Reversing the Damage: A QP-Aware Transformer-Diffusion Approach for 8K Video Restoration under Codec Compression

– Supplementary Materials –

Ali Mollaahmadi Dehaghi
University of Calgary
Calgary, AB, Canada
ali.mollaahmadidehag@ucalgary.ca

Reza Razavi
Userful Corporation
Calgary, AB, Canada
reza.razavi@userful.com

Mohammad Moshirpour
University of California, Irvine
Irvine, CA, USA
mmoshirp@uci.edu

A. More Results

Table 1. PSNR Performance of *DiQP* on AV1 Codec with Various QPs

QP	Dataset	Output	Input
51	SEPE8K	42.171	37.544
	UVG	37.001	35.124
102	SEPE8K	40.478	36.774
	UVG	35.772	34.286
153	SEPE8K	38.762	35.816
	UVG	34.962	33.644
204	SEPE8K	36.818	34.564
	UVG	33.871	32.740

Table 2. SSIM Performance of *DiQP* on AV1 Codec with Various QPs

QP	Dataset	Output	Input
51	SEPE8K	0.9304	0.9105
	UVG	0.9148	0.8981
102	SEPE8K	0.9047	0.8886
	UVG	0.8924	0.8788
153	SEPE8K	0.8852	0.8708
	UVG	0.8831	0.8699
204	SEPE8K	0.8730	0.8584
	UVG	0.8731	0.8587

In this section, we delve deeper into the quantitative and qualitative results obtained from our model. Fig. 1 showcases frame 179 from sequence 26 of the SEPE8K dataset, along with our model's output. To better illustrate our model's performance, we highlight three specific crops from this frame: 1) the in-focus "bag" crop, demonstrating the model's ability to handle sharp details, 2) the out-of-focus "hat" crop with its intricate winter hat texture, testing the model's performance on challenging patterns, and 3) the out-of-focus "jacket" crop, further evaluating the model's ability to reconstruct detailed textures in less-than-ideal conditions. Our findings highlight that our model con-

Table 3. PSNR Performance of *DiQP* on HEVC Codec with Various QPs

QP	Dataset	Output	Input
12	SEPE8K	43.063	38.036
	UVG	38.396	36.018
33	SEPE8K	38.099	35.808
	UVG	34.882	33.590
42	SEPE8K	35.829	34.385
	UVG	33.660	32.665

Table 4. SSIM Performance of *DiQP* on HEVC Codec with Various QPs

QP	Dataset	Output	Input
12	SEPE8K	0.9456	0.9241
	UVG	0.9307	0.9133
33	SEPE8K	0.8842	0.8704
	UVG	0.8835	0.8684
42	SEPE8K	0.8695	0.8564
	UVG	0.8728	0.8571

Figure 1. Visual comparisons of different methods on sequence 26 of SEPE8K encoded with QP=255 on AV1

sistently delivers high-quality results in both in-focus and out-of-focus scenarios, effectively handling various textures and sharpness levels, unlike other models that struggle in such diverse situations.

To validate our model's robustness, we evaluated its performance under a range of compression levels (QPs). This included not only the most challenging highest QP scenario but also lower QPs to ensure comprehensive learning. As detailed in Tables 1, 2, 3, and 4, our model consistently demonstrates strong performance across various compression levels, from the least to the most compressed videos. Intuitively, our *DiQP* exhibits a notably stronger performance advantage at lower QPs, with a margin of approximately 5 dB in the SEPE8K dataset and around 2 dB in the UVG dataset. This shows a superior ability to leverage the additional information available in less compressed videos.