化、反变换、重建；这个计算过程非常冗长和复杂；

解决方法：可以在变换和量化后，在频域把残差能量计算出来；

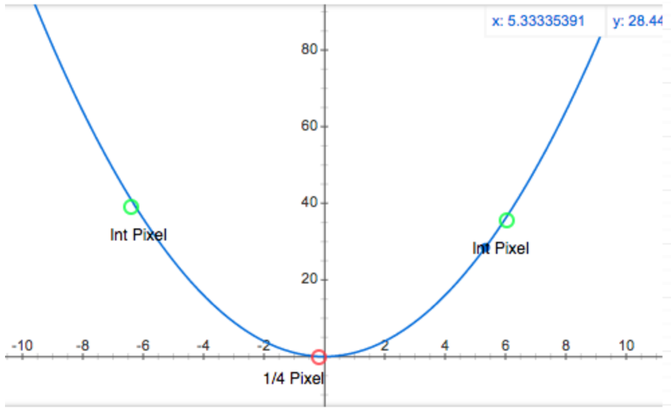
2. 其中码率统计$\lambda R(p)$是通过一次熵编码来实现，计算效率很低；

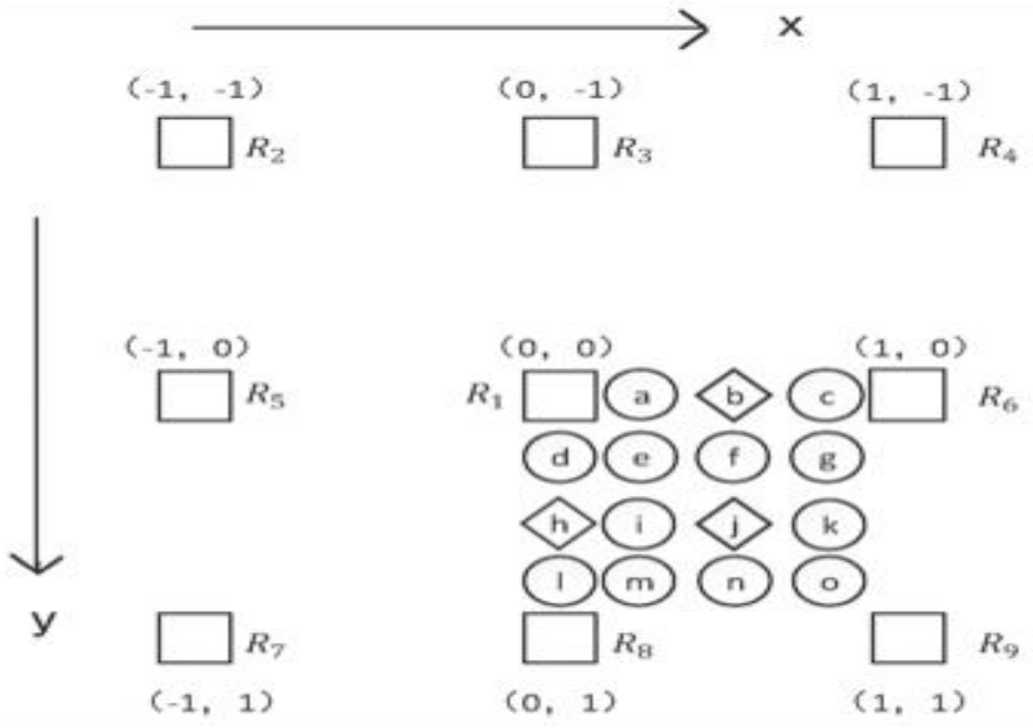
解决办法：对残差数据的码率统计建立线性估计模型，根据NxN变换矩阵量化后系数的特诊，估计其码率；

bdpsnr损失仅为-0.04db，但这部分的计算量可以减少50%以上；

RDO 优化— FME最优搜索位置估计

可通过整像素点及1/2像素点的sad值和其坐标值，估算出最优1/4像素点

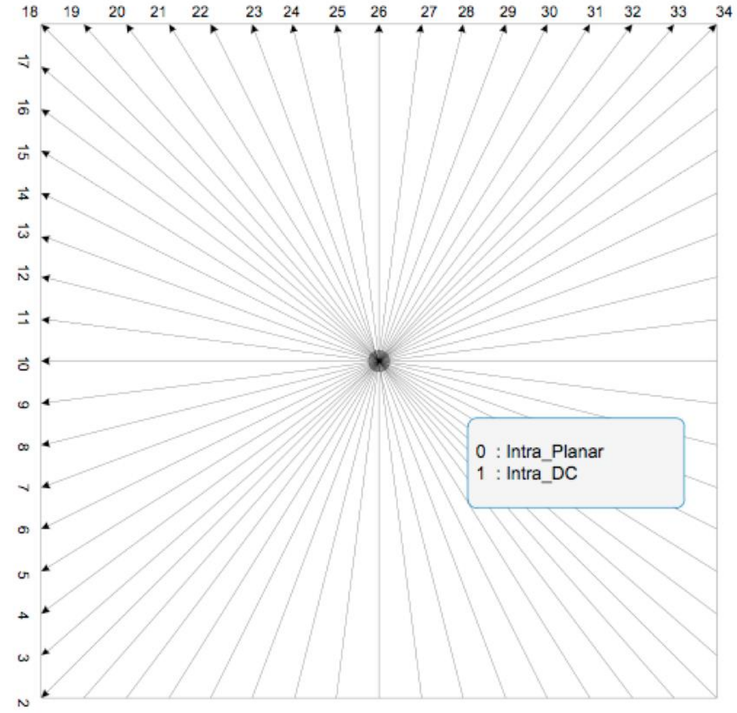


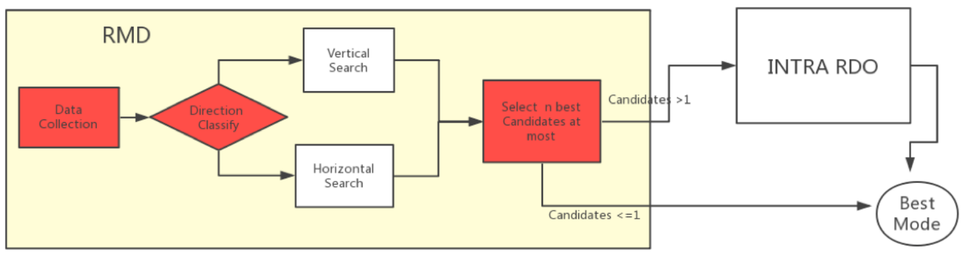


RDO 优化— 快速帧率预测方法(Patent)

基于贝叶斯模型的帧内预测快速决策方法

帧内预测速度提升100%, 仅有0.01db 损失

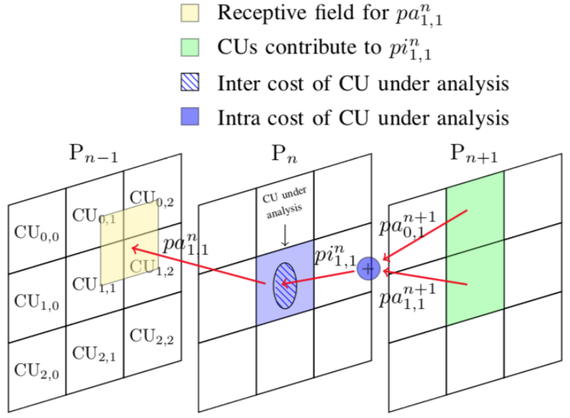




码率控制优化

- 码率控制及lookahead优化

1. 基于CuTree的信息传递来调整CUQP;
2. 基于rates和复杂度的IBP FrameQP ；
3. 基于参考强度的SliceType decision



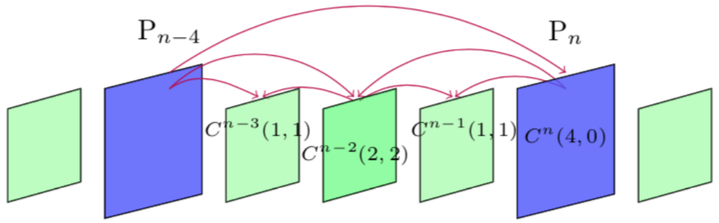
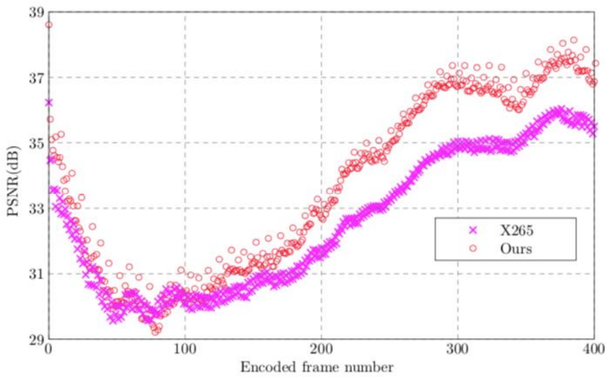
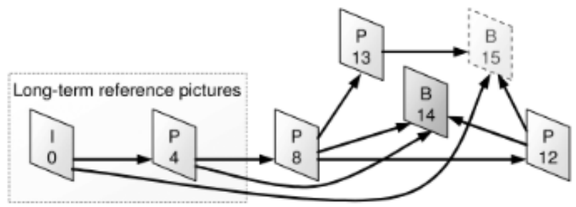


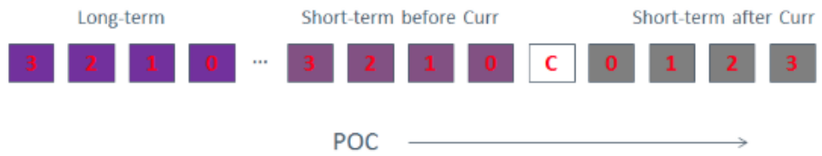
Fig. 3: C_P and C_B definition of a MiniGOP with $N_B = 3$ (MiniGOP under analyze is composed of the P-frame P_n and its leading three B-frames; $C_P = C^n(4,0)$; $C_{I|P} = C^n(0,0)$; C_B is the average of $C^{n-1}(1,1)$, $C^{n-2}(2,2)$ and $C^{n-3}(1,1)$.)

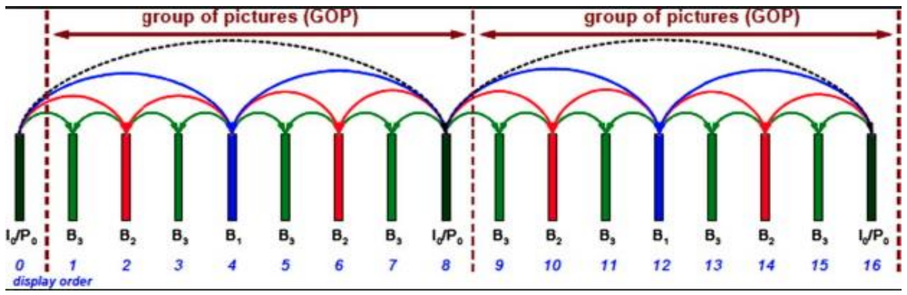


参考帧优化 – 长期参考帧

1. 长期参考帧往往有着更好的编码质量，有助于提高被参考帧质量；
2. 对于背景很少发生变化的直播场景，长期参考帧尤为重要，可减少信息经过多帧传递带来的损失。
3. 平均可提到0.25db。

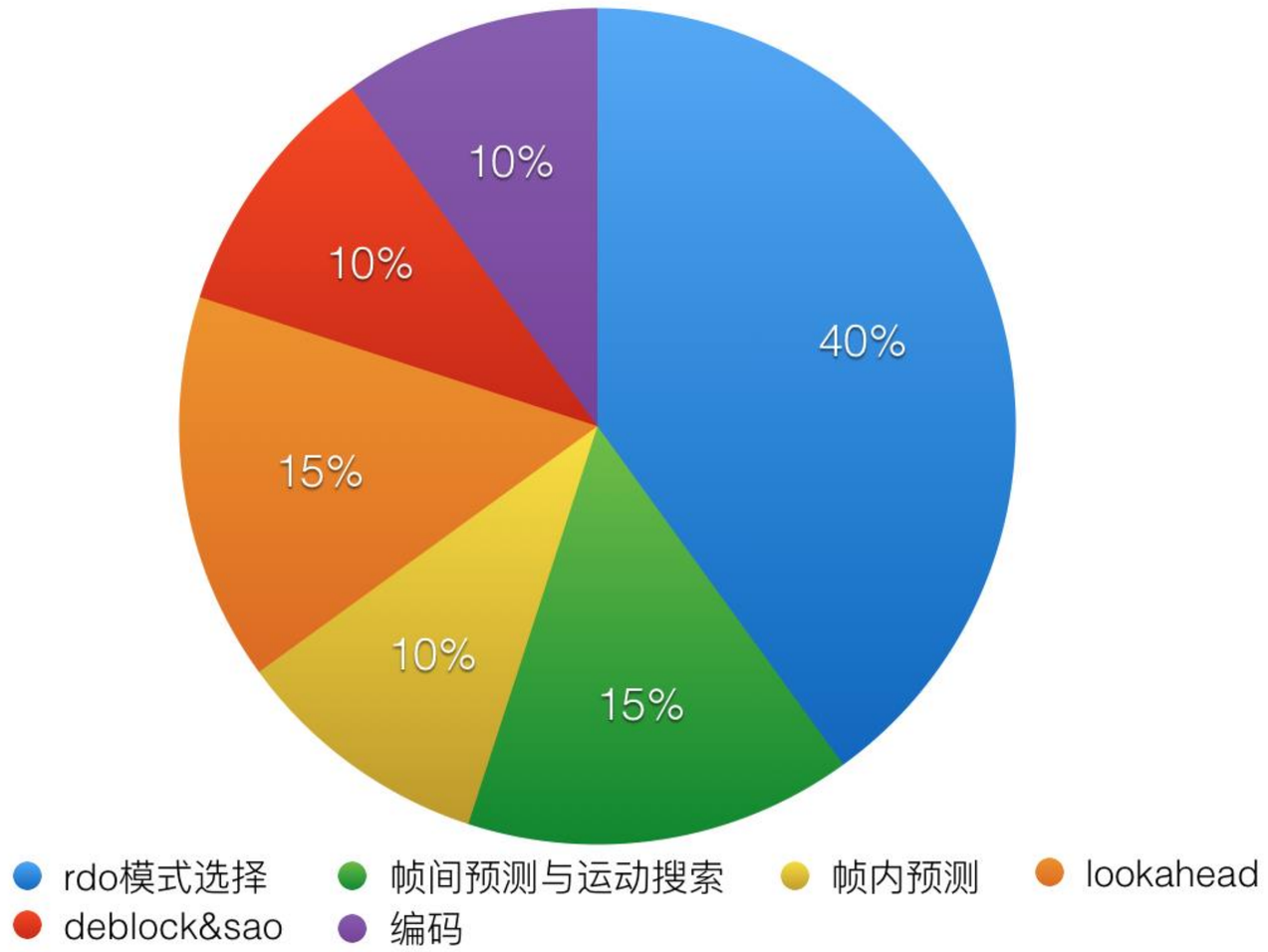






编码过程的计算瓶颈

各模块计算量占比



工程化优化方法

工程化优化

1. NEON指令集优化 ，典型计算提高2-4倍性能；
2. 多核并行计算，适应现代处理器架构；
3. 优化bottleneck的指令和访存，不可小觑的力量；

Neon 优化提升比例

模块	提升比例
RDO(SSE, SAD)	>300%
运动搜索(ME,MC)	>250%
帧内预测(Intra Pred)	>200%
变换与量化(T,Q,T-1, Q-1)	>200%
SAO	>200%
Deblock	>100%
Lookahead	>300%
熵编码	>50%

工程化优化方法

工程化优化

1. NEON指令集优化 ，典型计算提高2-4倍性能；
2. 多核并行算，适应现代处理器架构；
3. 优化bottleneck的指令和访存，不可小觑的力量；

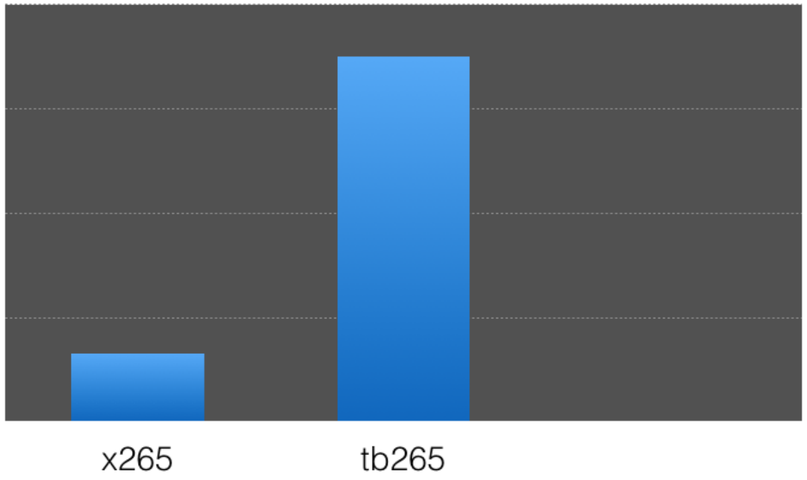
Neon 优化提升比例

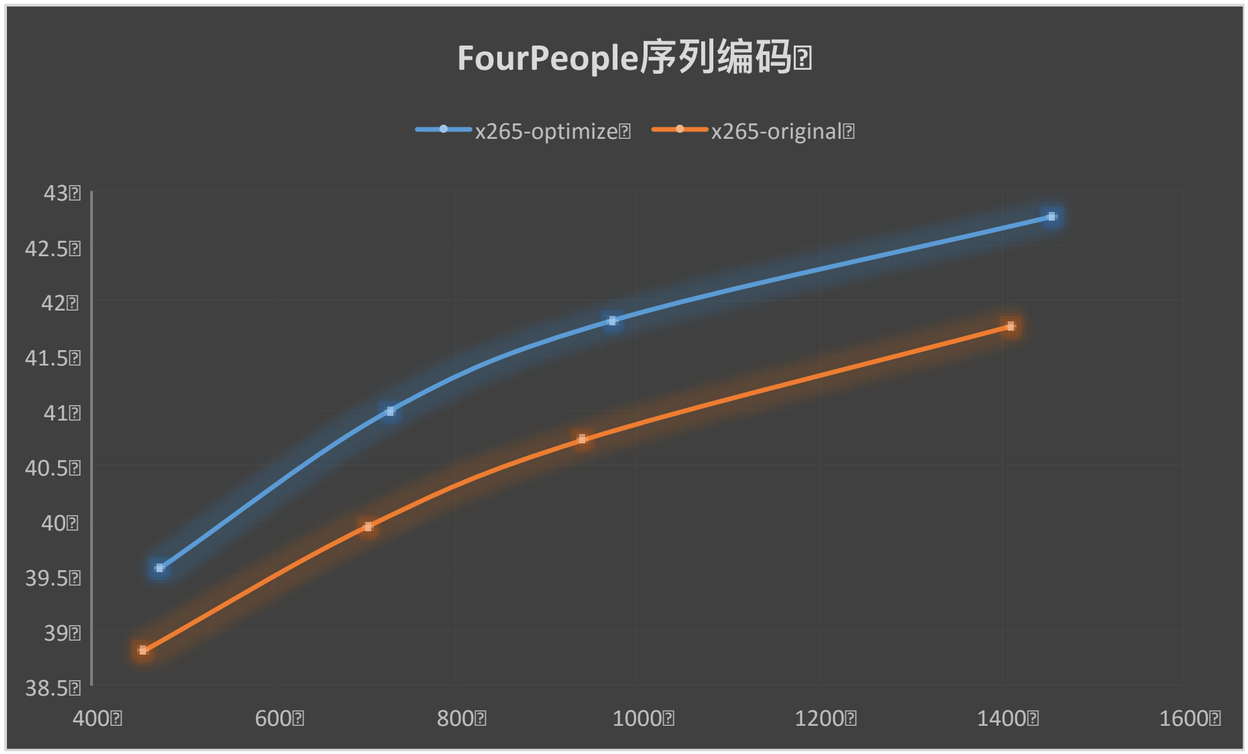
模块	提升比例
RDO(SSE, SAD)	>300%
运动搜索(ME,MC)	>250%
帧内预测(Intra Pred)	>200%
变换与量化(T,Q,T-1, Q-1)	>200%
SAO	>200%
Deblock	>100%
Lookahead	>300%
熵编码	>50%

手淘HEVC解码器

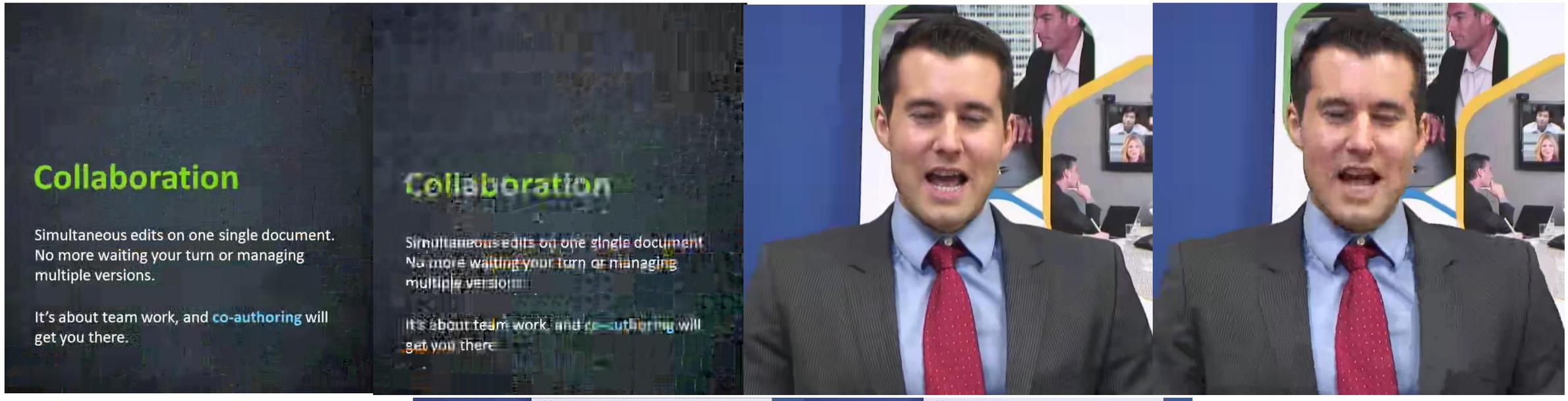
编码速度 VS X265 > 450%，在iphone6可完成720p 30fps实时编码；

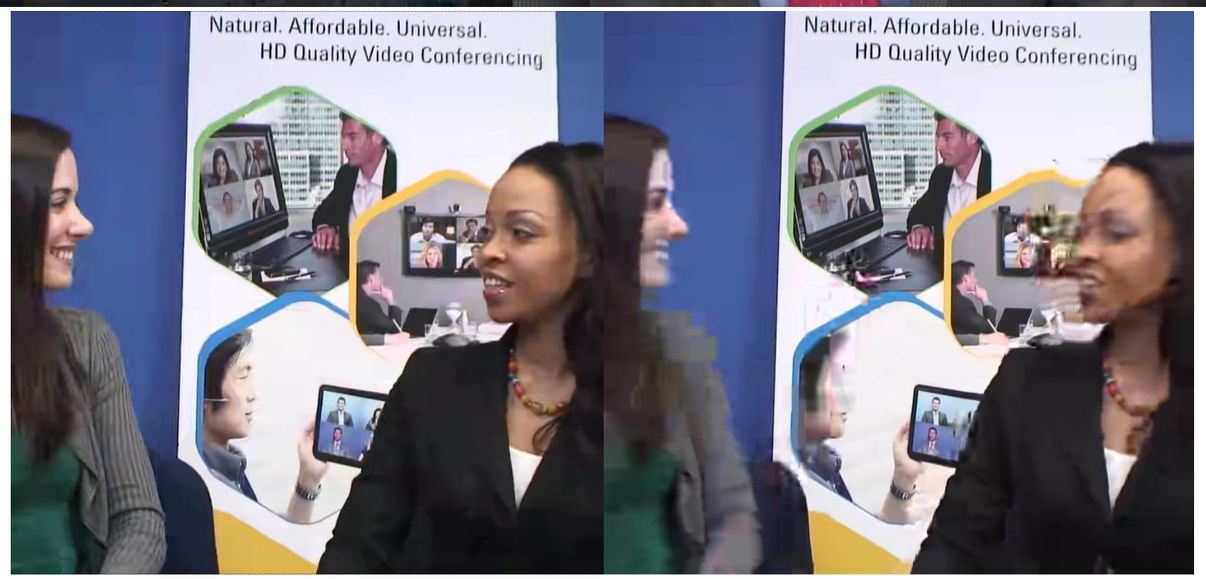
编码质量 VS X265　1Mbps psnr > 0.6db；
编码码率 VS X265　下降 > 15%;





手淘HEVC编码器质量比较





手淘HEVC编码器质量比较

Medium：
bdpsnr提升
>0.5db，
bdrate下降
>15%

medium	singlethread	vS X265	vS X265
class	video	bdpsnr	bdbr
d	BasketballPass_416x240_50	0.5123	-10.62
	BlowingBubbles_416x240_50	0.8624	-21.64
	BQSquare_416x240_60	1.7205	-41.37
	Flowervase_416x240_30	0.5058	-14.13
	RaceHorses_416x240_30	0.2321	-5.37
c	BasketballDrill_832x480_50	0.6759	-17.91
	BQMall_832x480_60	0.5025	-12.51
	Flowervase_832x480_30	0.3813	-12.56
	PartyScene_832x480_50	0.6773	-20.21
	RaceHorses_832x480_30	0.1576	-4.9
e	FourPeople_1280x720_60	0.9236	-28.27
	Johnny_1280x720_60	0.8728	-46.77
	KristenAndSara_1280x720_60	0.7858	-28.52
	vidyo1_720p_60	0.5723	-22.76
	vidyo3_720p_60	0.6529	-20.71
	vidyo4_720p_60	0.6304	-22.52
	2016zaowujie01_1280x720_30	0.4846	-8.42
b	BasketballDrive_1920x1080_50	0.1524	-5.38
	BQTerrace_1920x1080_60	0.5113	-29.52
	Cactus_1920x1080_50	0.4367	-15.62
	Kimono1_1920x1080_24	0.1358	-4.61
	ParkScene_1920x1080_24	0.6103	-19.61
f	BasketballDrillText_832x480_50	0.7094	-18.6
	ChinaSpeed_1024x768_30	0.3779	-10.84
	SlideEditing_1280x720_30	0.1211	-0.48
	SlideShow_1280x720_20	1.4936	-17.29
a	Traffic_2560x1600_30	0.6232	-21.28
	PeopleOnStreet_2560x1600_30	0.1029	-2.46
average		0.58659643	-17.317143

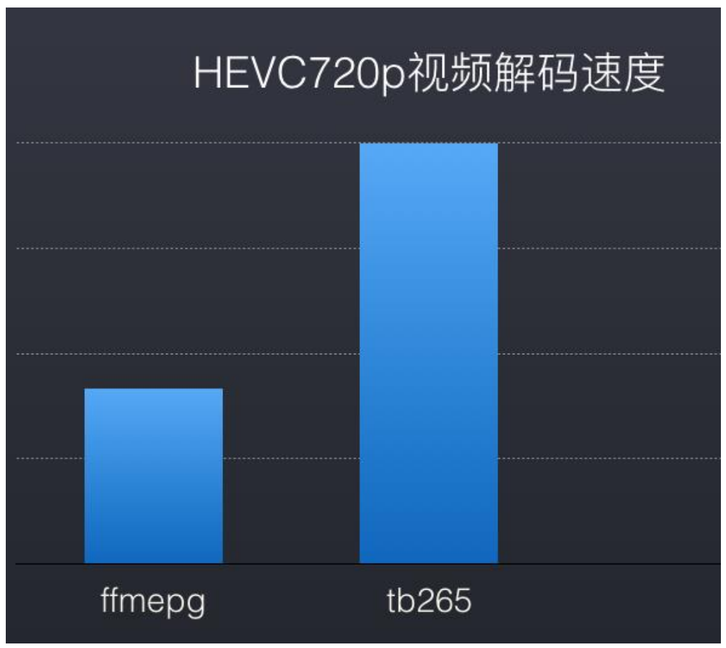
Slower：
bdpsnr提升
>0.5db

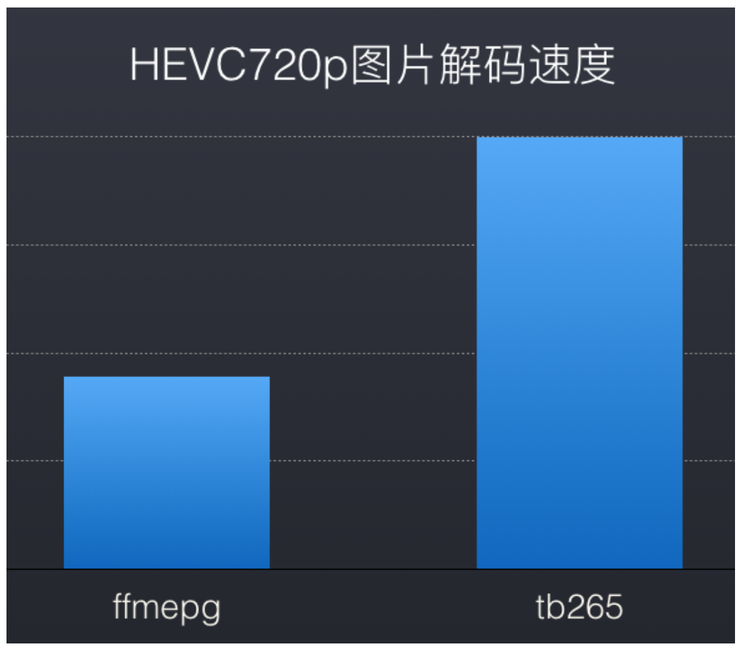
bdrate下降
>15%

slower	singlethread	vS X265	vS X265
class	video	bdpsnr	bdbr
d	BasketballPass_416x240_50	0.6203	-12.3
	BlowingBubbles_416x240_50	0.9004	-22.14
	BQSquare_416x240_60	1.7289	-41.11
	Flowervase_416x240_30	0.4939	-13.67
	RaceHorses_416x240_30	0.2473	-5.44
c	BasketballDrill_832x480_50	0.9707	-24.53
	BQMall_832x480_60	0.6141	-15.24
	Flowervase_832x480_30	0.3816	-12.81
	PartyScene_832x480_50	0.7367	-21.73
	RaceHorses_832x480_30	0.2284	-6.95
e	FourPeople_1280x720_60	0.9025	-28.48
	Johnny_1280x720_60	0.8856	-42.9
	KristenAndSara_1280x720_60	0.7298	-27.67
	vidyo1_720p_60	0.5399	-22.53
	vidyo3_720p_60	0.6725	-20.89
	vidyo4_720p_60	0.5572	-20.78
	2016zaowujie01_1280x720_30	0.5277	-9.57
b	BasketballDrive_1920x1080_50	0.255	-9.21
	BQTerrace_1920x1080_60	0.5571	-33.71
	Cactus_1920x1080_50	0.5438	-20.6
	Kimono1_1920x1080_24	0.171	-5.92
	ParkScene_1920x1080_24	0.6923	-21.85
f	BasketballDrillText_832x480_50	1.0172	-24.96
	ChinaSpeed_1024x768_30	0.4524	-11.29
	SlideEditing_1280x720_30	-0.0064	0.54
	SlideShow_1280x720_20	1.4582	-17.17
a	Traffic_2560x1600_30	0.6313	-18.65
	PeopleOnStreet_2560x1600_30	0.1434	-3.32
average		0.63045714	-18.388571

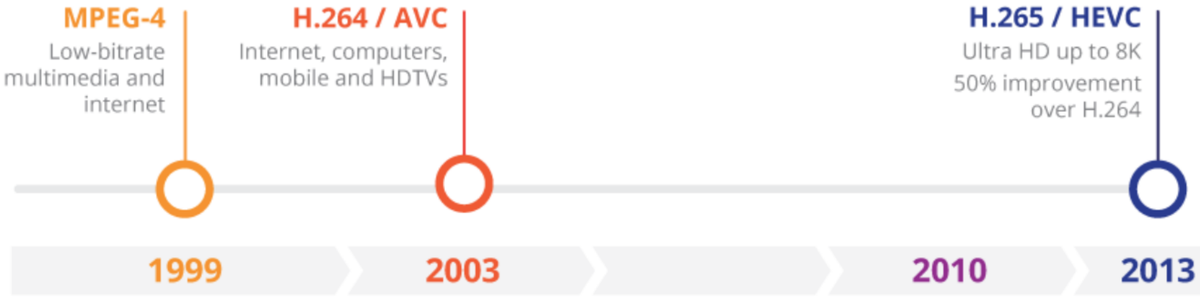
手淘HEVC软解码器

1. 视频解码速度比ffmpeg提升 >150%
2. 1Mbps 720p h265在小米5手机上解码速度>200fps，CPU占比控制在20%以下；
3. APG图片解码速度比ffmpeg提升>70%;

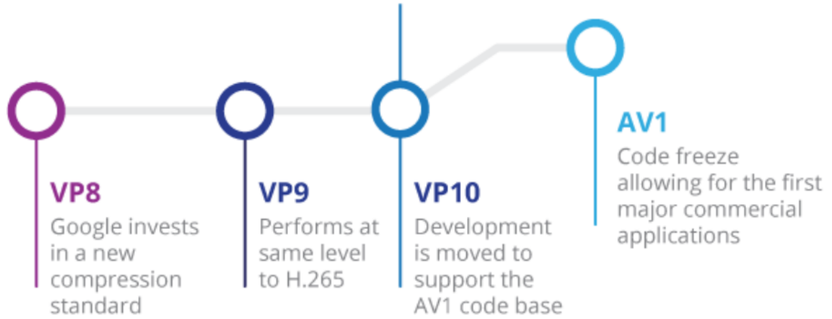




视频编码标准演进


MPEG-4
Low-bitrate multimedia and internet
H.264 / AVC
Internet, computers, mobile and HDTVs
H.265 / HEVC
Ultra HD up to 8K
50% improvement over H.264
1999
2003
2010
2013
2015
2018


Alliance for Open Media
AOM is formed and work on the AV1 codec begins
IBM
amazon web services
Google
mozilla
NETFLIX
cisco
BITMOVIN
intel


VP8
Google invests in a new compression standard
VP9
Performs at same level to H.265
VP10
Development is moved to support the AV1 code base
AV1
Code freeze allowing for the first major commercial applications

H266

四叉树二叉树结构（QTBT）bd-rate 提升4%

联发科公司提出的QTBT块结构被JVET采纳，集成在JEM3.0及其后的版本中，替代了原来HEVC的四叉树（QT）块结构。QTBT与四叉树划分的主要区别：

1、QTBT块结构在四叉树的叶结点，继续往下做二叉树递归遍历，其中二叉树可以水平划分或垂直划分。

2、在QTBT块结构中，CB=PB=TB，也就是说取消了PU/TU/PB/TB的概念，预测块即是变换块。

3、I slice的亮度分量的CTB和色度分量的CTB划分脱钩，单独划分。

4、四叉树结构的亮度分量的CTB最大为64x64，而QTBT结构亮度分量的CTB的大小默认从128x128开始，且最大可以设置为256x256

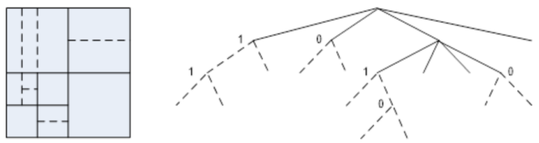
QTBT性能：

1、BD-Rate平均在-4%左右，LD配置增益最大，RA配置增益次之，AI配置增益最小。

2、编码时间是原来的2.2~5.6倍，AI配置增加的编码时间最多（5.6倍），该算法的编码复杂度在提案JVET-C0105中被诟病；解码时间是原来的105%~112%。

FRUC模式

QTBT块结构图例：



从上图可以看到，QTBT块结构先做四叉树划分，然后在四叉树叶结点做二叉树划分。

AV1

2015年9月1日，谷歌宣布与亚马逊、思科、英特尔、微软、火狐、奈飞（Netflix）成立开放媒体联盟（Alliance of Open Media，简称AOM），专注于下一代视频编码格式的开发。新一代视频编码格式命名为AV1，被认为是开源编码器VP9的继承者，具有以下几个特点：

1. Open and Free(Royalty)
2. Optimized for the Internet(W3C,webrtc)
3. Overwin of hevc(30% performance improve)

问题：

1. 编码速度慢，refereance model编码时间是x265的2500倍 - 3000倍
2. 虽然AV1号称自己编码性能优于hevc 25%-30%，但第三方测试表明bdrate不如hevc