A Smart Reference Picture Resampling Approach for VVC

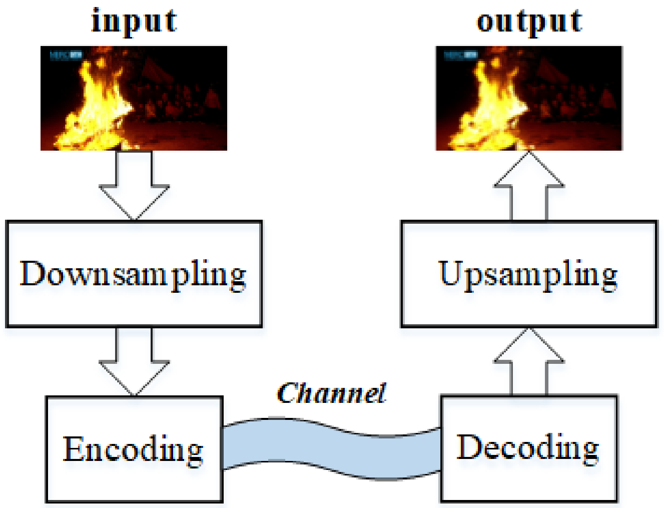
Tianliang Fu (PKU), Kai Zhang, Yue Li, Li Zhang (ByteDance)
Shanshe Wang, Siwei Ma (PKU)

Agenda

- Background
- Proposed method
- Simulation results
- Conclusion

Background-Reference Picture Resampling

- Resampling-based coding
  - Down-sampling before compression and up-sampling the decoded video for reconstruction
- Use case
  - Compressing high-resolution videos at low bitrate
  - Bitrate adaption in video telephony and conferencing
  - The frequent active speaker changes in the multi-party video conferencing
  - Fast start for streaming application
  - Adaptive stream switching in streaming application



Background-RPR in VVC

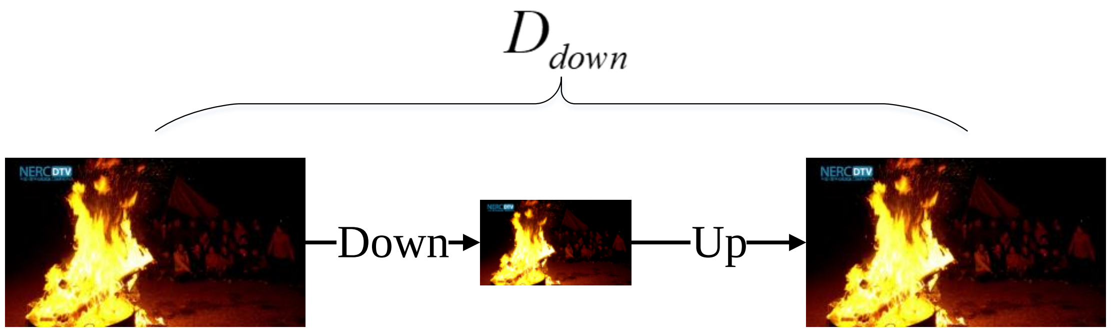
- Before VVC
  - Inserting an IDR or IRAP with a new resolution setting
  - Consuming more bits and time than an inter-coded picture
- RPR in VVC
  - Support the spatial resolution change without inserting an IDR or IRAP picture
  - The height and width of the current picture are signaled directly in PPS
  - The scaling ratios for horizontal and vertical directions are both arbitrary
- RPR as a coding tool
  - Multiple-pass encoded to enable a manual selection between low-resolution coding and full-resolution coding
  - A much higher encoding time and is not desirable in low-delay scenarios

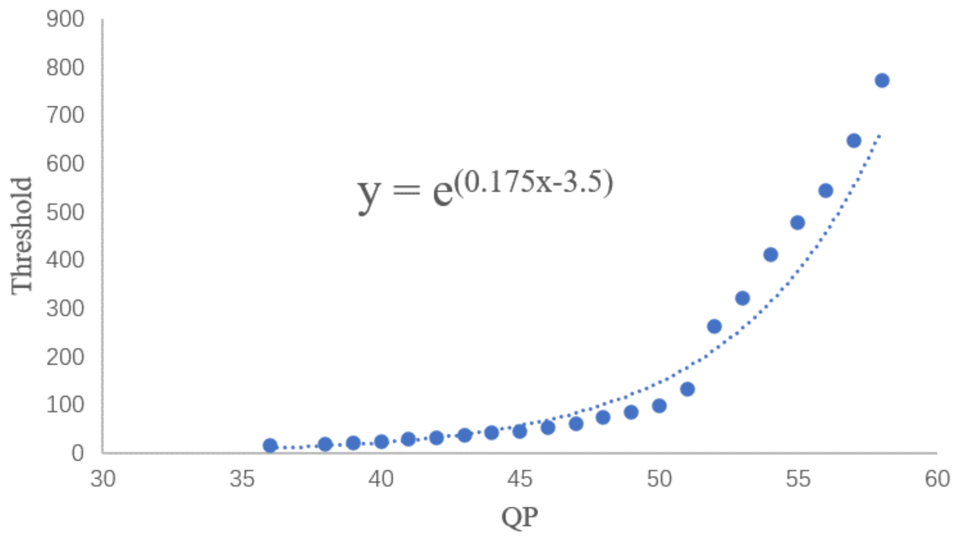
Smart RPR

- Adaptive decision on coding resolution

$$D_{down} < D_{thre}$$ ➡ Downsampling-based coding

$$D_{thre} = e^{a*QP-b}$$



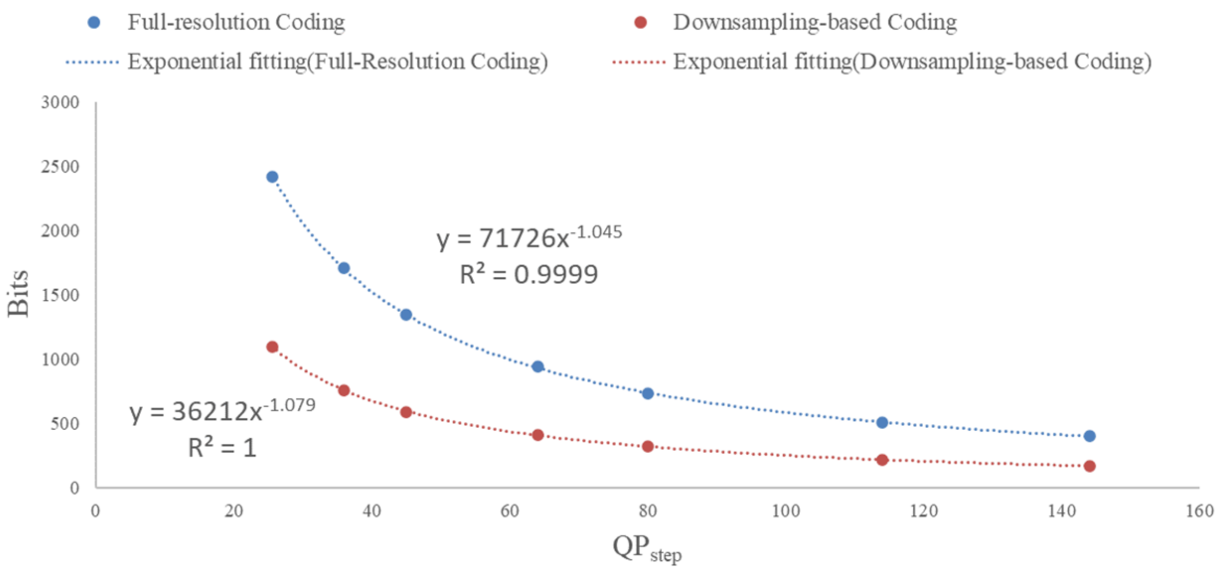


Smart RPR

- Coding parameters setting for Luma component
  - Align rate

$$R_{down} = R_{full}$$

$$R = \alpha \times QP_{step}^{\beta}$$



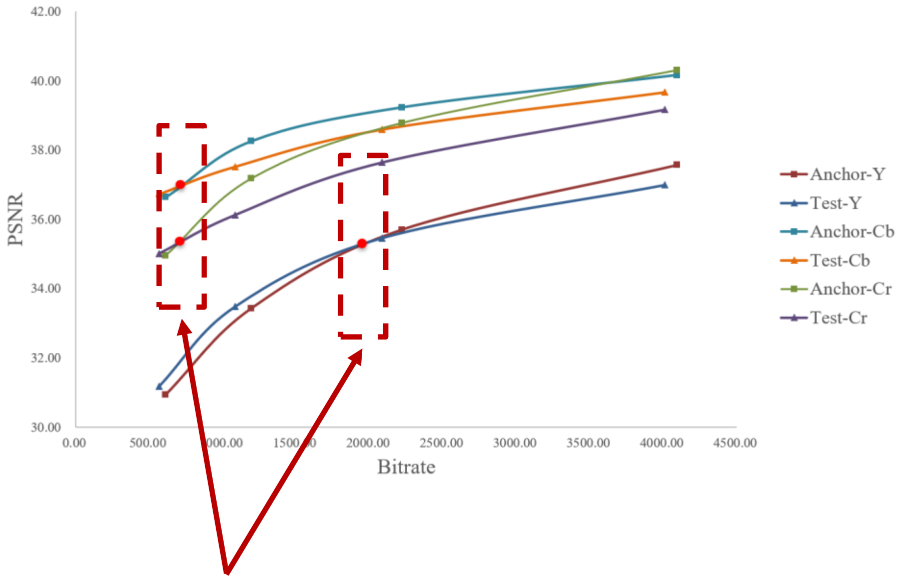
$$QP_{step} = 2^{\frac{QP-4}{6}}$$

$$\frac{QP_{step\_down}}{QP_{step\_full}} = \left(\frac{\alpha_{down}}{\alpha_{full}}\right)^{\frac{1}{\beta}}$$

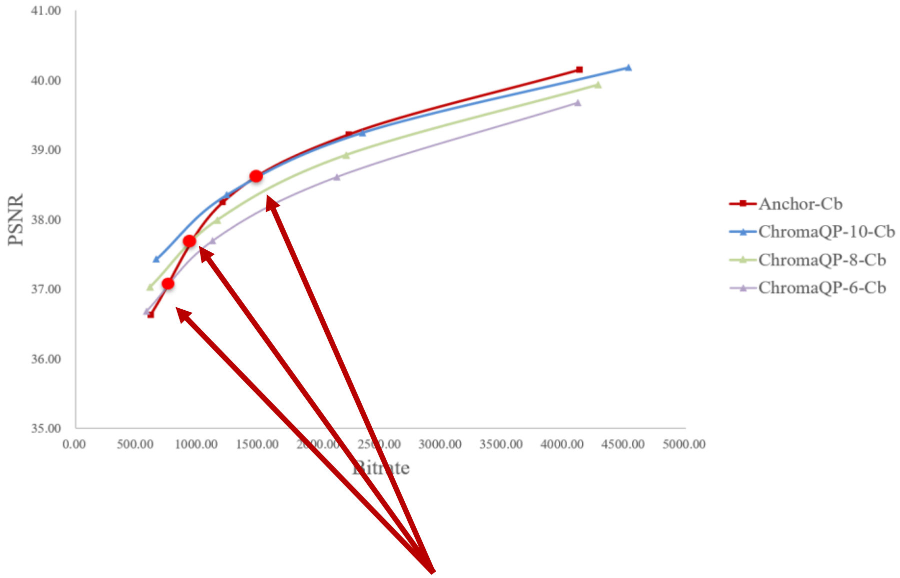
$$QP_{down} = QP_{full} - 6$$

Smart RPR

- Coding parameters setting for Chroma component
  - Eliminate the chroma loss



The critical bitrate of chroma components is lower than that of the luma component

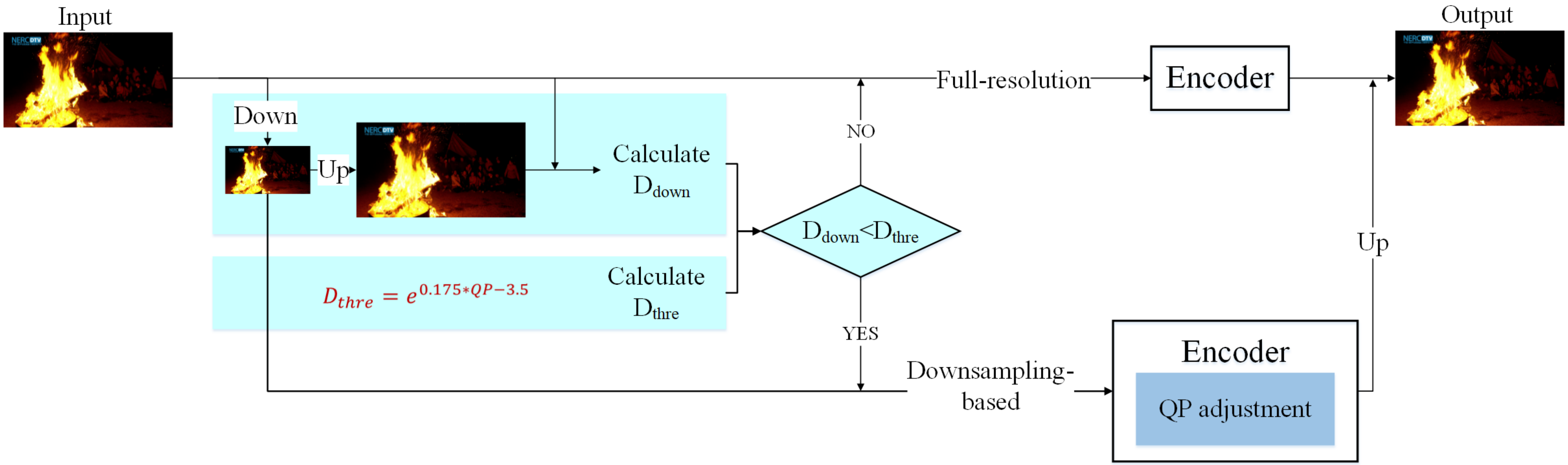


when the chroma QP offset increases from 6 to 10, the critical bitrate gradually gets larger

$$QP_{chroma\_down} = QP_{chroma\_full} - 9$$

Smart RPR

Framework



Simulation results

- Experiment configurations
  - Platform: VTM-12.0
  - Sequences: 4K

Table 1: The basic parameters of the test sequences

Class	Sequence	Width	Height	FrameRate	BitDepth
A1	*Tango2*	3840	2160	60	10
	FoodMarket4	3840	2160	60	10
	Campfire	3840	2160	30	10
A2	*CatRobot*	3840	2160	60	10
	DaylightRoad2	3840	2160	60	10
	ParkRunning3	3840	2160	50	10
Added	*MountainBay2*	3840	2160	30	10
	OberbaumSpree	3840	2160	60	10
	RaceNight	3840	2160	50	10
	TiergartenParkway	3840	2160	60	10

  - QPs: {32, 37, 42, 47}

Table 2: Performance of the proposed method in LDB

Sequence	Y-PSNR	U-PSNR	V-PSNR	EncT	DecT	$Delta_{Bitrate}$
Tango2	-7.40%	-16.94%	-22.26%	72%	48%	5.0%
FoodMarket4	-8.36%	-18.39%	-23.18%	73%	40%	2.0%
Campfire	2.37%	-0.58%	-1.46%	103%	84%	4.7%
CatRobot	-1.51%	0.46%	-2.24%	82%	67%	0.9%
DaylightRoad2	-0.99%	-4.25%	-5.08%	102%	84%	1.5%
ParkRunning3	-2.37%	32.46%	19.49%	79%	65%	4.2%
MountainBay2	-2.75%	-4.11%	-29.77%	69%	52%	6.0%
OberbaumSpree	-3.83%	-27.29%	-29.11%	65%	48%	5.6%
RaceNight	-2.32%	-11.28%	-12.13%	84%	70%	0.5%
TiergartenParkway	-0.08%	-2.93%	-2.44%	100%	91%	0.7%
Average	**-2.72%**	**-5.29%**	**-10.82%**	**83%**	**65%**	**3.1%**

Simulation results

- Subjective quality

original

VTM12.0 +
downsampling-based coding

VTM12.0 anchor

Conclusion

- We present a smart-RPR approach that enjoys improved compression efficiency and lower encoding/decoding complexity.
  - A frame is first down- and up-sampled without encoding and the resampling distortion is calculated by comparing the up-sampled frame with the original frame
  - The RPR for this frame will be enabled if the resampling distortion is less than a distortion threshold conditioning on QP.
  - The luma and chroma QPs of the downsampled frame are adjusted to avoid excessive bitrate fluctuation and to achieve a better coding performance.
- Smart RPR achieves about 2.72%, 5.29%, and 10.82% BD-rate savings on average for Y, Cb, and Cr components respectively
- Both encoding and decoding time savings are observed and the bitrate fluctuation is minor, i.e 3.1% on average

Thanks for your attention!

futl@pku.edu.cn