



2020
北京

多媒体开启
MULTIMEDIA BRIDGE
TO A WORLD OF VISION

新视界

基于人眼主观视觉 的画质评价与提升



目录

CONTENTS



2020
北京

01 业务形态及目标

02 直播视频质量评价

03 服务端窄带高清转码

04 客户端超分增强



2020
北京

PART 1

业务形态及目标

业务



2020
北京

很高兴
认识你



直播



短视频





2020
北京

目标

以人为本，给用户美的体验

画质

码率

视频画质**增强**

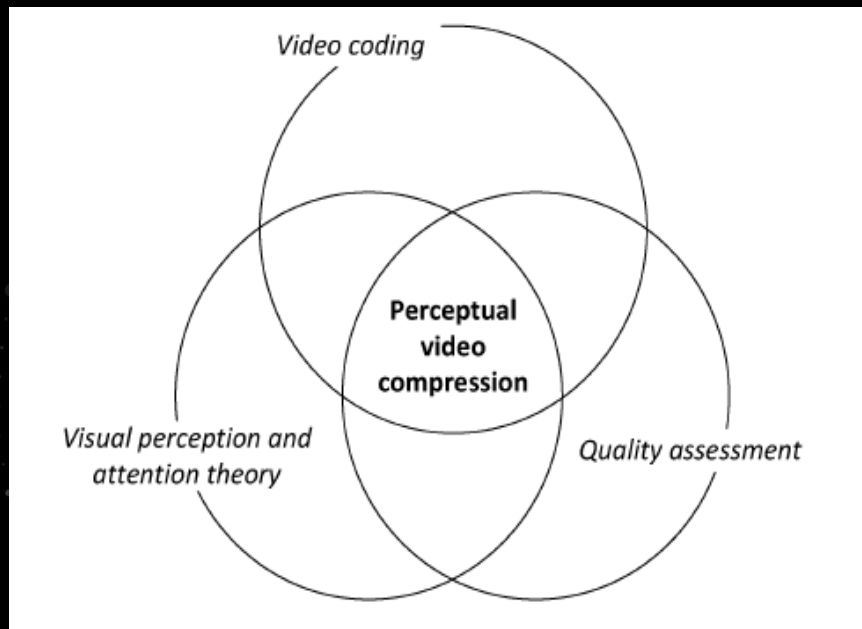
- 码率压缩，体验不变
- 码率压缩，**体验变好**

面向提升质量的视频处理和编码

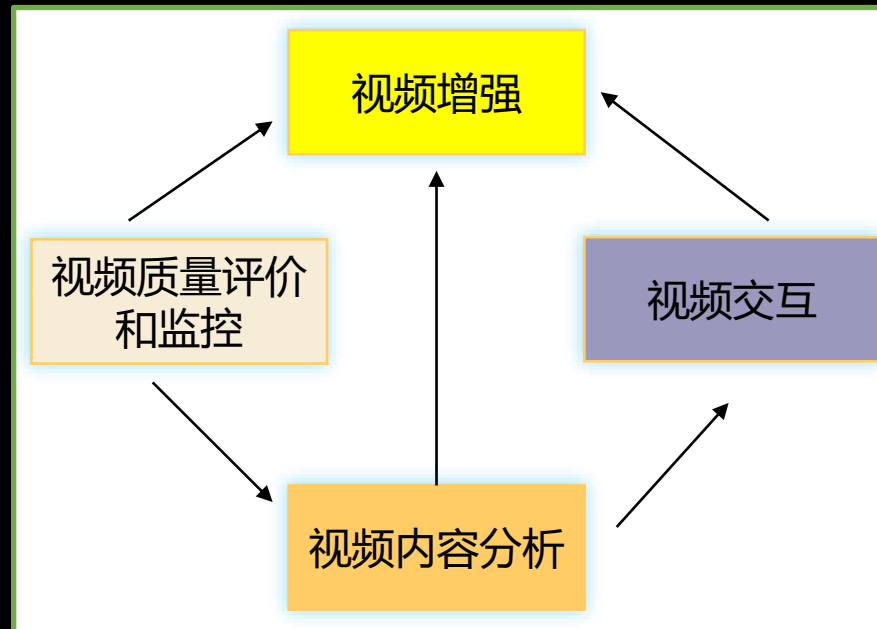


2020
北京

- 感知编码



- 超分增强





2020
北京

PART 2

视频质量评价



2020
北京

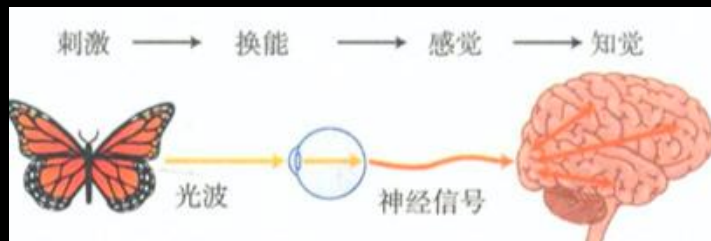
评价的定义

从人眼主观特性出发

- 换能和阈值

因人而异，趋势收敛

- 绝对阈值: absolute threshold
- 差别阈值: difference threshold, JND Just Noticeable Difference
 1. 一个人50%概率区分不同的界限
 2. 人眼最大不可区分的界限



“一个比另一个清晰很多”
“我看不出什么差别”

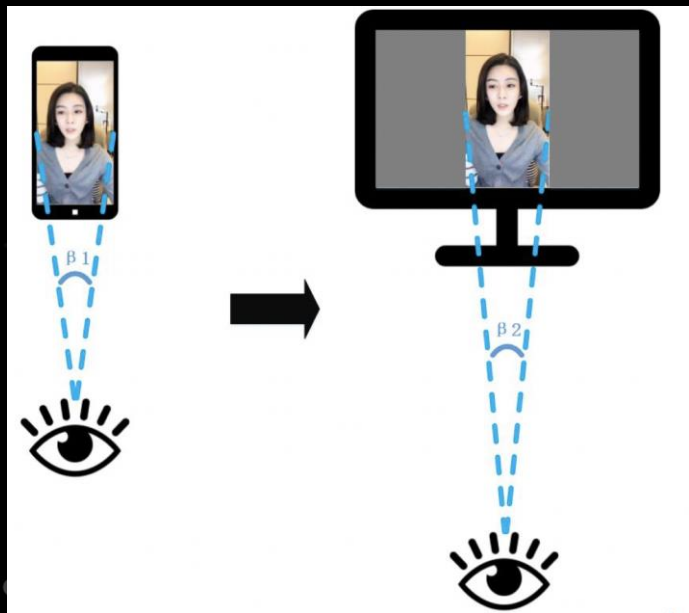


2020
北京

正确的执行测试

- ITU-R BT.2022指出显示器更新

	屏幕尺寸 (英寸)	分辨率
iphoneX	6.5	2688x1242
Mac pro	15.4	2880x1800
Dell P2414H	23.5	1920x1080



显示设备对于观感有一定影响

画质评价标准



2020
北京

α 美学分 + β 清晰度分 = 质量分



22分



39分



60分



69分



79分



88分



97分



2020
北京

评分模型应用及效果



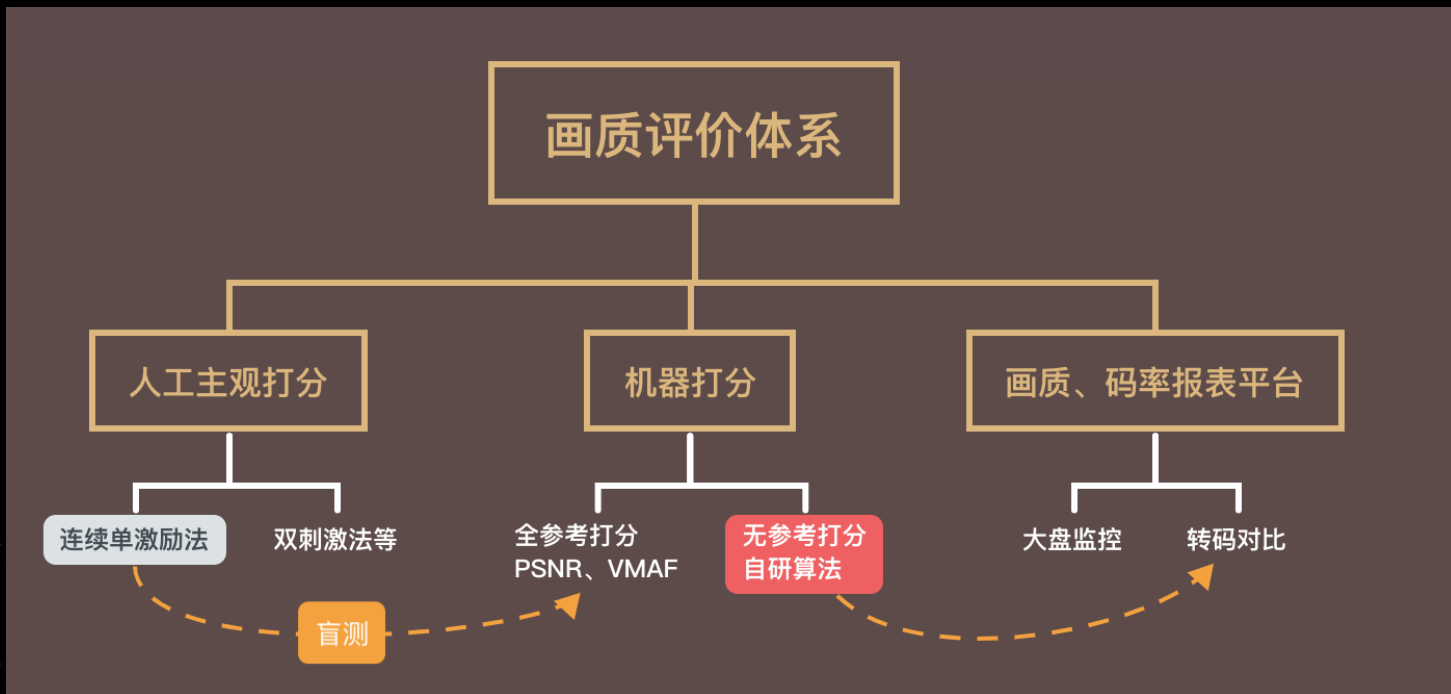
黄色; 人工打分
绿色: 算法打分

基本趋势一致，具体有差异



2020
北京

画质评价体系



专家作用



2020
北京

- 黄金眼
 - 上帝视角
 - 宏观效果到细节
- 大规模主观测试：验证 和 纠偏
 - 解决规模问题
 - 大众喜好



2020
北京

PART 3

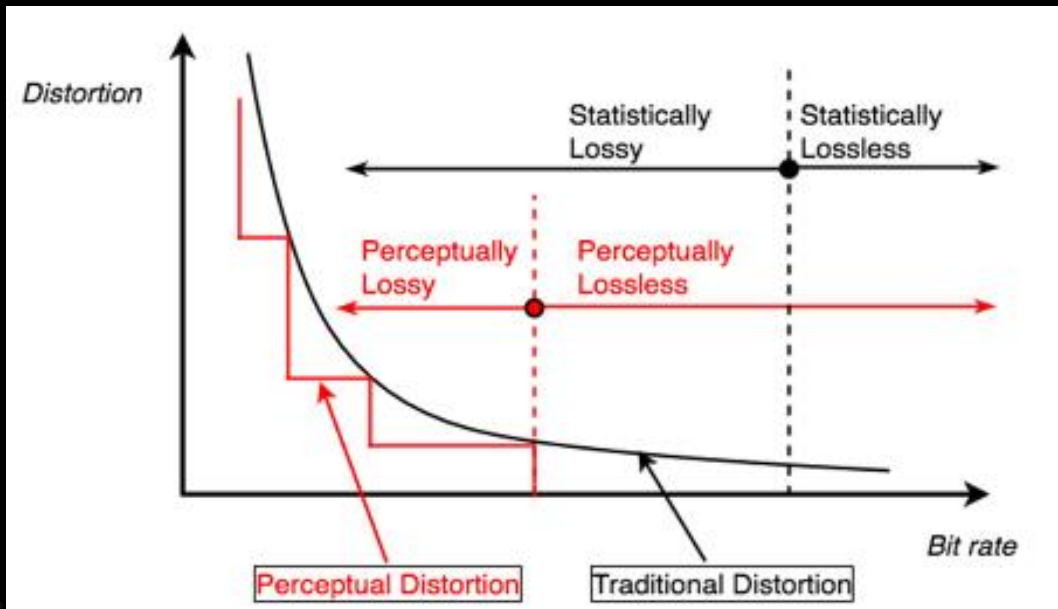
智美高清

- 基于人眼视觉特性的编码优化
- 画质增强



2020
北京

压缩后主观体验不变： why?



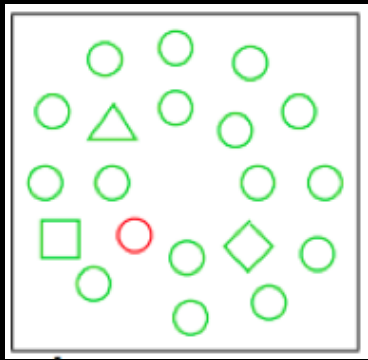
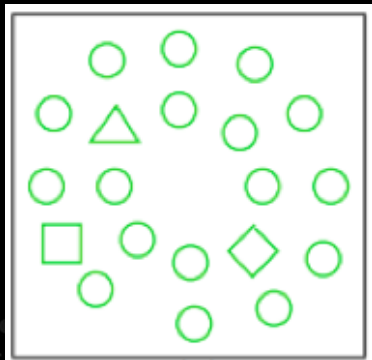
传统RD曲线没有考虑人眼特性

- 人眼无法感知像素细小变化
- 人眼感知明显JND特性



2020
北京

压缩后体验不变：ROI-显著性



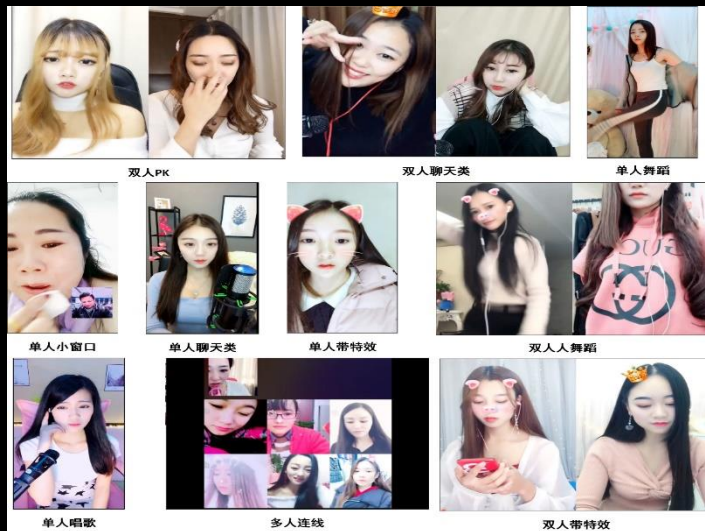
人脑对整图的关注度不均衡，对注意力焦点投入了更多的资源

深度学习显著性估计



2020
北京

- 设备：眼动仪（Eye Link 1000 plus）
- 视频数：597个，包括单人多人，聊天类，唱歌类，舞蹈类等类别
- 采集人数：35（男22，女13）



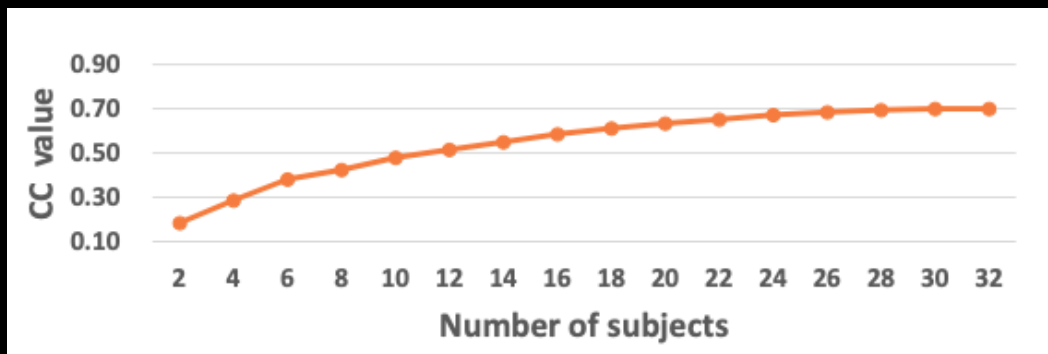
数据示例及一致性



2020
北京



显著性实验一致性及人数



人数增长，趋势收敛

DeepVS, 2018

主观实验



2020
北京



分两个Session测试，每个Session在30min左右，测试前先观看5个训练视频



播放视频顺序随机，视频质量类型随机，保证同一视频的不同类型不连续播放



测试者根据主观感受对视频质量进行打分，记录原始的MOS分



打分条采用百分制，且分成了相等的5段，每段表示不同的质量等级

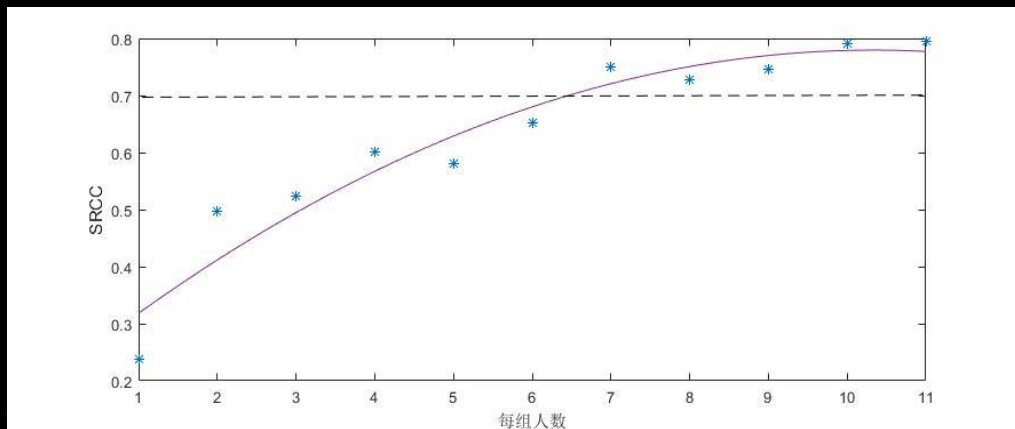
The interface shows a progress bar at the top with the text '进度: 2 / 26' and a '下一个>>' button on the left and a '打分' button on the right. Below this, the text '整体质量' (Overall Quality) is displayed above the numerical value '73.4509'. At the bottom, there is a horizontal scale bar divided into five equal segments, labeled 'Bad', 'Poor', 'Fair', 'Good', and 'Excellent' from left to right. A red vertical line is positioned at the boundary between the 'Fair' and 'Good' segments, indicating the current score.

主观实验一致性分析



2020
北京

1. 将 $2n$ 名测试者随机等分为两组，分别计算两组测试者的MOS分数
2. 对两组测试者的MOS分数求斯皮尔曼相关系数
3. 重复以上实验15次，并求15个斯皮尔曼相关系数的平均数



- **ITU-R BT.500/1788**
筛选后的观测者数目应至少为15

随着每组人数的增多，相关系数有增长并收敛的趋势

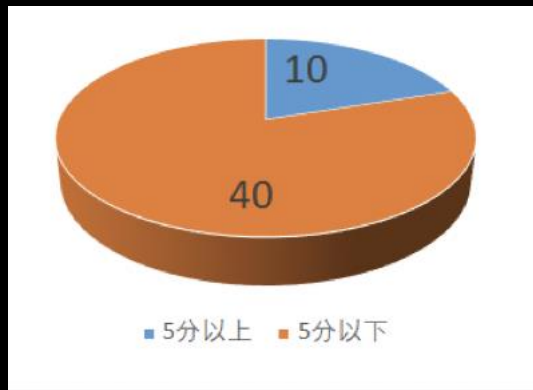
基于显著性编码压缩



2020
北京



标准方法压缩



基于显著性压缩

橙色码率压缩40%视频，蓝色原视频。

大分差视频分析



2020
北京



观察发现，视频中人脸轮廓区域存在不同程度的失真。经过分析，针对上述类别(户外直播、跳舞、头部运动)的视频，人对运动(例如人脸由远及近)过程中出现的轮廓失真情况较敏感。

人脸边缘优化



2020
北京



(a) 原始视频



(b) 压缩视频



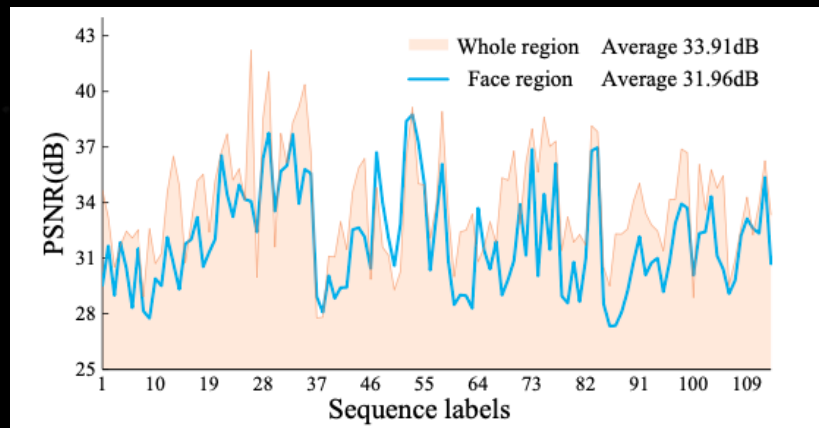
(c) 轮廓优化

MRS-Net

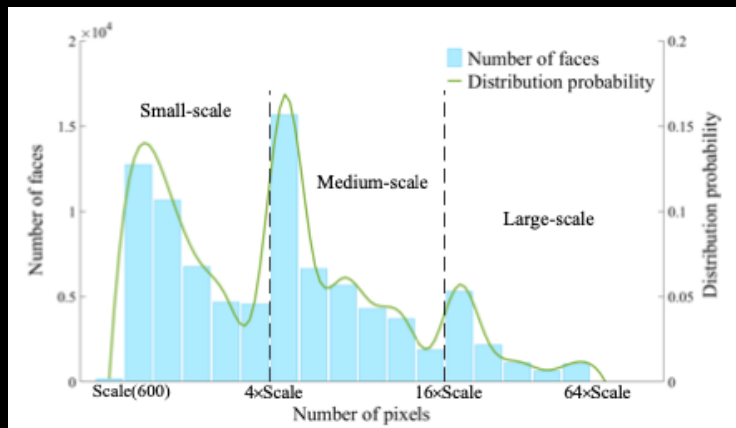
Muti-Scale Recurrent Scalable Network for Face Quality Enhancement of Compressed Video



2020
北京



人脸的质量明显比整图要低

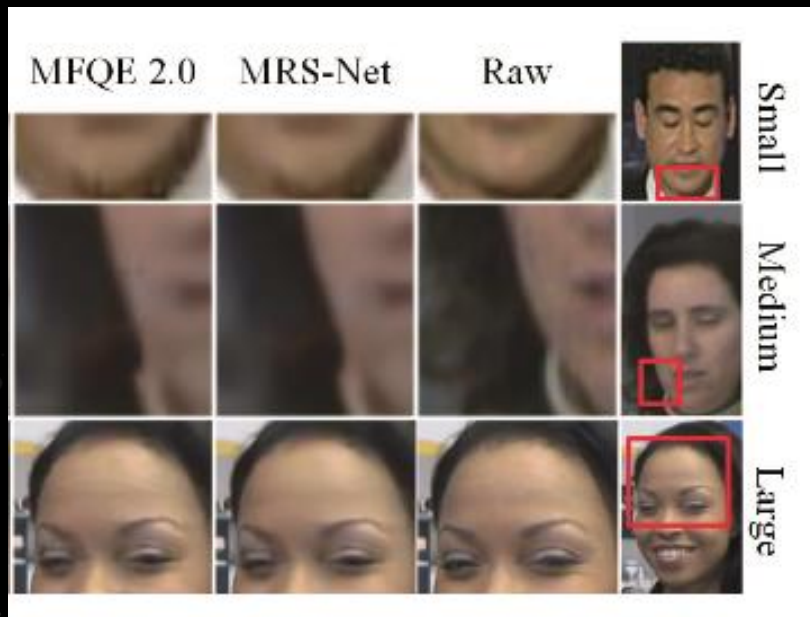


人脸区域呈现明显分布

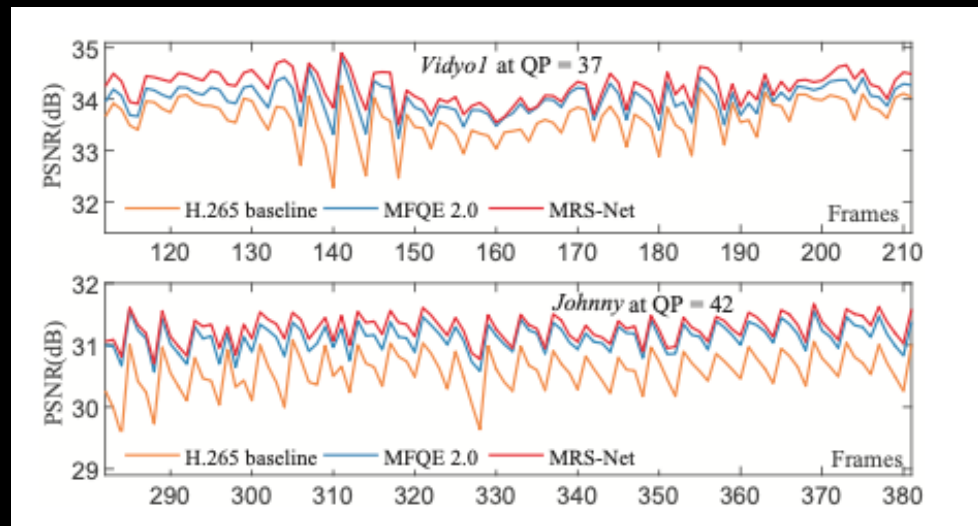
MRS-Net



2020
北京



效果对比

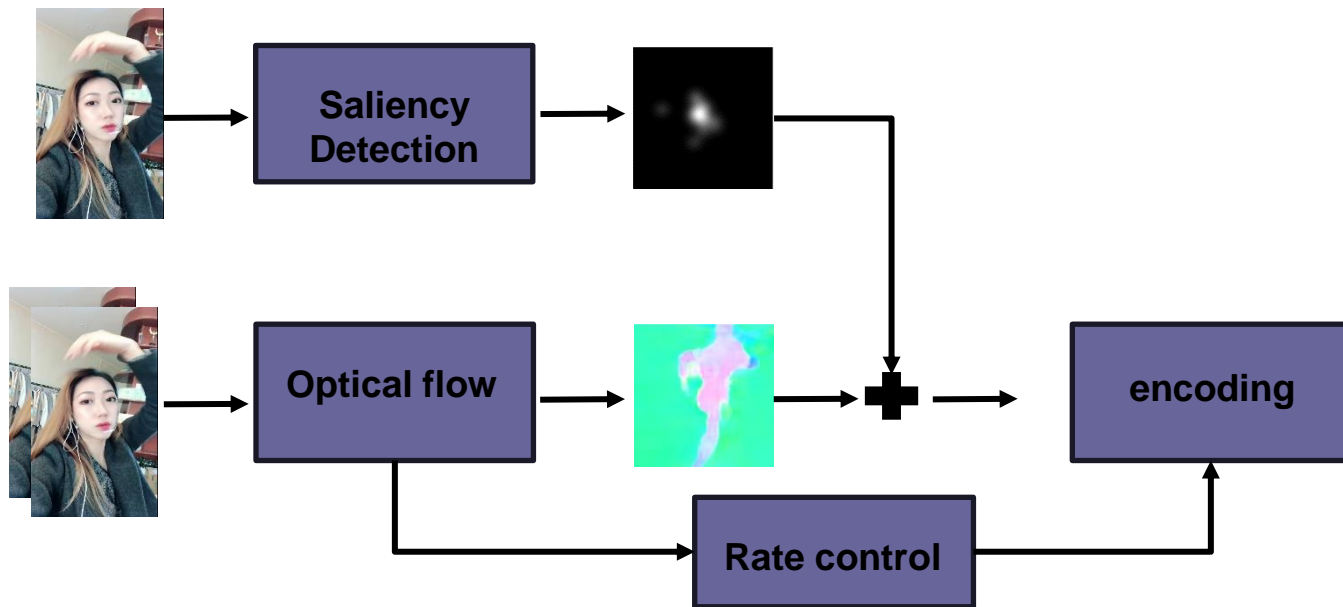


更小的波动

根据光流调整显著性和码率



2020
北京



光流改进



2020
北京



w/o flow



w flow

model	metric	w/o flow	w flow
CC	ave	0.755	0.801
	std	0.118	0.092
KL	ave	0.689	0.544
	std	0.213	0.178
NSS	ave	2.606	2.764
	std	0.597	0.534

JND阈值模型

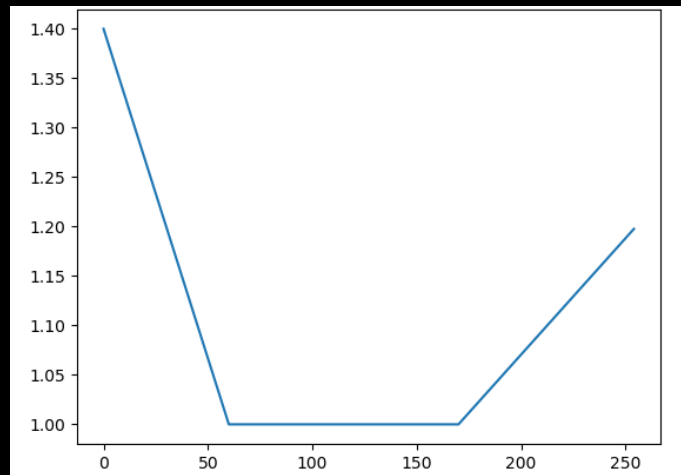


2020
北京

重分配后，可能会导致负面效果



对比度遮蔽效应

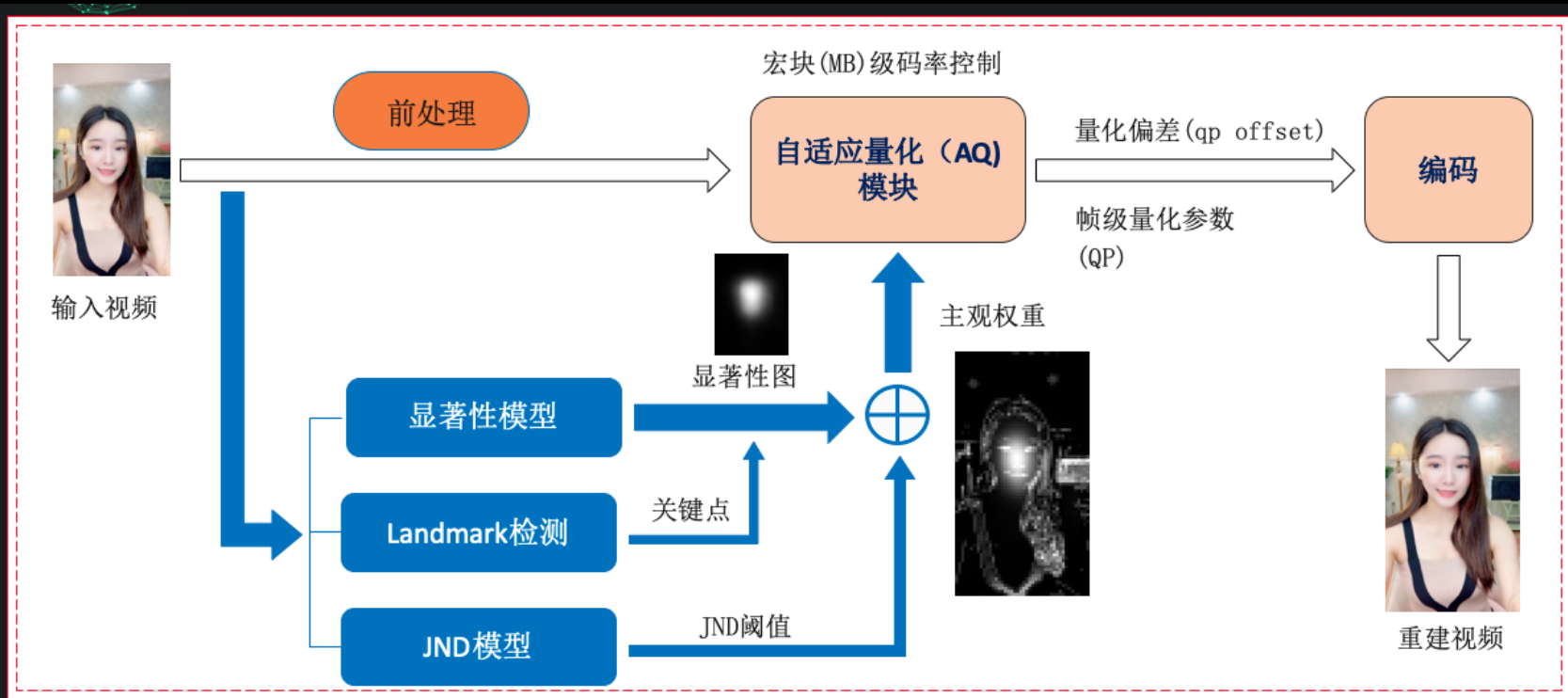


背景亮度遮蔽效应

基于人眼感知的编码框架



2020
北京



画质不变，码率下降



2020
北京

仅单人脸显著性实验结果:

	标准方法	our
大于5分视频数	25	10
合格率	50%	80%

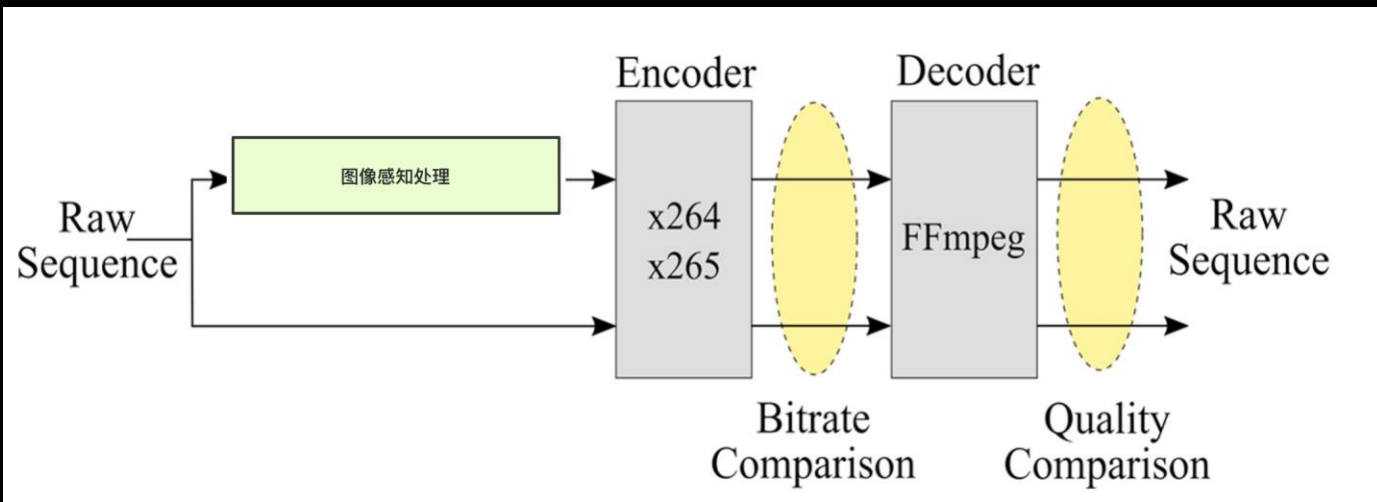
叠加人脸关键点等实验结果:

	标准方法	our
大于5分视频数	10	3
合格率	75%	92.5%

画质增强



2020
北京



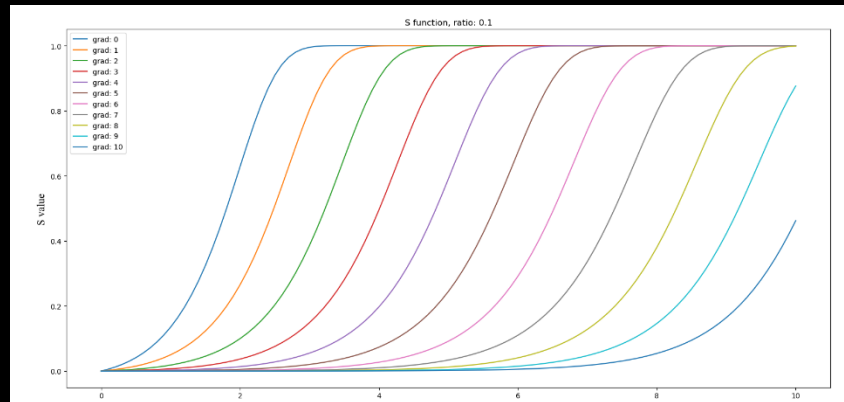
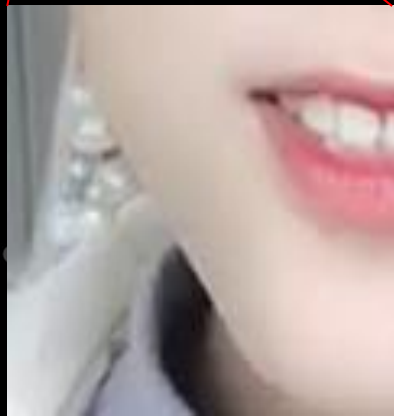
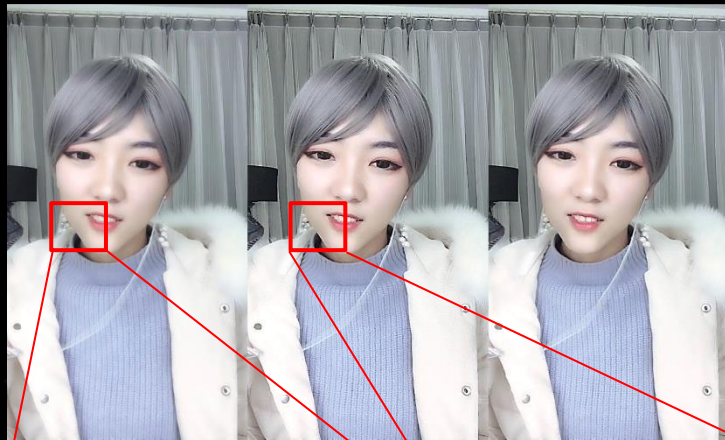
前置处理能够有效提升编码画质，降低码率

2017, New adaptive filters as perceptual preprocessing for rate- quality performance

算法增强



2020
北京



基于S函数的自适应增强

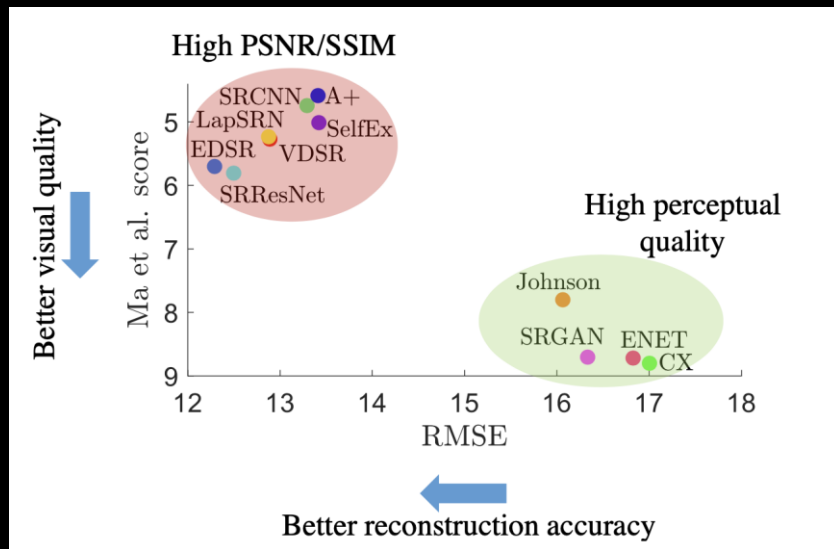
力度过大，光环效应



2020
北京

深度学习增强

增强的本质：主观感觉变好



2018, The 2018 PIRM Challenge on Perceptual Image Super-resolution

码率下降，画质变好



2020
北京



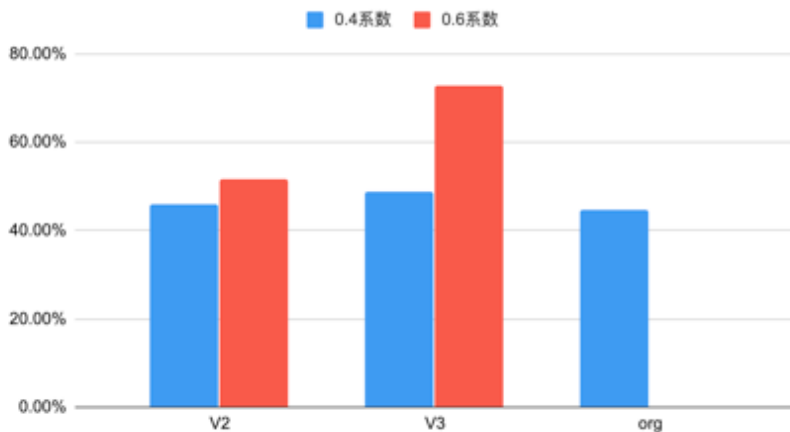
右：原始
左：码率压缩40%

码率下降，画质提升

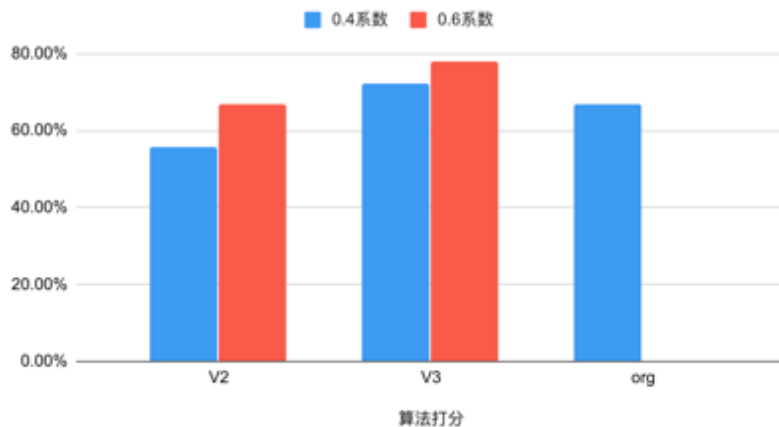


2020
北京

盲测打分



算法打分



直播评测：码率压缩40%，画质仍比源好

合理利用资源



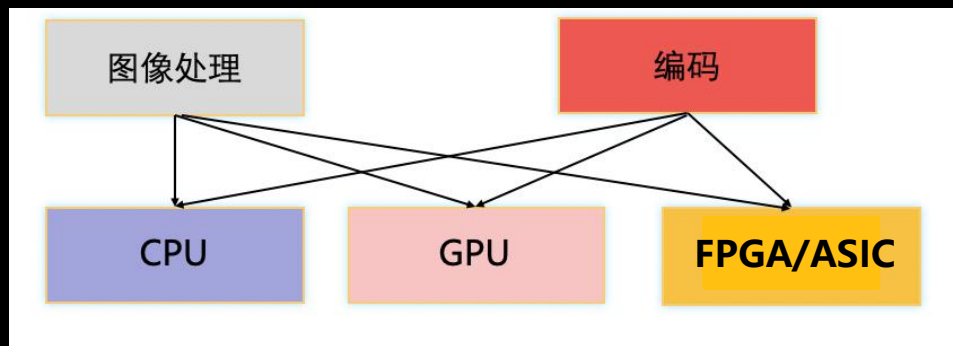
2020
北京

编码复杂度增高

图像处理算法需要较大算力

直播：实时、码率波动小

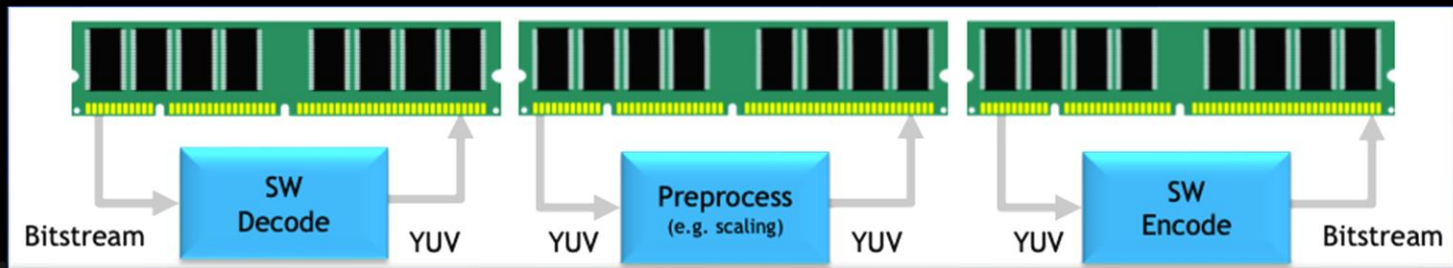
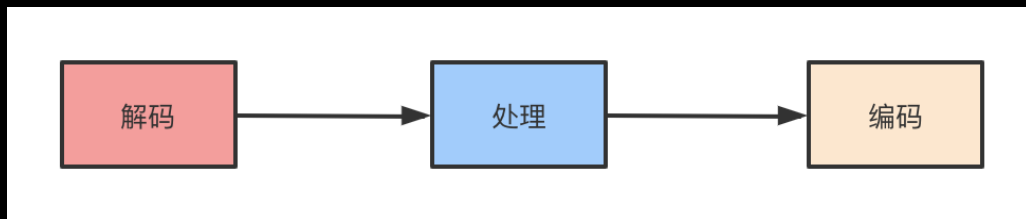
短视频：成本



软硬件适配



2020
北京



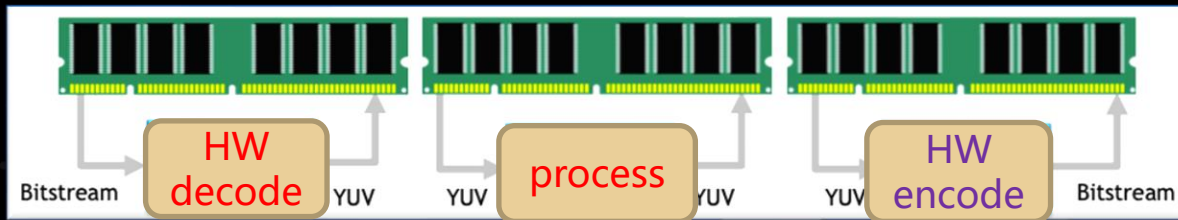
硬件支持：GPU/AISC 图像处理、专用硬件编解码器
加快处理速度，减少数据传输耗时

硬件编解码



2020
北京

- 硬件特点
 - 成本低
 - 可控性低，比如恒定质量编码，开始提供ROI编码的接口
 - 编码特性与软件不同
- 长远看
 - 独立硬件应对复杂度——处理、编解码 复杂度越来越高
 - 支持特性越来越多

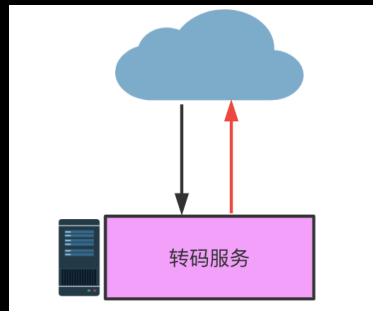


收益及损耗



2020
北京

- 直播---cpu编码
 - 卡顿率降低
 - 清晰度提高
 - 延时增加，主要为网络延时



- 短视频---混合编码
 - 成本相比只使用CPU编码降低50%
 - CPU编码：264，清晰率从26.2%提高到42.9%；265，清晰率从24.3%提高到42.5%
 - GPU编码：264，清晰率从26.2%提高到38.9%；265，清晰率从24.3%提高到39.6%

复杂现场美颜



2020
北京

现场娱乐节目的美颜



复杂现场美颜



2020
北京



图片去噪增强



2020
北京





2020
北京

客户端超分增强

PART 4

客户端超分增强



2020
北京



服务端:

1. 易于开发灰度实验
2. 扩展性好

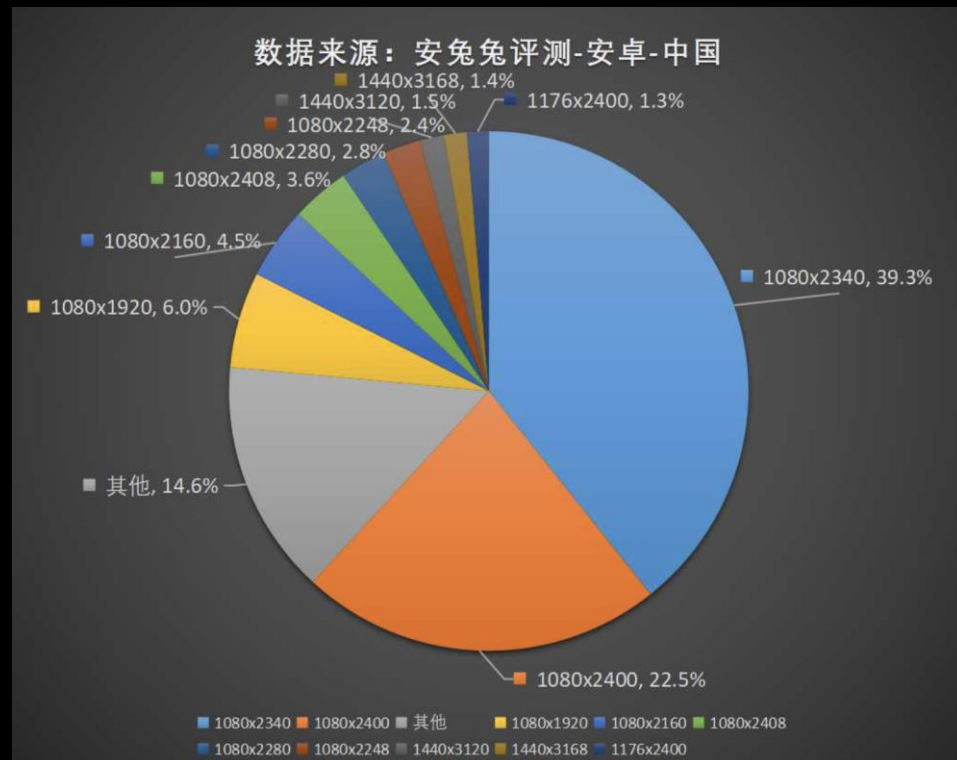
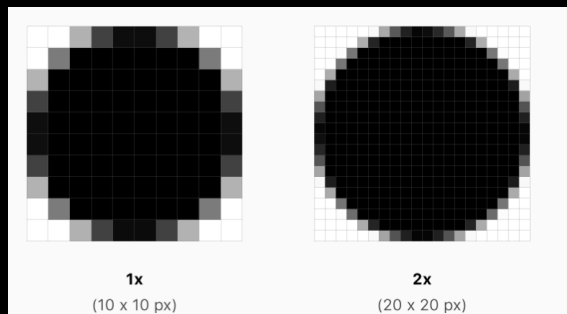
客户端:

1. 不引入额外延迟
2. 不引入额外成本

客户端超分增强



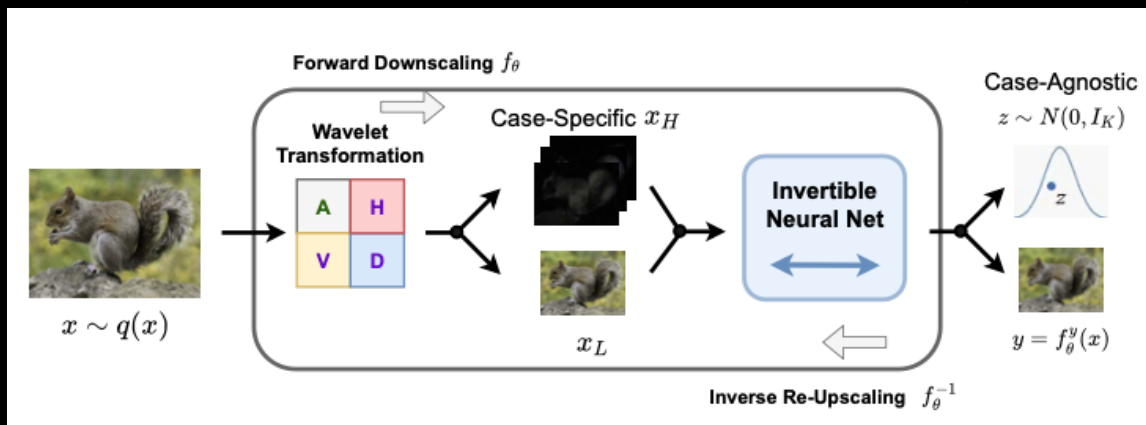
2020
北京



可逆图像缩放



2020
北京



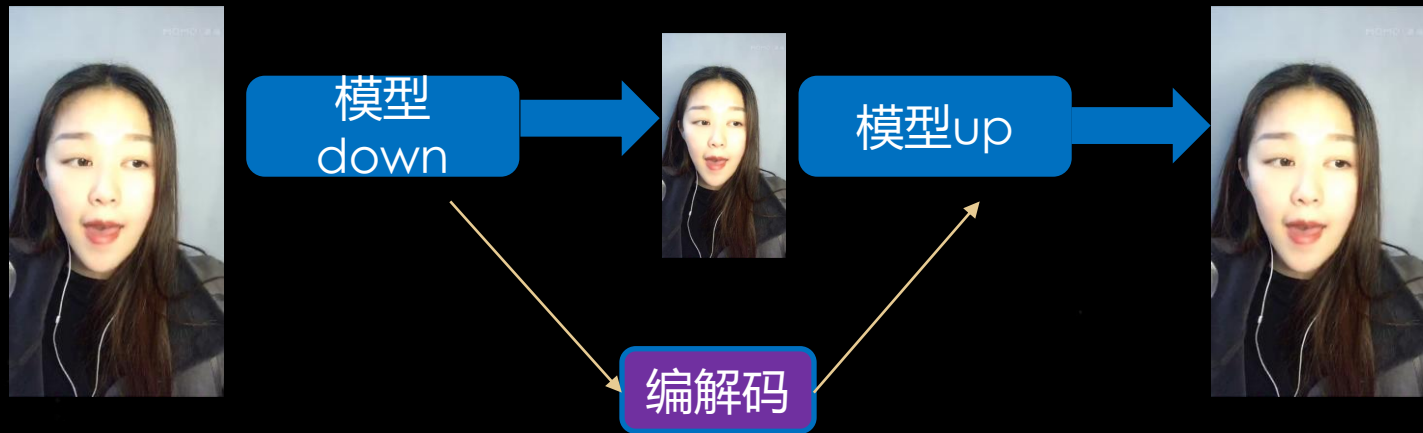
IRN,eccv2020

Downscaling & Upscaling	Scale	Param	Set5	Set14	BSD100	Urban100	DIV2K
Bicubic & Bicubic	2×	/	33.66 / 0.9299	30.24 / 0.8688	29.56 / 0.8431	26.88 / 0.8403	31.01 / 0.9393
Bicubic & SRCNN [17]	2×	57.3K	36.66 / 0.9542	32.45 / 0.9067	31.36 / 0.8879	29.50 / 0.8946	—
Bicubic & EDSR [36]	2×	40.7M	38.20 / 0.9606	34.02 / 0.9204	32.37 / 0.9018	33.10 / 0.9363	35.12 / 0.9699
Bicubic & RDN [60]	2×	22.1M	38.24 / 0.9614	34.01 / 0.9212	32.34 / 0.9017	32.89 / 0.9353	—
Bicubic & RCAN [59]	2×	15.4M	38.27 / 0.9614	34.12 / 0.9216	32.41 / 0.9027	33.34 / 0.9384	—
Bicubic & SAN [14]	2×	15.7M	38.31 / 0.9620	34.07 / 0.9213	32.42 / 0.9028	33.10 / 0.9370	—
TAD & TAU [26]	2×	—	38.46 / —	35.52 / —	36.68 / —	35.03 / —	39.01 / —
CNN-CR & CNN-SR [34]	2×	—	38.88 / —	35.40 / —	33.92 / —	33.68 / —	—
CAR & EDSR [49]	2×	51.1M	38.94 / 0.9658	35.61 / 0.9404	33.83 / 0.9262	35.24 / 0.9572	38.26 / 0.9599
IRN (ours)	2×	1.66M	43.99 / 0.9871	40.79 / 0.9778	41.32 / 0.9876	39.92 / 0.9865	44.32 / 0.9908

可逆图像缩放



2020
北京



超分增强效果



2020
北京



原画

增强

超分增强

总结



2020
北京

- 结合特性业务，进行评价和画质增强
- 结合服务端和客户端，系统性解决画质问题
- 算法、软硬件一体，合理利用资源

展望



2020
北京

技术进展

- 新的传感器：采集的图像更清晰
- 新的处理器：更强处理能力
- 新的视频压缩标准：更高压缩比
- 新的移动通信技术，5G，包括全链路的通信：高带宽、低延迟、大流量
- 新的图像处理方式
- . . .

技术趋势

- AI会赋能新技术发展
- 垂直、结合会是导致差异化主要途径
- 编解码标准短期较少融入深度学习能力
- AI使能编码压缩+质量提升，将会是一个长期方向



2020
北京

多媒体开启
MULTIMEDIA BRIDGE
TO A WORLD OF VISION

新视界

Thank you

