

Image and Video Compression using Generative Sparse Representation with Fidelity Controls

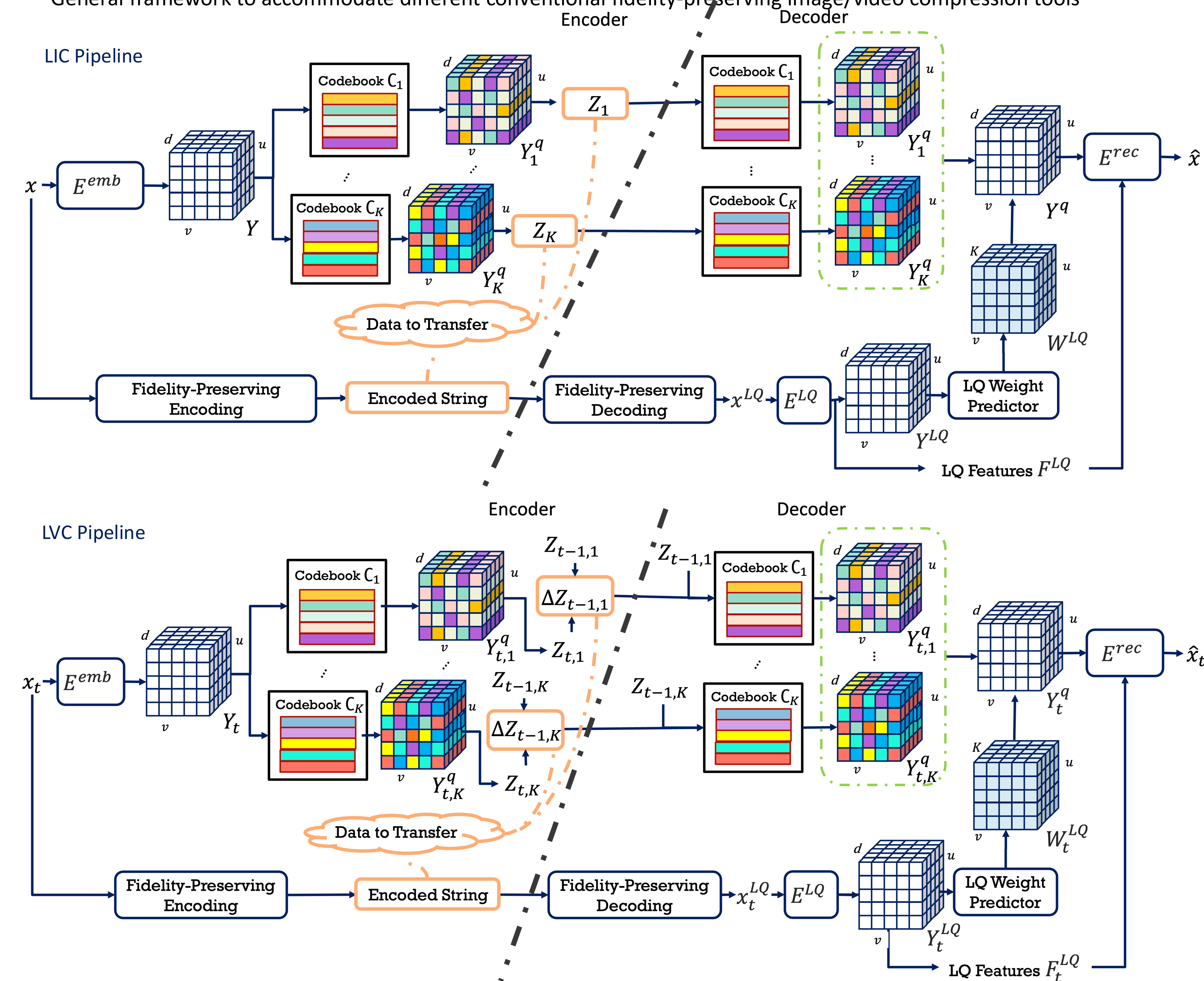
Introduction

Dual-stream framework for learned image compression (LIC) and video compression (LVC) by Sparse Visual Representation (SVR)

- Main stream: high-quality (HQ) codebook-based SVR for HQ baseline reconstruction with high perceptual quality
- Control stream: derive fidelity-preserving controls from low-quality (LQ) input to guide conditional generation in main stream

Advantages over conventional pipeline:

- Balanced perceptual quality and fidelity with flexible bitrate control
- Fully recovered SVR-based reconstruction of all frames that mitigates error propagation
- Unified pipeline for both LIC and LVC
- General framework to accommodate different conventional fidelity-preserving image/video compression tools



Experiments

LIC: JPEG-AI Dataset with 5283/350/50 training/validation/test images

LQ fidelity-preserving LQ by MLIC [1]

SVR codebook from AdaCode [2]

LVC: Mixed Dataset from AOM, JVET, MPEG, and AVS, 134/16 training/validation videos

LQ fidelity-preserving LQ by VVC with qp=42

SVR codebook from FeMaSR [3]

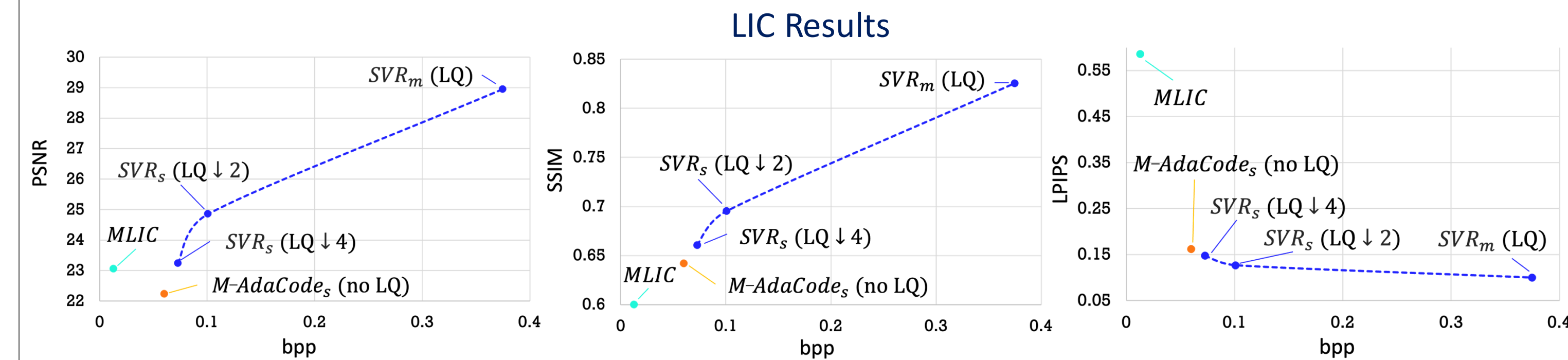
$$B = b_c + b_{LQ}$$

$$\text{LIC } b_c = \sum_{k=1}^K b(Z_k)$$

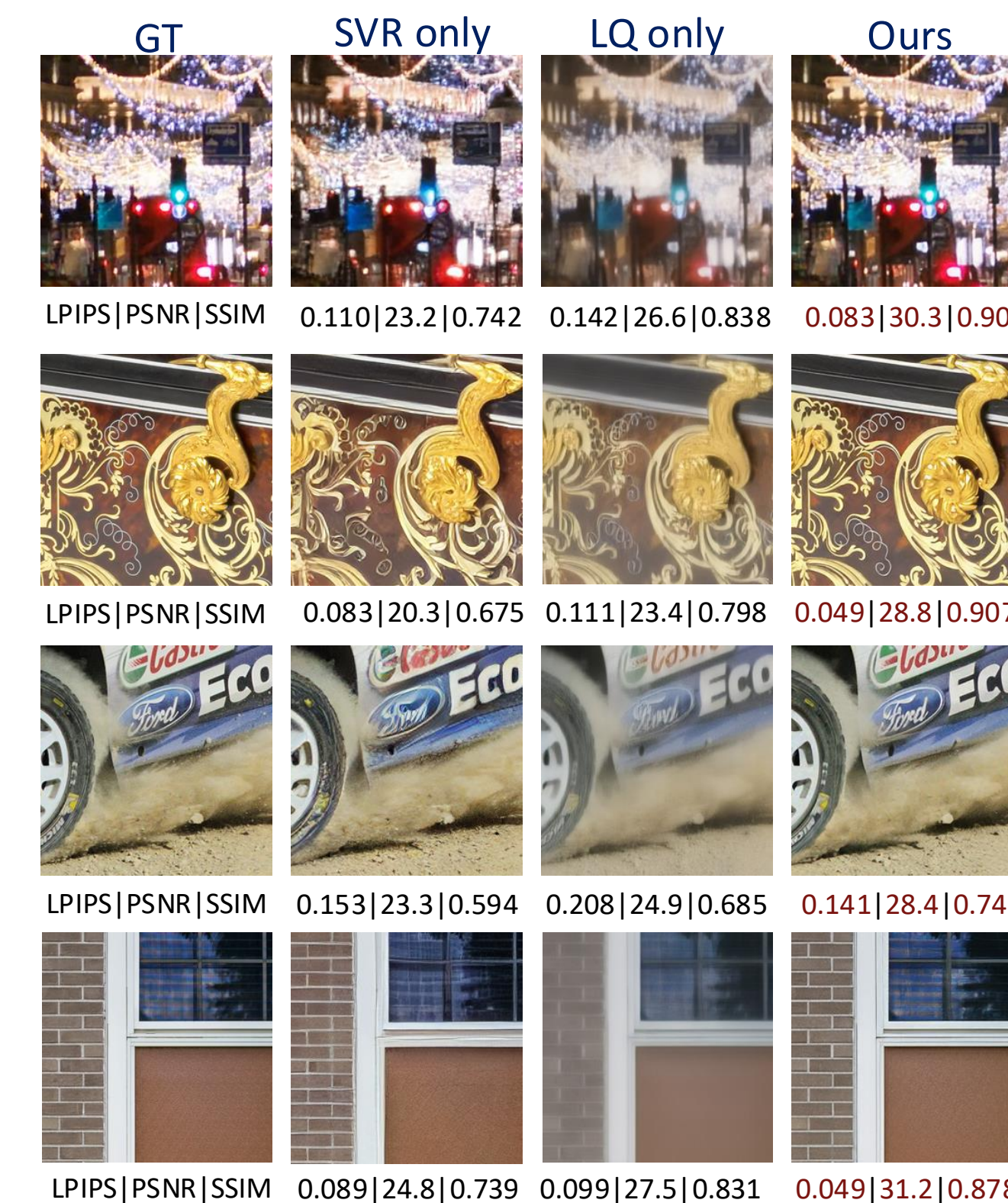
$$\text{LVC } b_c = \sum_{k=1}^K \left[b(Z_{1,k}) + \sum_{t=2}^T b(\Delta Z_{t,k}) \right] / T$$

LVC Results

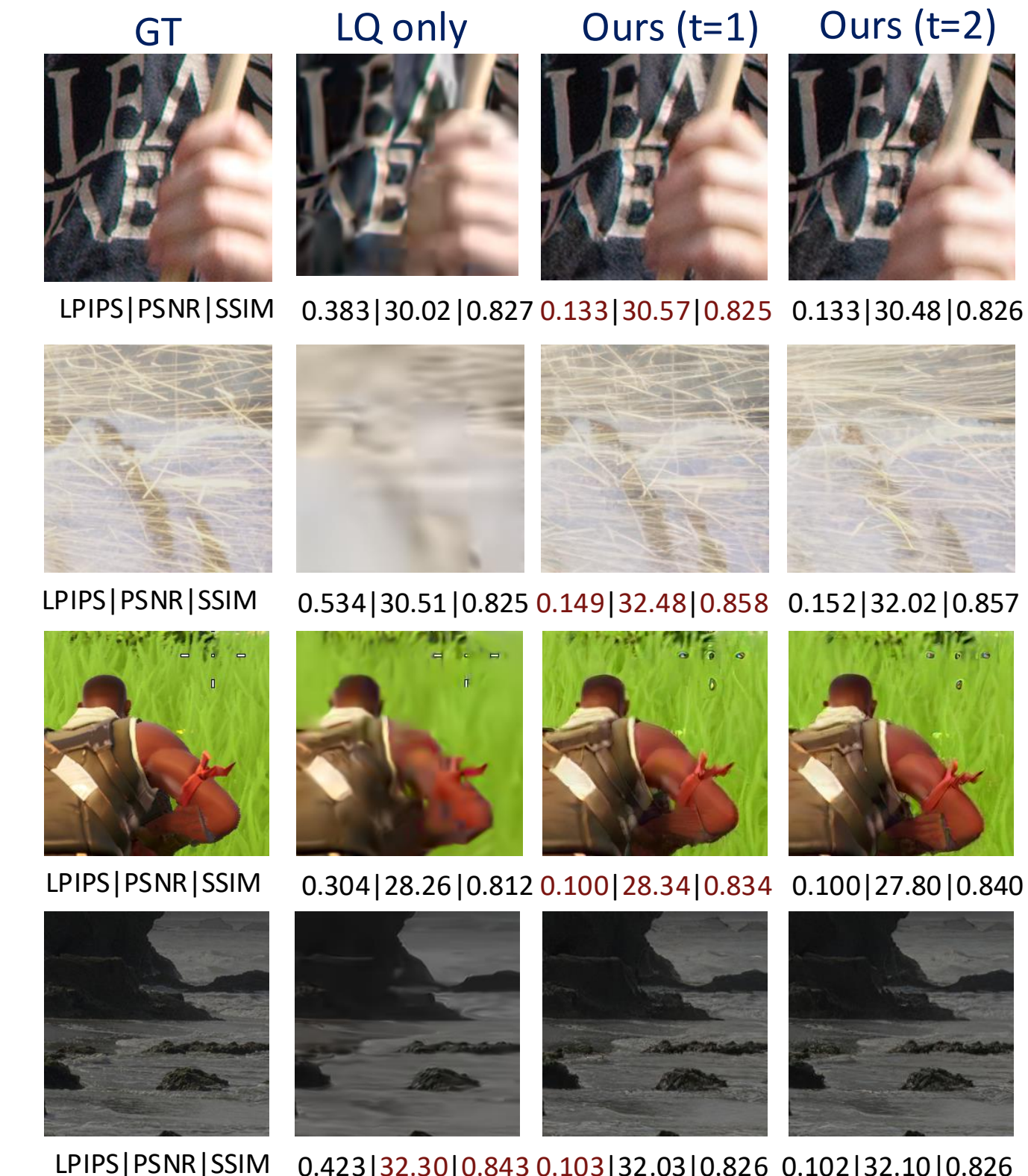
	psnr↑	ssim↑	lpi↓	bpp
Ours	28.15	0.812	0.109	0.095
VVC	28.09	0.806	0.310	0.063



LIC Results



LVC Results



[1] Jiang, W., et al. Mlic: Multi-reference entropy model for learned image compression. ACM Multimedia (2023)

[2] Liu, K., Jiang, Y., Choi, I., Gu, J.: Learning image-adaptive codebooks for class-agnostic image restoration. ICCV (2023)

[3] Chen, C., et al., Real-world blind super-resolution via feature matching with implicit high-resolution priors. ACM Multimedia (2022)