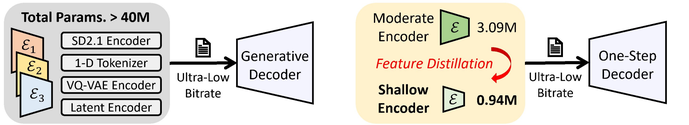
Ultra-Low Bitrate Perceptual Image Compression with Shallow Encoder

Tianyu Zhang [1,2], Dong Liu [1], Chang Wen Chen [2]
[1] University of Science and Technology of China, [2] The Hong Kong Polytechnic University

CVPR JUNE 3-7, 2026 DENVER COLORADO

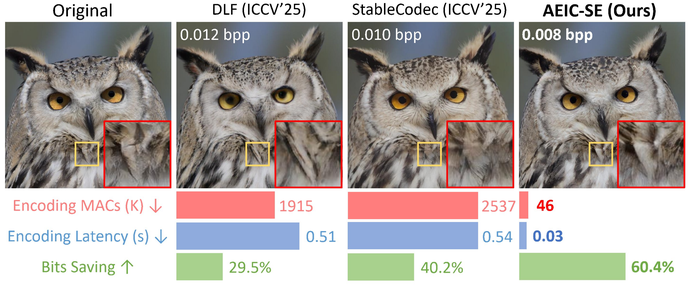
Github

Overview



(a) Extreme image compression with multiple complex encoders

(b) **A**symmetric **E**xtreme **I**mage **C**ompression with **S**hallow **E**ncoder **(Ours)**



(c) Comparison with existing extreme image compression methods

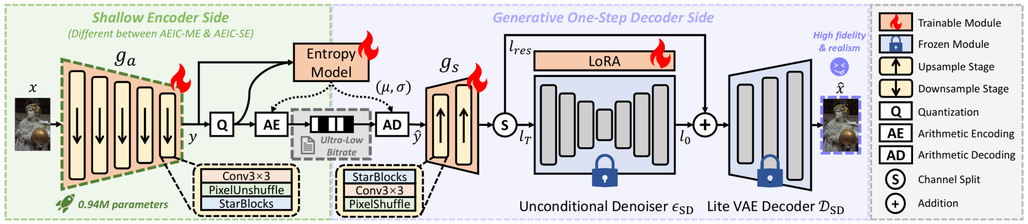
TL;DR:

- Ultra-low bitrate image compression (<0.05bpp) is increasingly critical for **bandwidth-constrained and computation-limited encoding scenarios** such as edge devices, whereas mainstream methods relies on multiple complex encoders for generative latent-space compression and generation.
- We demonstrate that ultra-low bitrate perceptual image compression allows for shallow encoders to reach SOTA performance, and propose **Asymmetric Extreme Image Compression (AEIC)** framework that pursues simultaneously encoding simplicity and decoding quality at extreme bitrates.

Contributions

1. We provide a **theoretical and empirical analysis** revealing the potential of applying shallow and low-complexity encoders for ultra-low bitrate compression.
2. We propose **asymmetric extreme image compression (AEIC) composed of lightweight encoders and a one-step diffusion decoder. A dual-side feature distillation strategy is introduced** to enhance our shallow encoder variant (AEIC-SE) with efficient knowledge transfer.
3. Specifically, AEIC-SE obtains:
   - ✓ **State-of-the-art perceptual performance at ultra-low bitrates** in terms of LPIPS, DISTS, FID, and KID, while preserving strong fidelity.
   - ✓ **Exceptional encoding efficiency for 35.8 FPS@1080P** on 1x RTX 4090D.
   - ✓ **Competitive decoding speed** compared to existing advanced methods.

Methodology



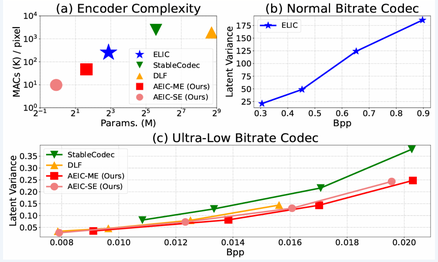
Pipeline of **A**symmetric **E**xtreme **I**mage **C**ompression with **S**hallow **E**ncoder **(AEIC-SE)**

Exploring Encoder with Two Variants

Model	g_a (each downsample stage)		Entropy Model	
	Depth	Dimension	Depth	Dimension
AEIC-ME	2	(64, 128, 192, 256, 320)	4	960
AEIC-SE	1	(32, 64, 128, 192, 256)	3	512

Encoder-side model configurations

Potential of Shallow Encoder at Low Bitrates



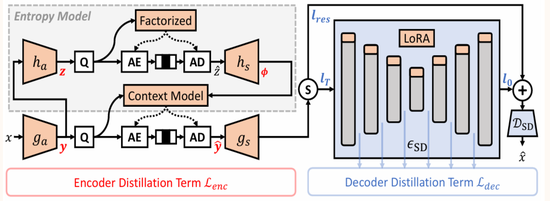
Encoder complexity & Latent variance at different bitrates

- Latent variance diminishes sharply as bitrate decreases, allowing shallow encoders to express this data range at ultra-low bitrates.
- A small latent variance in the continuous domain restricts the range of probably sampled values, which is analogous to a discrete codebook with fewer elements (allows low-complexity encoding).

Analysis on Implicit Bitrate Pruning

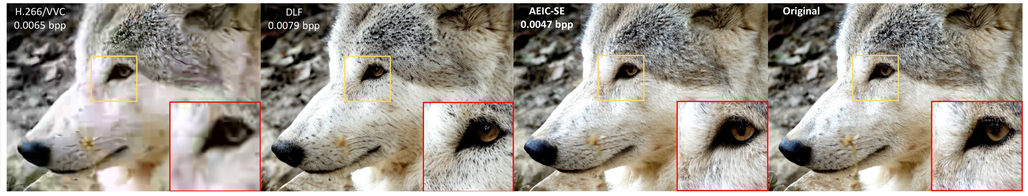
Bitrate and reconstruction evolution of AEIC-ME under the two-stage training

Dual-Side Feature Distillation



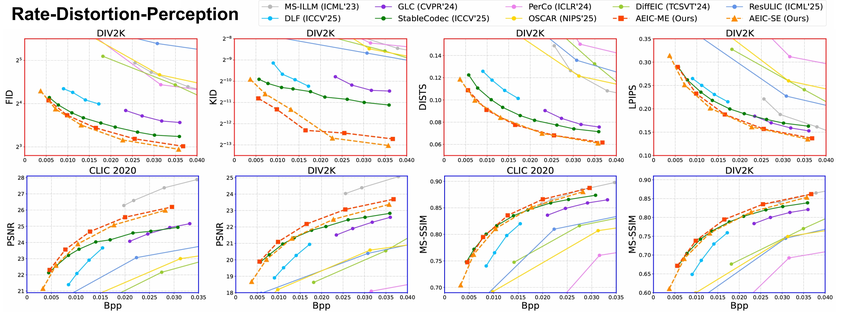
Feature distillation terms from AEIC-ME to AEIC SE

- Due to the limited capacity of the shallow encoder, training AEIC-SE from scratch results in a significant performance drop.
- We propose to transfer knowledge from AEIC-ME to AEIC-SE, enhancing the shallow encoder's representational ability.
- We design two separate objectives, supervising the encoder in the 1st stage and the decoder finetuning in the 2nd stage.

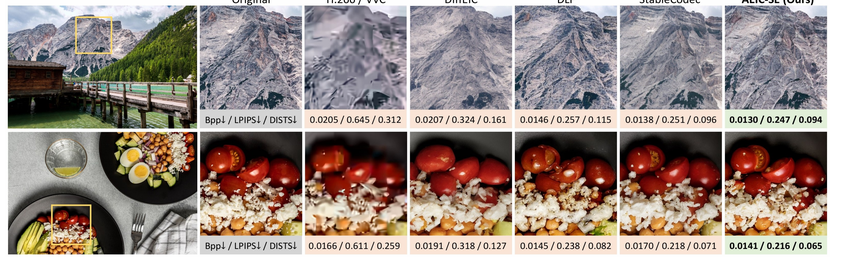


Experiments

Rate-Distortion-Perception



Visual Comparison



Computational Complexity

Method	Steps	Encoding		Decoding	
		Params.	MACs	Params.	MACs
ELIC [20]	-	33.38	346.03	33.38	590.823
EVC-Small [18]	-	**11.64**	71.12	**11.96**	**71.416**
PerCo [10]	20	>1B	2666.47	>1B	$> 10^4$
DiffEIC [38]	50	73.85	2253.69	>1B	$> 10^4$
DLF [62]	-	437.35	1915.35	561.63	4160.99
StableCodec [70]	1	102.23	2537.51	985.27	6201.66
AEIC-ME (Ours)	1	55.91	204.26	951.47	2884.88
AEIC-SE (Ours)	1	16.10	**46.02**	913.66	2762.22

Ablation Study

Method (Sp. Comp. Ratio)	BD-rate (↓%) on Kodak			
	PSNR	MS-SSIM	LPIPS	DISTS
StableCodec (64×)	0	0	0	0
AEIC-ME (64×)	+14.30	-4.54	-6.97	-6.95
AEIC-ME (32×)	-2.21	-4.85	**-13.67**	**-24.91**
AEIC-ME (16×)	**-5.53**	**-8.00**	-1.91	-9.03

Model	Distillation Term		BD-rate (↓%) on DIV2K		
	$\mathcal{L}_{enc}$	$\mathcal{L}_{dec}$	LPIPS	DISTS	FID
AEIC-ME	-	-	0	0	0
AEIC-SE	-	-	+8.47	+23.75	+22.10
	✓	-	+3.92	+7.68	+4.75
	✓	✓	**+0.60**	**+2.55**	**+2.98**

Encoding & Decoding Speed

Type	Method	Steps	Encoding (ms) [FPS]				Decoding (ms)			
			GTX 1080Ti		RTX 4090D		GTX 1080Ti		RTX 4090D	
			768×512	1920×1088	768×512	1920×1088	768×512	1920×1088	768×512	1920×1088
Normal Bitrate	ELIC [20]	-	302.0	1010.1	90.4	300.5	458.0	1648.9	171.1	465.3
	EVC-Small [18]	-	34.7	146.6	20.0	35.3	**32.1**	**120.5**	**14.4**	**33.6**
Ultra-Low Bitrate	PerCo [10]	20	OOM	OOM	245.2	OOM	OOM	OOM	2841.6	OOM
	DiffEIC [38]	50	1002.4	OOM	153.7	1935.4	12742.0	OOM	4785.1	50786.5
	DLF [62]	-	493.9	OOM	94.6	508.1	663.1	OOM	152.1	962.5
	StableCodec [70]	1	328.2	OOM	98.9	538.4	709.3	OOM	192.9	1225.0
	AEIC-ME (Ours)	1	60.3 [16.6]	284.0 [3.5]	37.1 [27.0]	58.7 [17.0]	433.4	3783.1	106.7	845.3
	AEIC-SE (Ours)	1	**22.6** [44.2]	**110.4** [9.1]	**14.0** [71.4]	**27.9** [35.8]	399.6	3629.4	104.2	836.2