

JVET-AH0113

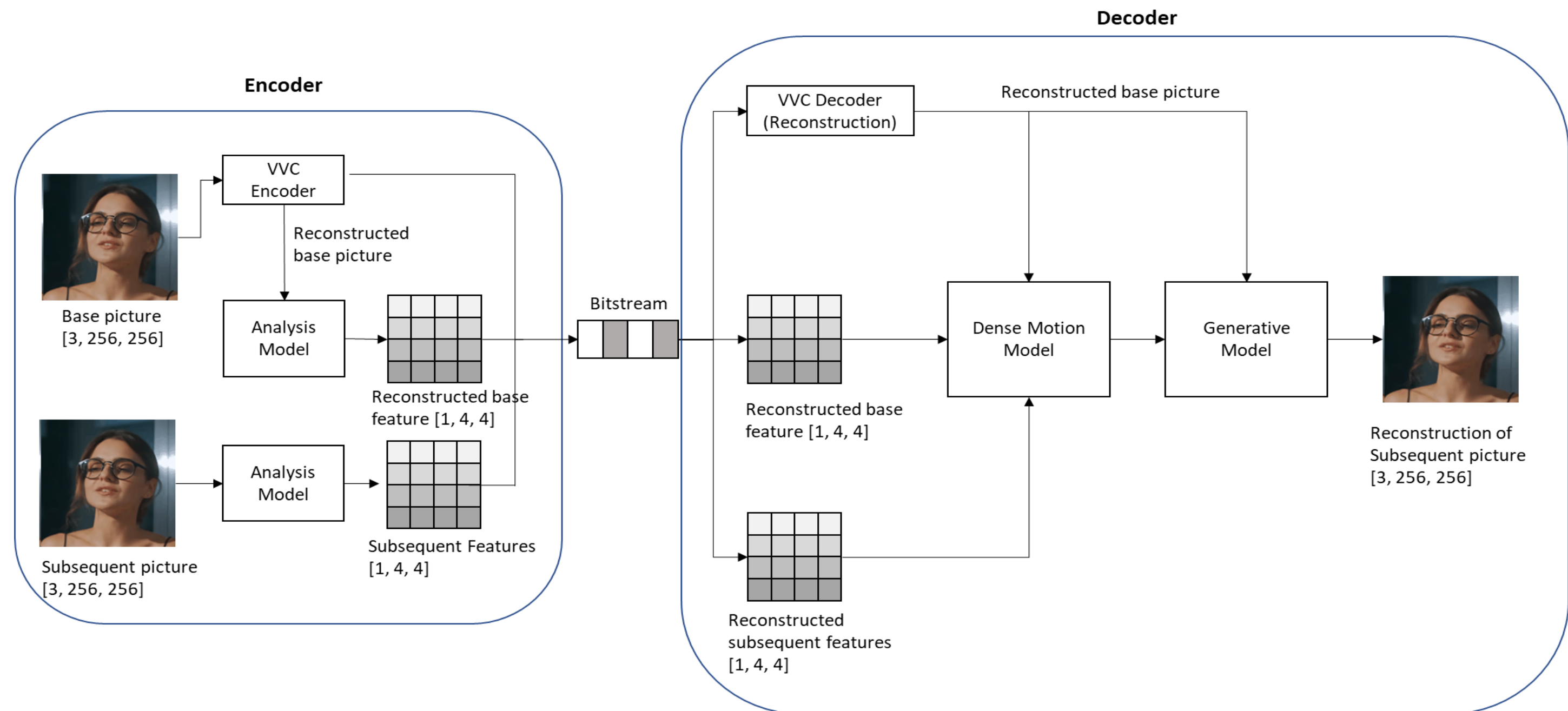
AHG16: Lightweight CFTE with Multi-Resolution support

Renjie Zou, Bolin Chen, Ru-Ling Liao, Jie Chen, Yan Ye

Alibaba group

Introduction

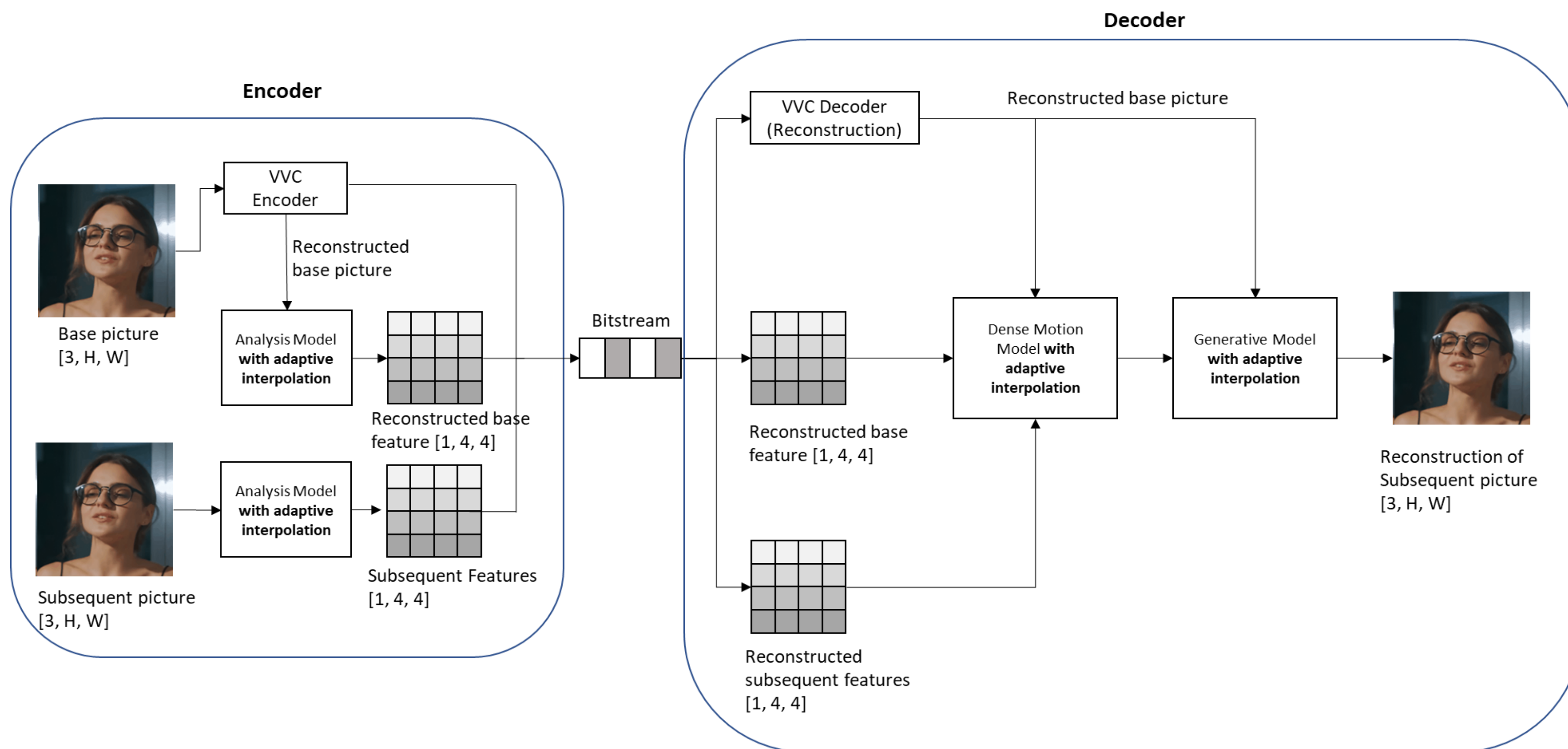
- The current AHG16 GFVC software only supports coding video sequences with a resolution of 256x256
- We propose to extend the lightweight CFTE method in JVET-AH0109 to support higher resolution sequences such as 512x512 and 640x640



Lightweight CFTE method proposed in JVET-AH0109

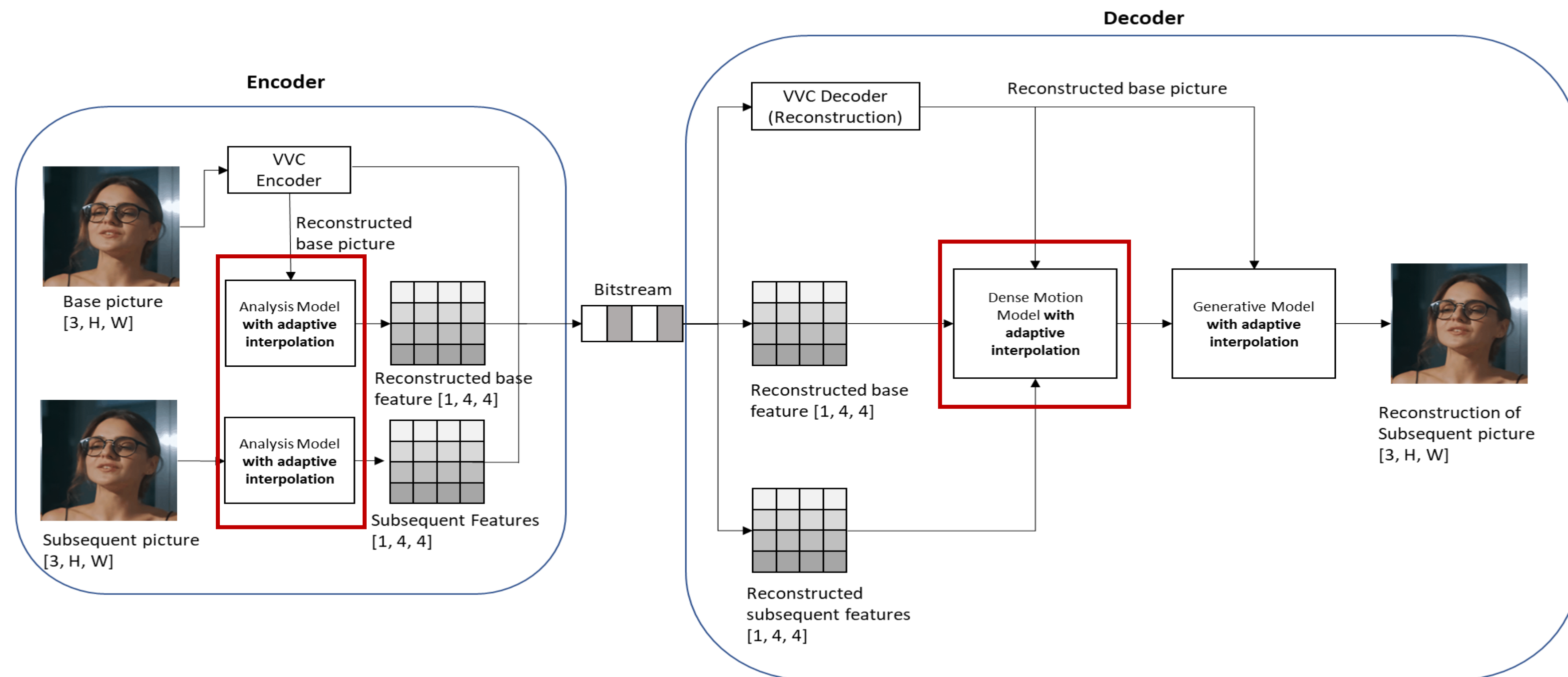
Proposed methods

- Adaptive interpolation is incorporated into analysis model, dense motion model, and generative model



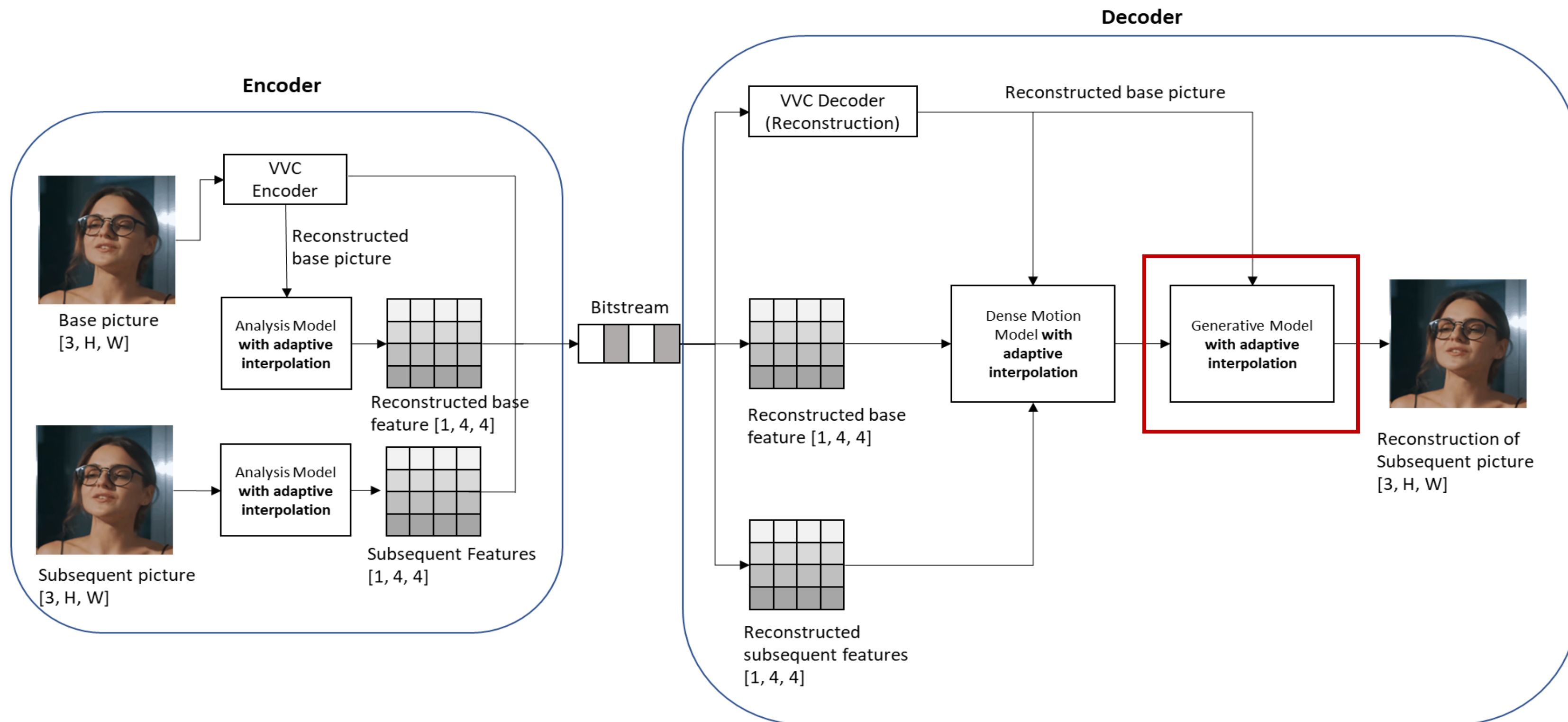
Modifications to analysis and dense motion models

- An adaptive down-sampling filter is employed to get intermediate output of a uniform size
- A gaussian blur filter followed by a nearest-neighbor down-sampling is used
- Down-sampling factor is set to 0.25 for 256x256, 0.125 for 512x512, 0.1 for 640x640



Modification to generative model

- An up-sampling operator is applied to the output of the dense motion model to upscale back to the respective target resolution
- Bilinear interpolation is used in up-sampling operator



Simulation results at resolution 256x256

- The proposed method is evaluated using the test conditions (sequence of resolution 256x256) and evaluation procedures specified in JVET-AG2035

	Proposed method at resolution 256x256					
	Over VTM-22.2					
	DISTS	LPIPS	RGB444_PSNR	RGB444_SSIM	YUV420_PSNR	YUV420_SSIM
Class A	-54.42%	-51.30%	0.00%	0.0%	0.0%	0.0%
Class B	-45.68%	-41.16%	1.14%	-0.9%	3.0%	0.2%
Overall	-50.05%	-46.23%	0.57%	-0.4%	1.5%	0.1%

* For resolution 512x512, 640x640, both bilinear and Lanczos interpolations are used to upscale test sets from a base resolution of 256x256

Simulation results at resolution 512x512

- Test sets obtained through bilinear interpolation

	Proposed method at resolution 512x512					
	Over VTM-22.2					
	DISTS	LPIPS	RGB444_PSNR	RGB444_SSIM	YUV420_PSNR	YUV420_SSIM
Class A	-69.13%	-66.55%	0.00%	0.0%	0.0%	0.0%
Class B	-57.45%	-46.72%	0.00%	0.0%	0.0%	-1.8%
Overall	-63.29%	-56.63%	0.00%	0.0%	0.0%	-0.9%

- Test sets obtained through Lanczos interpolation

	Proposed method at resolution 512x512					
	Over VTM-22.2					
	DISTS	LPIPS	RGB444_PSNR	RGB444_SSIM	YUV420_PSNR	YUV420_SSIM
Class A	-69.83%	-66.58%	0.00%	0.0%	0.0%	0.0%
Class B	-58.33%	-42.37%	0.00%	-2.0%	0.0%	-1.8%
Overall	-64.08%	-54.47%	0.00%	-1.0%	0.0%	-0.9%

Simulation results at resolution 640x640

- Test sets obtained through bilinear interpolation

	Proposed method at resolution 640x640					
	Over VTM-22.2					
	DISTS	LPIPS	RGB444_PSNR	RGB444_SSIM	YUV420_PSNR	YUV420_SSIM
Class A	-70.09%	-61.47%	0.00%	0.0%	0.0%	#VALUE!
Class B	-56.85%	-41.75%	0.00%	#VALUE!	0.0%	-2.0%
Overall	-63.47%	-51.61%	0.00%	#VALUE!	0.0%	#VALUE!

- Test sets obtained through Lanczos interpolation

	Proposed method at resolution 640x640					
	Over VTM-22.2					
	DISTS	LPIPS	RGB444_PSNR	RGB444_SSIM	YUV420_PSNR	YUV420_SSIM
Class A	-71.31%	-61.70%	0.00%	0.0%	0.0%	#VALUE!
Class B	-58.98%	-36.66%	0.00%	-1.8%	0.0%	-1.5%
Overall	-65.15%	-49.18%	0.00%	-0.9%	0.0%	#VALUE!

Visualization at resolution 512x512

Original sequence



Proposed method: 7.7648kbps



VVC: 8.5728kbps



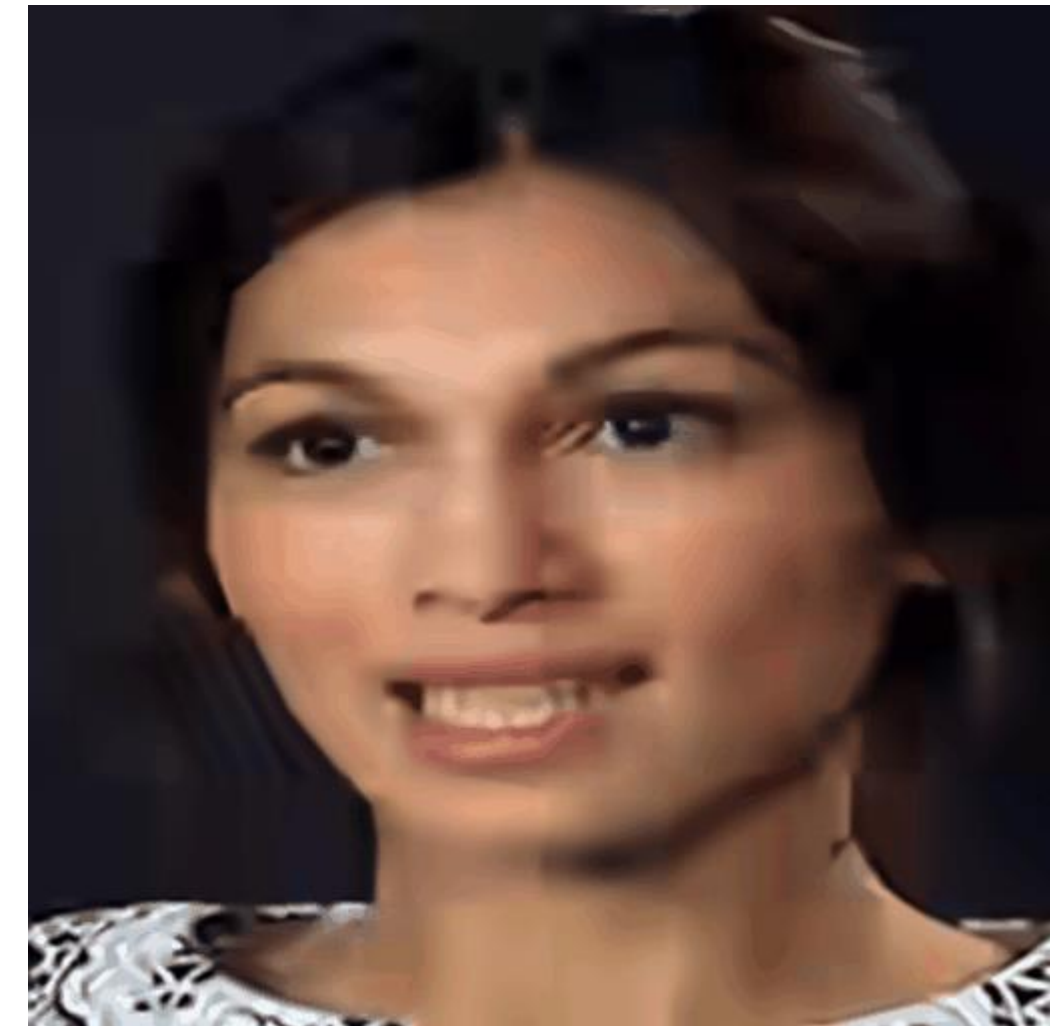
Original sequence



Proposed method: 8.3088kbps



VVC: 12.7976kbps



Visualization at resolution 640x640

Original sequence



Proposed method: 8.8984kbps



VVC: 12.5656kbps



* The original sequences are upsampled from resolution of 256x256 using Lanczos interpolation.

Conclusion

- The contribution extends the capabilities of the lightweight CFTE method by incorporating support for higher resolution video sequences up to 512x512 and 640x640
- The proposed multi-resolution lightweight CFTE method demonstrates the ability to efficiently process video sequences at multiple resolutions simultaneously
- We recommend to adopt this contribution to the AHG16 GFVC software

Thanks

