

# Appendix

## A. Architecture details

### A.1. I-P-frame codec

The details of the main autoencoder and the hyperprior in Image-AE, Flow-AE, and Residual-AE are shown in Figs. 14 and 15, respectively.

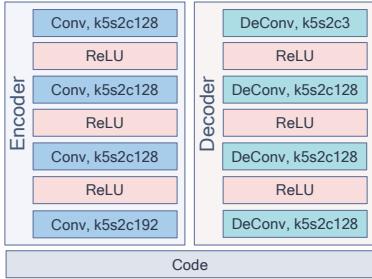


Figure 14. Main autoencoder details.  $k$ ,  $s$ , and  $c$  denote kernel size, stride, and the number of output channels, respectively.

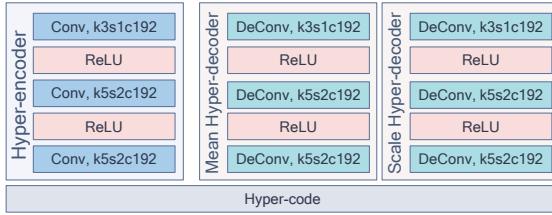


Figure 15. Hyperprior details.  $k$ ,  $s$ , and  $c$  denote kernel size, stride, and the number of output channels, respectively.

In our implementation of the hyperprior, we followed [5] which differs slightly from SSF [2] as pointed below:

- in the scale hyper-decoder, QReLU activations are replaced by ReLU and the last QReLU is removed. To have a lower bound on standard deviation values, we clamp the scale hyper-decoder output at 0.11,
- in both hyper-decoders, the last layer is implemented as a DeConv  $3 \times 3$  with stride 1 as opposed to DeConv  $5 \times 5$  with stride 2 in SSF,
- the hyper-encoder, the first layer is implemented as a Conv  $3 \times 3$  with stride 1 as opposed to Conv  $5 \times 5$  with stride 2 in SSF.

### A.2. Frame interpolation

In the frame interpolation component, FlowNet is a pre-trained PWC-Net [35] without modifications and RefineNet is a U-Net shown in Fig. 16.

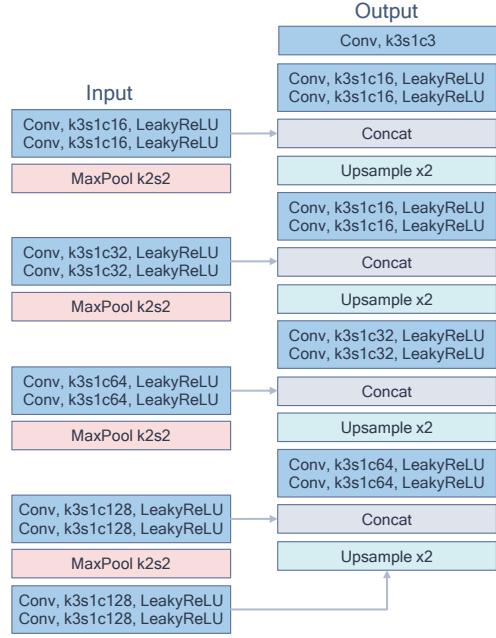


Figure 16. U-Net architecture details.  $k$ ,  $s$ , and  $c$  denote kernel size, stride, and the number of output channels, respectively.

## B. FFMPEG commands

We generated H.264 and H.265 baselines using FFMPEG. The command that we used to run FFMPEG with all the default configurations is as follows:

```
ffmpeg -pix_fmt yuv420p -s [W]x[H]
-r [FR] -i [IN].yuv -c:v libx[ENC]
-b:v -crf [CRF] [OUT].mkv
```

and the command that we used to run FFMPEG with GoP=12 is as follows:

```
ffmpeg -pix_fmt yuv420p -s [W]x[H]
-r [FR] -i [IN].yuv -c:v libx[ENC]
-b:v -crf [CRF] -x[ENC]-params
"keyint=[GOP]:min-keyint=[GOP]:verbose=1"
[OUT].mkv
```

where the values in brackets represent the encoder parameters as follows:  $H$  and  $W$  are the frame dimensions,  $FR$  is the frame rate,  $ENC$  is the encoder type ( $\times 264$  or  $\times 265$ ),  $GOP$  is the GoP size (12),  $INPUT$  and  $OUTPUT$  are the input and the output filenames, respectively, and  $CRF$  controls the bit-rate (We tried  $CRF = \{9, 12, 15, 18, 21, 24, 27, 30\}$ ).

In order to measure MS-SSIM and PSNR in the RGB color space, we saved all video frames as PNG files using FFMPEG using the following commands for YUV and MKV files:

```
ffmpeg -pix_fmt yuv420p -s [W]x[H]
-i [IN].yuv %9d.png
ffmpeg [IN].mkv %9d.png
```

## C. Extended results

**Bjøntegaard delta rate (BD-rate) comparison:** We report the BD-rate [6] gains for a scenario where both H. 264 and H. 265 are configured to use all the default parameters as opposed to the results for GoP=12 reported in 4.4 (see Table 2).

Dataset	PSNR BD-rate gain (%)			MS-SSIM BD-rate gain (%)		
	H. 265 (MSE)	SSF (MSE)	B-EPIC (MSE)	H. 265 (MS-SSIM)	SSF (MS-SSIM)	B-EPIC (MS-SSIM)
UVG	<b>-29.02</b>	11.05	-26.93	-24.28	-11.55	<b>-25.99</b>
MCL-JCV	<b>-20.06</b>	3.77	-17.27	-16.76	-29.69	<b>-40.76</b>
HEVC-B	<b>-22.44</b>	6.69	-18.21	-16.54	-32.36	<b>-41.80</b>
HEVC-C	<b>-12.96</b>	91.98	43.38	-8.21	-5.24	<b>-24.66</b>
HEVC-D	<b>-7.64</b>	149.80	72.16	-3.13	5.71	<b>-27.71</b>
HEVC-E	<b>-27.31</b>	34.28	-18.12	-27.24	-4.36	<b>-30.77</b>
HEVC-Avg	<b>-17.59</b>	70.69	19.80	-13.78	-9.05	<b>-31.23</b>

Table 2. Average BD-rate gain versus H. 264 (with FFmpeg default parameters) on different datasets.

### Qualitative results:

In Fig. 17 we show the intermediate visualizations as well as the detailed rate-distortion results across a GoP of seven frames for both B-EPIC and SSF. Figures 18 and 19 show qualitative comparisons of B-EPIC and SSF for the first GoP of two videos.

**Rate-distortion results:** We report the per-video performance of our B-EPIC(MSE) and B-EPIC( MS-SSIM) models on the UVG [38], MCL-JCV [39], and HEVC [7] datasets in Tables 3 through 14.

Video	Performance across models - PSNR (dB) vs Rate (bits-per-pixel)															
	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR
Beauty	0.956	37.92	0.553	36.8	0.229	35.53	0.061	34.51	0.029	34.16	0.018	33.9	0.012	33.54	0.009	33.1
Bosphorus	0.235	42.71	0.12	41.41	0.067	40.09	0.04	38.8	0.025	37.34	0.016	36.0	0.013	34.45	0.01	33.1
HoneyBee	0.415	39.87	0.112	38.56	0.036	37.66	0.017	36.68	0.01	35.49	0.007	34.4	0.007	33.01	0.006	31.68
Jockey	0.419	40.79	0.193	39.73	0.11	39.0	0.071	38.29	0.053	37.42	0.034	36.59	0.026	35.44	0.02	34.18
ReadySetGo	0.455	41.38	0.28	40.14	0.178	38.75	0.119	37.26	0.084	35.65	0.054	34.12	0.039	32.46	0.031	30.77
ShakeNDry	0.599	40.13	0.292	38.7	0.159	37.42	0.094	36.17	0.058	34.81	0.035	33.46	0.02	31.92	0.013	30.68
YachtRide	0.474	41.87	0.316	40.64	0.21	39.28	0.139	37.75	0.091	36.1	0.057	34.46	0.036	32.6	0.024	30.96
Average	0.508	40.67	0.267	39.43	0.141	38.25	0.077	37.07	0.05	35.85	0.032	34.71	0.022	33.35	0.016	32.07

Table 3. Detailed rate-distortion performance of B-EPIC(MSE) on the UVG dataset.

Video	Performance across models - MS-SSIM vs Rate (bits-per-pixel)							
	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM
Beauty	0.842	0.98	0.607	0.973	0.377	0.961	0.21	0.946
Bosphorus	0.37	0.995	0.203	0.993	0.112	0.99	0.065	0.986
HoneyBee	0.47	0.992	0.253	0.989	0.105	0.985	0.054	0.981
Jockey	0.589	0.991	0.365	0.987	0.207	0.981	0.115	0.976
ReadySetGo	0.413	0.995	0.245	0.993	0.146	0.99	0.095	0.986
ShakeNDry	0.573	0.992	0.362	0.988	0.207	0.983	0.135	0.976
YachtRide	0.47	0.995	0.296	0.993	0.19	0.99	0.121	0.986
Average	0.532	0.991	0.333	0.988	0.192	0.983	0.114	0.977

Table 4. Detailed rate-distortion performance of B-EPIC(MS-SSIM) on the UVG dataset.

Video	Performance across models - PSNR (dB) vs Rate (bits-per-pixel)															
	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR				
videoSRC01	0.371	40.88	0.093	39.54	0.036	38.96	0.017	38.49	0.011	37.97	0.007	37.39	0.007	36.65	0.006	35.88
videoSRC02	0.19	44.26	0.124	43.36	0.086	42.36	0.061	41.23	0.046	39.94	0.032	38.64	0.025	37.09	0.02	35.36
videoSRC03	0.328	41.5	0.166	40.29	0.09	39.08	0.05	37.78	0.029	36.31	0.018	35.07	0.013	33.62	0.01	32.07
videoSRC04	0.858	41.28	0.648	39.92	0.481	38.46	0.339	36.71	0.231	35.07	0.152	33.55	0.093	31.88	0.056	30.36
videoSRC05	0.905	37.94	0.467	36.77	0.266	35.89	0.169	35.0	0.113	34.01	0.074	33.0	0.051	31.68	0.038	30.25
videoSRC06	1.374	33.86	1.021	33.42	0.647	32.51	0.169	31.15	0.016	30.63	0.009	30.54	0.007	30.43	0.006	30.3
videoSRC07	0.895	37.22	0.55	36.31	0.253	35.18	0.111	34.33	0.061	33.64	0.038	32.99	0.025	32.16	0.017	31.37
videoSRC08	0.62	38.96	0.205	37.69	0.1	37.1	0.054	36.5	0.031	35.8	0.019	35.03	0.014	34.14	0.011	33.2
videoSRC09	1.107	38.1	0.744	36.83	0.446	35.4	0.233	33.74	0.128	32.05	0.078	30.42	0.05	28.59	0.035	26.83
videoSRC10	0.899	40.18	0.612	38.8	0.417	37.32	0.286	35.81	0.202	34.29	0.132	32.77	0.093	31.08	0.084	29.35
videoSRC11	0.274	44.58	0.186	43.39	0.126	42.07	0.086	40.65	0.061	39.1	0.039	37.58	0.028	36.0	0.022	34.33
videoSRC12	0.338	41.55	0.19	39.79	0.11	38.04	0.061	36.18	0.034	34.47	0.02	32.94	0.014	31.55	0.011	30.18
videoSRC13	0.781	39.28	0.544	38.11	0.347	36.65	0.176	34.74	0.067	32.58	0.024	30.35	0.012	27.8	0.009	25.8
videoSRC14	0.544	40.75	0.284	39.46	0.166	38.37	0.107	37.26	0.072	36.05	0.048	34.8	0.035	33.38	0.026	31.89
videoSRC15	1.0	38.48	0.63	37.19	0.329	35.69	0.139	33.99	0.054	32.41	0.03	31.18	0.02	29.87	0.014	28.55
videoSRC16	0.232	41.51	0.09	40.68	0.05	40.14	0.03	39.49	0.02	38.75	0.013	37.94	0.011	36.98	0.008	35.81
videoSRC17	0.525	40.34	0.244	39.15	0.152	38.29	0.097	37.36	0.065	36.28	0.043	35.12	0.029	33.73	0.021	32.49
videoSRC18	0.316	39.3	0.191	37.82	0.12	36.89	0.079	35.19	0.055	33.49	0.034	31.83	0.024	29.98	0.017	28.45
videoSRC19	0.517	41.81	0.333	40.35	0.219	38.82	0.145	37.24	0.096	35.61	0.061	34.05	0.041	32.51	0.03	31.07
videoSRC20	0.409	42.06	0.284	40.84	0.202	39.6	0.14	38.08	0.1	36.57	0.068	35.13	0.05	33.66	0.038	32.13
videoSRC21	0.172	45.04	0.111	44.25	0.075	43.46	0.052	42.54	0.036	41.5	0.025	40.38	0.019	39.1	0.015	37.78
videoSRC22	0.525	42.84	0.348	41.5	0.242	40.23	0.172	38.91	0.124	37.49	0.088	35.95	0.063	34.29	0.046	32.58
videoSRC23	0.317	43.45	0.186	42.05	0.113	40.61	0.072	39.13	0.047	37.57	0.031	36.04	0.022	34.47	0.016	32.88
videoSRC24	0.333	41.48	0.216	40.51	0.146	39.68	0.095	38.53	0.061	37.07	0.039	35.76	0.027	33.95	0.02	32.22
videoSRC25	1.084	36.19	0.809	35.27	0.594	34.45	0.411	33.06	0.27	31.49	0.174	30.02	0.107	28.09	0.067	26.31
videoSRC26	0.243	42.55	0.138	41.65	0.093	40.85	0.065	39.94	0.047	38.89	0.034	37.8	0.026	36.48	0.021	34.9
videoSRC27	0.447	42.4	0.299	40.89	0.2	39.28	0.132	37.59	0.086	35.82	0.054	34.12	0.034	32.3	0.023	30.7
videoSRC28	0.178	42.72	0.085	41.74	0.046	40.74	0.027	39.63	0.018	38.44	0.012	37.16	0.011	35.8	0.009	34.39
videoSRC29	0.071	45.3	0.035	44.59	0.021	43.96	0.013	43.28	0.009	42.5	0.007	41.71	0.006	40.84	0.005	39.79
videoSRC30	0.286	40.02	0.144	39.26	0.054	38.36	0.022	37.52	0.013	36.72	0.009	35.94	0.008	35.06	0.007	34.08
Average	0.538	40.86	0.333	39.71	0.208	38.62	0.12	37.37	0.073	36.08	0.047	34.84	0.032	33.44	0.024	32.04

Table 5. Detailed rate-distortion performance of B-EPIC(MSE) on the MCL-JCV dataset.

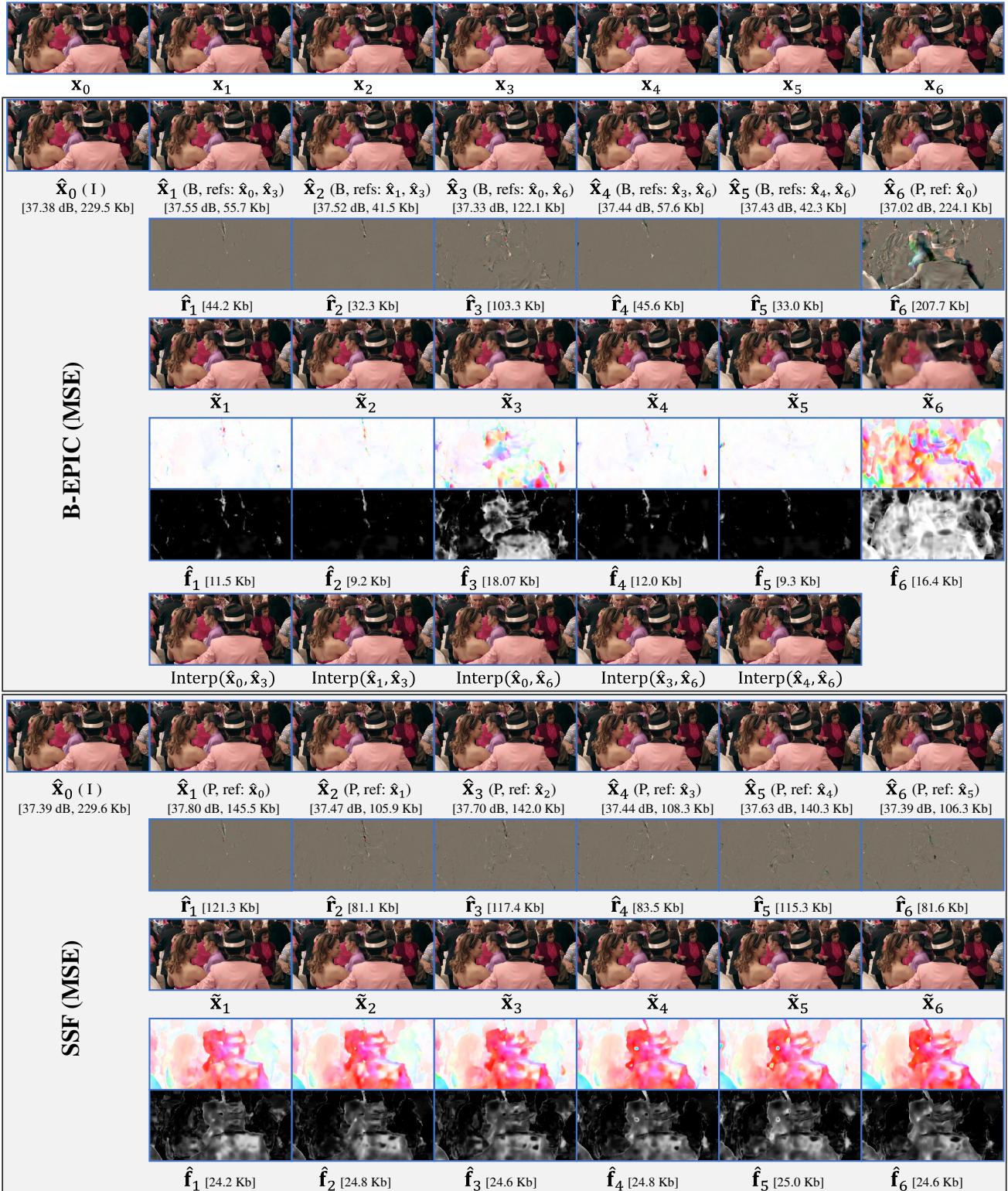


Figure 17. Qualitative results for the first GoP (GoP=7) of Tango video from Netflix Tango in Netflix El Fuente [44] (resolution  $1024 \times 2048$ , zoom in for more details). Both B-EPIC and SSF models are trained on MSE with  $\beta = 0.0016$ .  $\mathbf{x}_0$  is an intra frame with similar performance on B-EPIC and SSF. The other six frames are inter frames (P or B). SSF yields consistent bit-rates across the inter frames due to similar level of details in the optical flow and the residuals. That is mainly because all the inter frames are P-frames where the immediate previous decoded frame is used as reference. In B-EPIC, the inter frames are mostly B-frames where the distances to the references are quite variable. As a result, it delivers different bit-rates across the inter frames. The average results for the inter-frames for B-EPIC and SSF are as follows: PSNR: 37.37dB vs 37.57dB, bit-rate: 90.5 Kb vs 124.7 kb, residual bit-rate: 77.7 Kb vs 100.0 Kb, flow bit-rate: 12.8 Kb vs 24.7 Kb. Here, the bit-rate values are obtained by multiplying Rate (bits-per-pixel) by frame resolution.

[Video produced by Netflix, with CC BY-NC-ND 4.0 license: [https://media.xiph.org/video/derf/ElFuente/Netflix\\_Tango\\_Copyright.txt](https://media.xiph.org/video/derf/ElFuente/Netflix_Tango_Copyright.txt)]

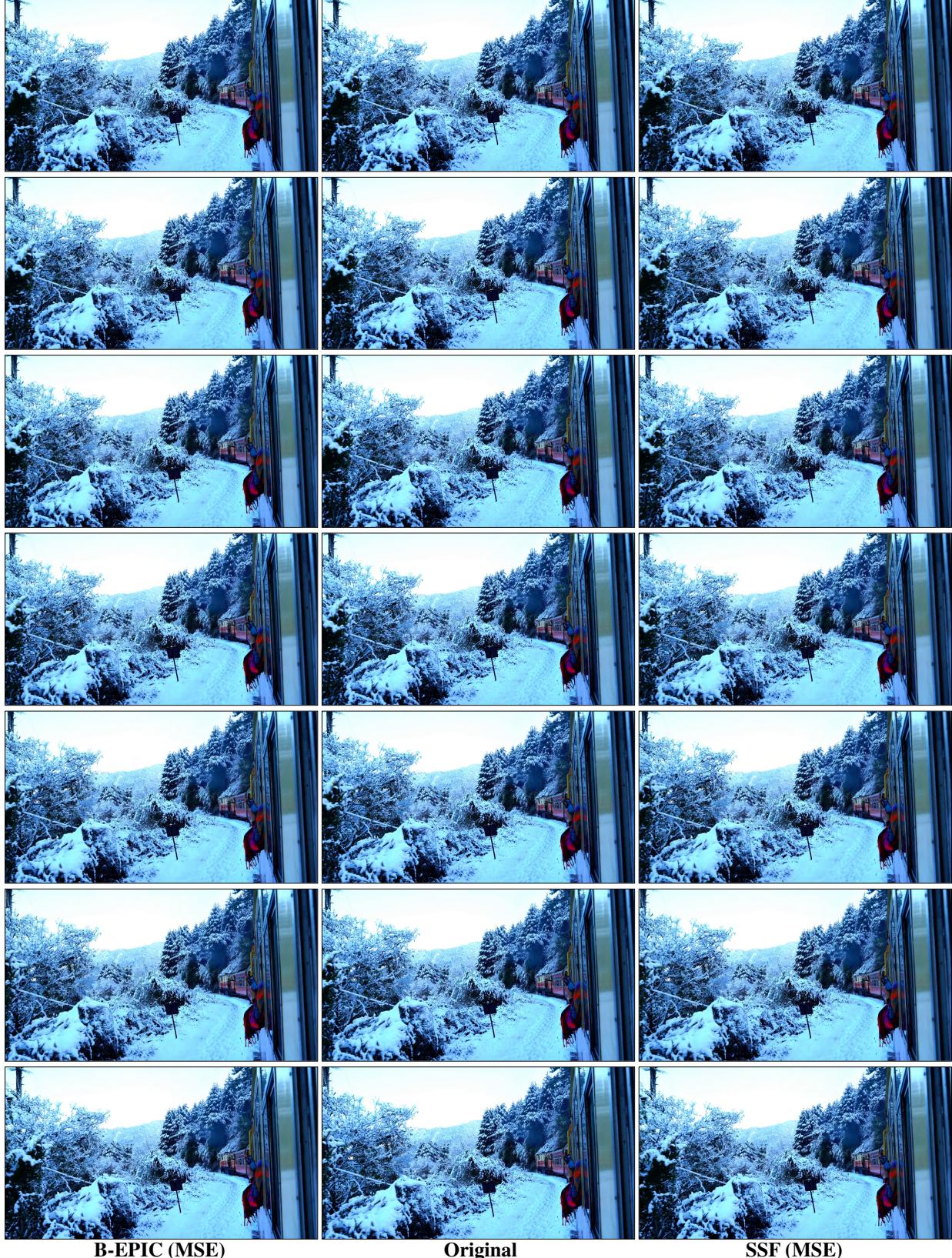


Figure 18. Qualitative comparisons of B-EPIC and SSF. The columns represent a sequence of 7 frames (top to bottom) compressed using B-EPIC and SSF as well as the uncompressed sequence where the GoP structures for B-EPIC and SSF are `IBBBBBP` and `IPPPPPP`, respectively. The average (size, PSNR) for B-EPIC and SSF are (43.47KB, 28.80dB) and (51.39KB, 28.85dB). While B-EPIC generates similar PSNR and visual quality, it consumes 15.4% less bits compared to SSF. [Video obtained from Pexels.]

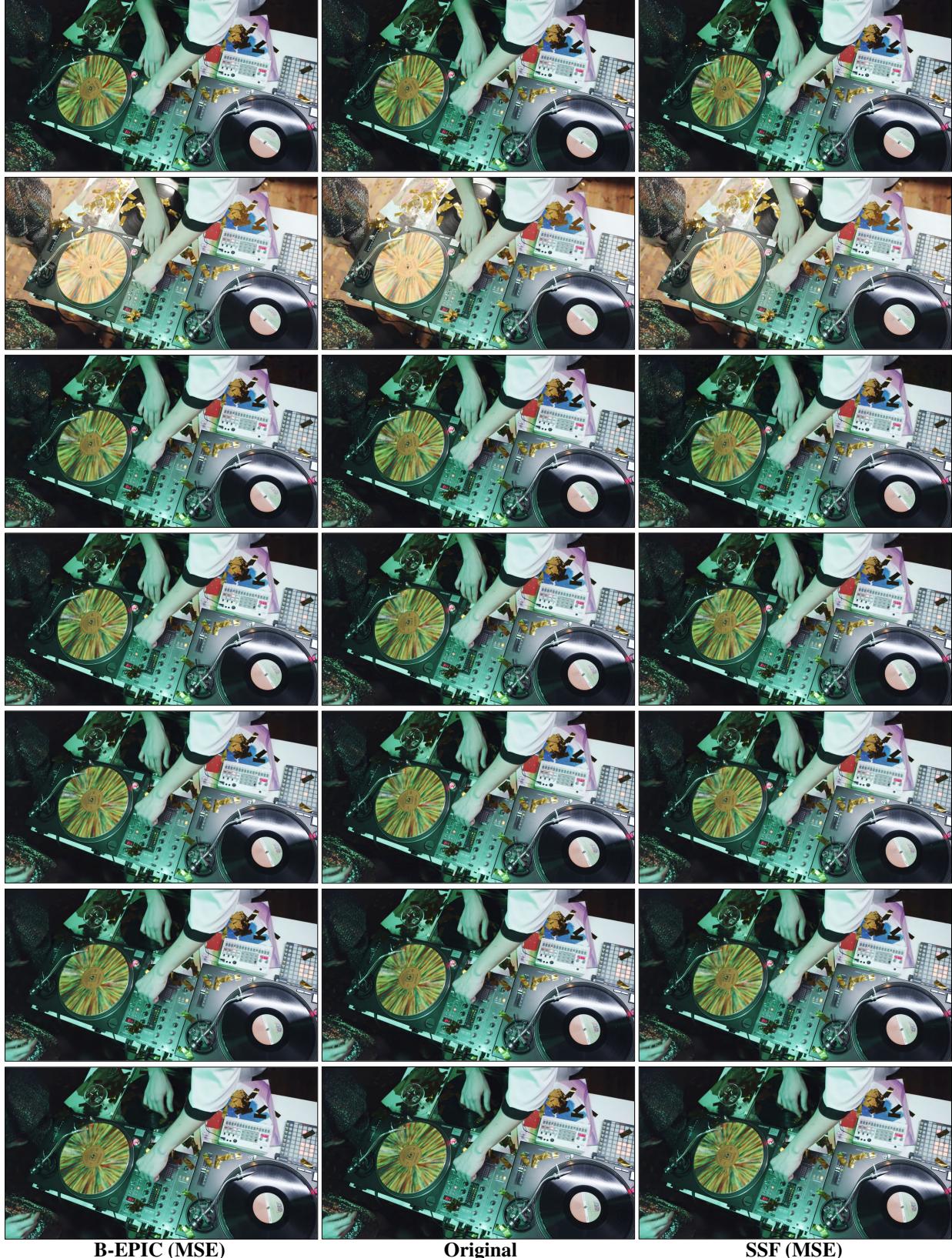


Figure 19. Qualitative comparisons of B-EPIC and SSF. The columns represent a sequence of 7 frames (top to bottom) compressed using B-EPIC and SSF as well as the uncompressed sequence where the GoP structures for B-EPIC and SSF are IBBBBBP and IPPPPP, respectively. The average (size, PSNR) for B-EPIC and SSF are (23.84KB, 33.81dB) and (27.01KB, 33.94dB). While B-EPIC generates similar PSNR and visual quality, it consumes 11.7% less bits compared to SSF. [Video obtained from Pexels.]

Performance across models - MS-SSIM vs Rate (bits-per-pixel)												
Video	Rate MS-SSIM											
videoSRC01	0.567	0.991	0.347	0.987	0.178	0.982	0.078	0.976	0.021	0.97	0.009	0.966
videoSRC02	0.312	0.996	0.177	0.995	0.106	0.993	0.068	0.99	0.047	0.987	0.031	0.982
videoSRC03	0.467	0.994	0.273	0.992	0.136	0.988	0.064	0.983	0.029	0.976	0.016	0.968
videoSRC04	0.634	0.994	0.45	0.991	0.318	0.986	0.223	0.98	0.154	0.971	0.1	0.957
videoSRC05	0.642	0.988	0.415	0.984	0.249	0.979	0.144	0.972	0.085	0.963	0.051	0.952
videoSRC06	0.9	0.95	0.652	0.938	0.433	0.919	0.269	0.896	0.108	0.865	0.016	0.839
videoSRC07	0.756	0.981	0.512	0.975	0.307	0.966	0.158	0.955	0.08	0.944	0.043	0.932
videoSRC08	0.708	0.986	0.487	0.982	0.272	0.975	0.117	0.967	0.041	0.96	0.018	0.953
videoSRC09	0.62	0.994	0.371	0.991	0.211	0.985	0.125	0.978	0.067	0.966	0.041	0.948
videoSRC10	0.58	0.995	0.388	0.993	0.263	0.99	0.171	0.985	0.114	0.979	0.071	0.971
videoSRC11	0.29	0.996	0.175	0.995	0.114	0.993	0.077	0.991	0.054	0.987	0.034	0.982
videoSRC12	0.259	0.995	0.146	0.992	0.086	0.988	0.066	0.983	0.032	0.973	0.019	0.961
videoSRC13	0.338	0.995	0.158	0.991	0.072	0.986	0.036	0.979	0.011	0.971	0.006	0.959
videoSRC14	0.526	0.994	0.326	0.991	0.191	0.987	0.111	0.983	0.067	0.976	0.041	0.968
videoSRC15	0.727	0.992	0.501	0.988	0.315	0.982	0.174	0.972	0.075	0.956	0.039	0.939
videoSRC16	0.461	0.992	0.264	0.988	0.127	0.984	0.055	0.98	0.026	0.976	0.015	0.971
videoSRC17	0.604	0.992	0.385	0.989	0.227	0.985	0.133	0.979	0.078	0.972	0.051	0.963
videoSRC18	0.215	0.994	0.136	0.991	0.089	0.986	0.066	0.98	0.041	0.969	0.026	0.954
videoSRC19	0.496	0.995	0.326	0.993	0.213	0.989	0.139	0.984	0.089	0.976	0.055	0.965
videoSRC20	0.326	0.996	0.228	0.994	0.16	0.991	0.116	0.987	0.079	0.98	0.053	0.97
videoSRC21	0.261	0.995	0.146	0.992	0.085	0.99	0.053	0.988	0.036	0.985	0.024	0.981
videoSRC22	0.48	0.995	0.323	0.993	0.216	0.99	0.148	0.986	0.106	0.981	0.074	0.974
videoSRC23	0.334	0.996	0.196	0.995	0.113	0.992	0.068	0.989	0.04	0.984	0.027	0.978
videoSRC24	0.303	0.996	0.182	0.995	0.112	0.992	0.071	0.99	0.045	0.986	0.028	0.98
videoSRC25	0.637	0.994	0.447	0.991	0.302	0.987	0.207	0.98	0.135	0.968	0.085	0.951
videoSRC26	0.388	0.994	0.216	0.991	0.115	0.988	0.069	0.985	0.045	0.981	0.031	0.977
videoSRC27	0.374	0.996	0.238	0.994	0.158	0.992	0.107	0.988	0.066	0.982	0.043	0.973
videoSRC28	0.272	0.994	0.128	0.992	0.055	0.99	0.027	0.988	0.015	0.985	0.01	0.981
videoSRC29	0.19	0.995	0.089	0.993	0.042	0.991	0.02	0.989	0.012	0.987	0.007	0.984
videoSRC30	0.294	0.989	0.138	0.985	0.048	0.981	0.021	0.976	0.009	0.971	0.006	0.965
Average	0.465	0.992	0.294	0.989	0.177	0.984	0.106	0.979	0.06	0.971	0.036	0.961
	0.025	0.949									0.017	0.93

Table 6. Detailed rate-distortion performance of B-EPIC(PSNR) on the MCL-JCV dataset.

Performance across models - PSNR (dB) vs Rate (bits-per-pixel)												
Video	Rate PSNR											
BQTerrace	1.18	36.28	0.799	35.4	0.414	34.13	0.185	32.9	0.089	31.61	0.046	30.32
BasketballDrive	0.858	37.98	0.411	36.76	0.219	35.85	0.13	34.93	0.083	33.93	0.053	32.91
Cactus	0.986	36.74	0.525	35.69	0.239	34.64	0.108	33.51	0.06	32.39	0.036	31.27
Kimono	0.55	39.96	0.247	38.69	0.138	37.7	0.082	36.67	0.052	35.49	0.034	34.32
ParkScene	0.805	38.02	0.445	36.73	0.241	35.39	0.122	33.83	0.061	32.2	0.034	30.78
Average	0.876	37.8	0.486	36.66	0.25	35.54	0.125	34.37	0.069	33.13	0.041	31.92
	0.026	28.75	0.018	27.15								

Table 7. Detailed rate-distortion performance of B-EPIC(MSE) on the HEVC class-B dataset.

Performance across models - MS-SSIM vs Rate (bits-per-pixel)												
Video	Rate MS-SSIM											
BQTerrace	0.635	0.987	0.413	0.983	0.238	0.977	0.12	0.968	0.049	0.956	0.026	0.94
BasketballDrive	0.608	0.989	0.381	0.985	0.22	0.98	0.121	0.974	0.065	0.964	0.038	0.952
Cactus	0.7	0.987	0.452	0.982	0.243	0.975	0.122	0.965	0.048	0.952	0.028	0.939
Kimono	0.587	0.992	0.371	0.988	0.211	0.983	0.119	0.977	0.063	0.968	0.04	0.958
ParkScene	0.628	0.991	0.416	0.986	0.244	0.98	0.139	0.971	0.064	0.956	0.036	0.941
average	0.632	0.989	0.407	0.985	0.231	0.979	0.124	0.971	0.058	0.959	0.033	0.945
	0.02	0.921	0.014	0.892								

Table 8. Detailed rate-distortion performance of B-EPIC(PSNR) on the HEVC class-B dataset.

Performance across models - PSNR (dB) vs Rate (bits-per-pixel)												
Video	Rate PSNR											
BQMall	0.741	38.11	0.436	36.97	0.269	35.83	0.168	34.33	0.105	32.65	0.063	31.06
BasketballDrill	0.68	36.6	0.392	35.48	0.257	34.4	0.167	33.06	0.106	31.63	0.066	30.35
PartyScene	1.182	33.32	0.799	32.6	0.546	31.72	0.354	30.27	0.202	28.38	0.107	26.61
RaceHorses	1.175	36.33	0.834	35.44	0.589	34.43	0.395	32.96	0.249	31.27	0.145	29.7
Average	0.945	36.09	0.615	35.12	0.415	34.09	0.271	32.65	0.165	30.98	0.095	29.43
	0.053	27.59	0.033	25.93								

Table 9. Detailed rate-distortion performance of B-EPIC(MSE) on the HEVC class-C dataset.

Performance across models - MS-SSIM vs Rate (bits-per-pixel)												
Video	Rate MS-SSIM											
BQMall	0.439	0.993	0.235	0.989	0.13	0.985	0.077	0.978	0.045	0.968	0.025	0.953
BasketballDrill	0.451	0.991	0.275	0.987	0.162	0.98	0.096	0.968	0.055	0.951	0.033	0.928
PartyScene	0.474	0.99	0.288	0.986	0.165	0.979	0.098	0.965	0.053	0.942	0.03	0.904
RaceHorses	0.758	0.991	0.524	0.987	0.35	0.981	0.231	0.973	0.138	0.957	0.075	0.932
average	0.531	0.991	0.33	0.987	0.202	0.981	0.125	0.971	0.073	0.954	0.041	0.93
	0.026	0.895	0.017	0.849								

Table 10. Detailed rate-distortion performance of B-EPIC(PSNR) on the HEVC class-C dataset.

Video	Performance across models - PSNR (dB) vs Rate (bits-per-pixel)									
	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR
BQSquare	1.249	34.09	0.87	33.36	0.593	32.21	0.391	30.58	0.217	28.49
BasketballPass	0.883	36.99	0.623	35.85	0.449	34.74	0.308	33.21	0.197	31.45
BlowingBubbles	1.124	33.82	0.729	32.92	0.468	31.96	0.281	30.46	0.154	28.63
RaceHorses	1.253	36.63	0.903	35.55	0.648	34.3	0.434	32.68	0.268	30.81
Average	1.127	35.38	0.781	34.42	0.539	33.31	0.354	31.73	0.209	29.85
									0.116	28.12
									0.06	26.19
									0.036	24.44

Table 11. Detailed rate-distortion performance of B-EPIC(MSE) on the HEVC class-D dataset.

Video	Performance across models - MS-SSIM vs Rate (bits-per-pixel)									
	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM
BQSquare	0.4	0.991	0.228	0.987	0.121	0.981	0.066	0.972	0.029	0.952
BasketballPass	0.477	0.993	0.311	0.99	0.208	0.984	0.137	0.975	0.084	0.961
BlowingBubbles	0.393	0.989	0.22	0.984	0.117	0.975	0.066	0.96	0.035	0.936
RaceHorses	0.697	0.992	0.474	0.989	0.317	0.983	0.206	0.973	0.125	0.956
average	0.492	0.992	0.308	0.987	0.191	0.981	0.119	0.97	0.068	0.951
									0.036	0.921
									0.023	0.885
									0.015	0.835

Table 12. Detailed rate-distortion performance of B-EPIC(MS-SSIM) on the HEVC class-D dataset.

Video	Performance across models - PSNR (dB) vs Rate (bits-per-pixel)									
	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR
Vidyo1	0.234	42.17	0.108	40.84	0.052	39.65	0.029	38.44	0.018	37.06
Vidyo3	0.243	42.48	0.125	41.17	0.066	39.81	0.037	38.46	0.022	36.78
Vidyo4	0.241	42.61	0.118	41.17	0.058	39.73	0.032	38.36	0.019	36.85
Average	0.239	42.42	0.117	41.06	0.059	39.73	0.033	38.42	0.02	36.89
									0.013	35.39
									0.01	33.58
									0.009	31.87

Table 13. Detailed rate-distortion performance of B-EPIC(MSE) on the HEVC class-E dataset.

Video	Performance across models - MS-SSIM vs Rate (bits-per-pixel)									
	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM	Rate	MS-SSIM
Vidyo1	0.298	0.994	0.154	0.992	0.07	0.988	0.032	0.984	0.013	0.979
Vidyo3	0.254	0.995	0.112	0.992	0.047	0.989	0.026	0.985	0.012	0.979
Vidyo4	0.243	0.995	0.118	0.992	0.054	0.989	0.031	0.985	0.014	0.979
average	0.265	0.995	0.128	0.992	0.057	0.989	0.029	0.985	0.013	0.979
									0.009	0.971
									0.009	0.959
									0.008	0.946

Table 14. Detailed rate-distortion performance of B-EPIC(MS-SSIM) on the HEVC class-E dataset.