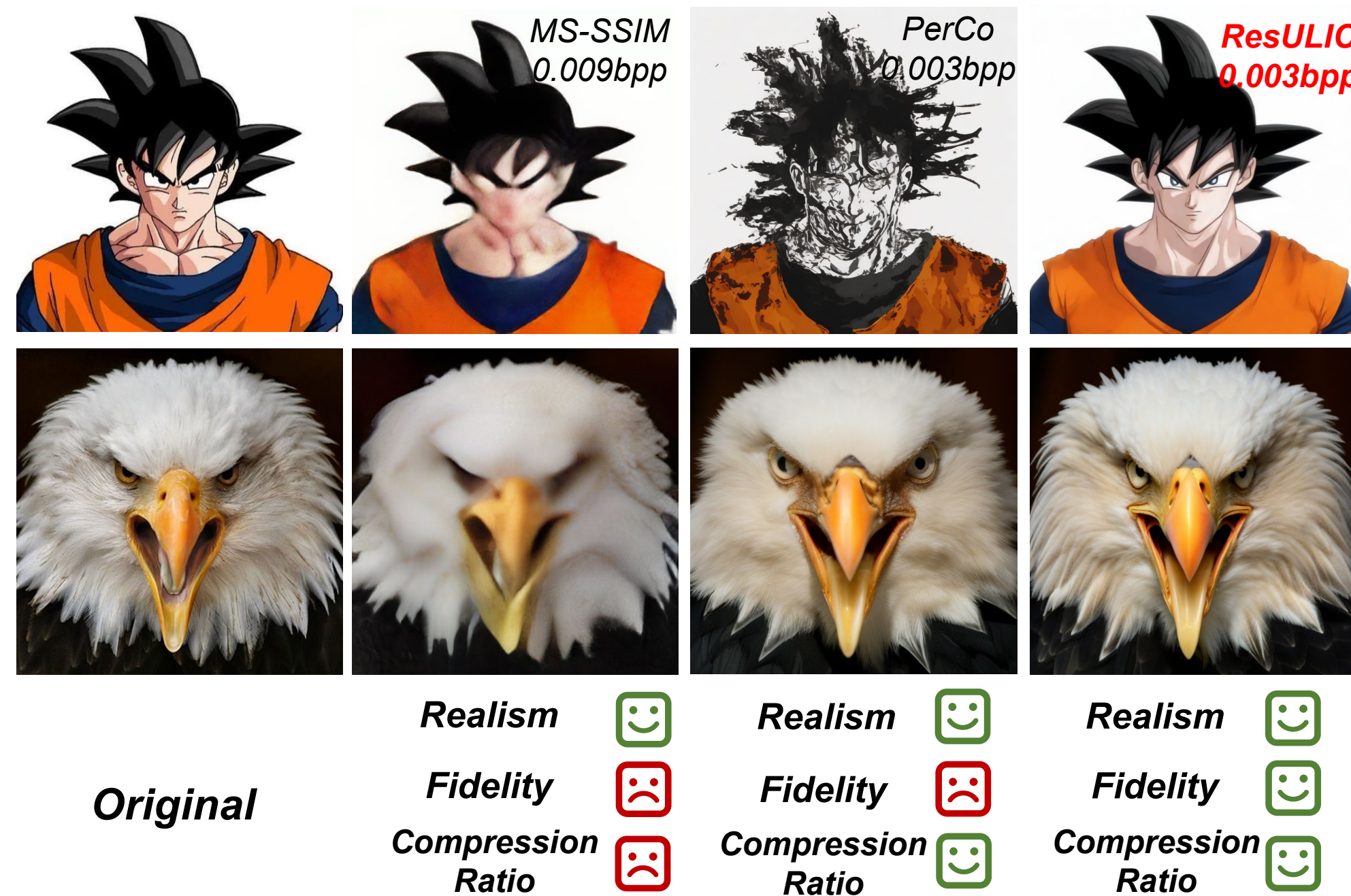


Ultra Lowrate Image Compression with Semantic Residual Coding and Compression-aware Diffusion

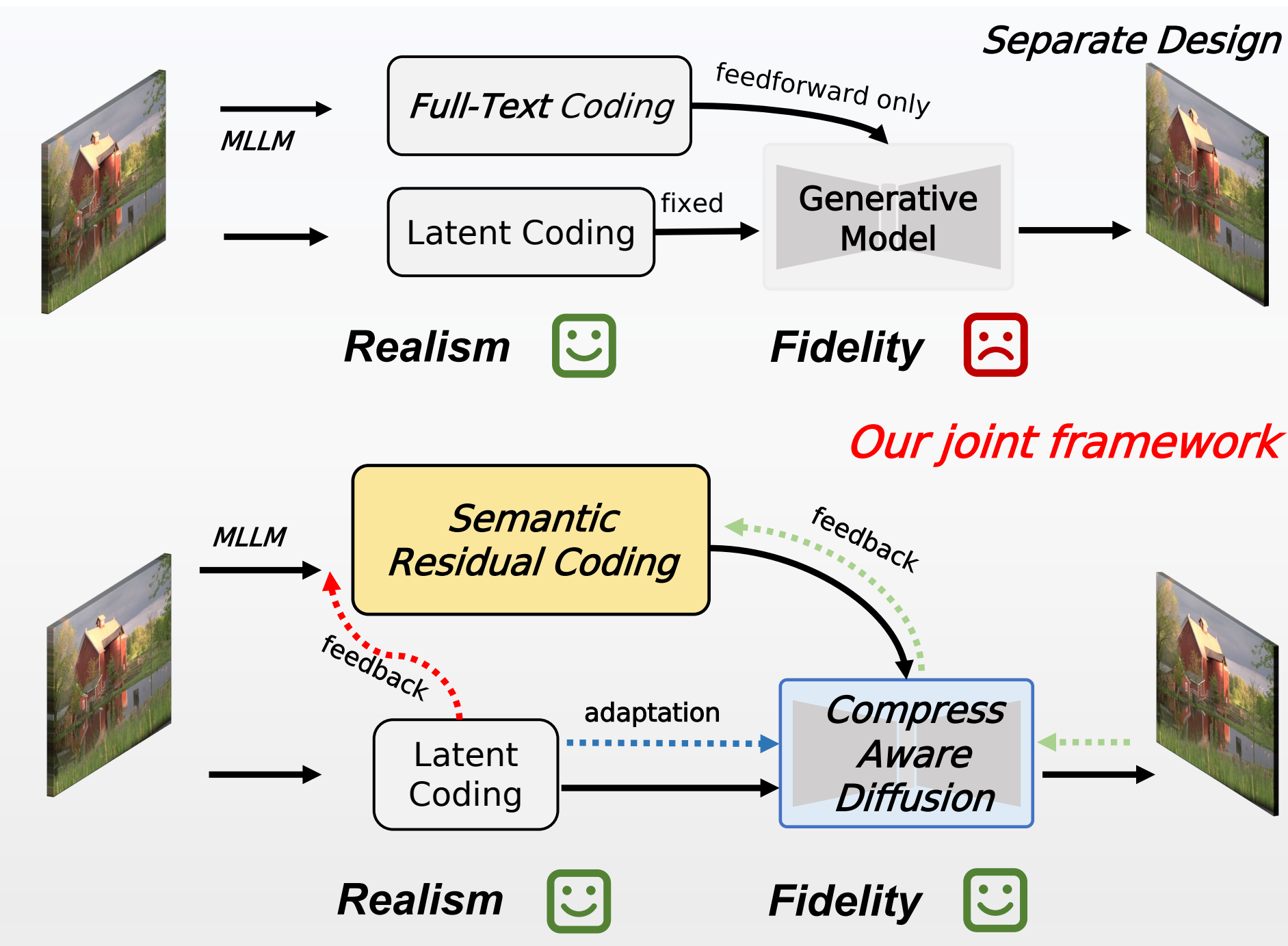
Anle Ke¹, Xu Zhang¹, Tong Chen^{1*}, Ming Lu¹, Chao Zhou², Jiawen Gu², Zhan Ma¹

¹Nanjing University ²Kuaishou Technology

Motivation



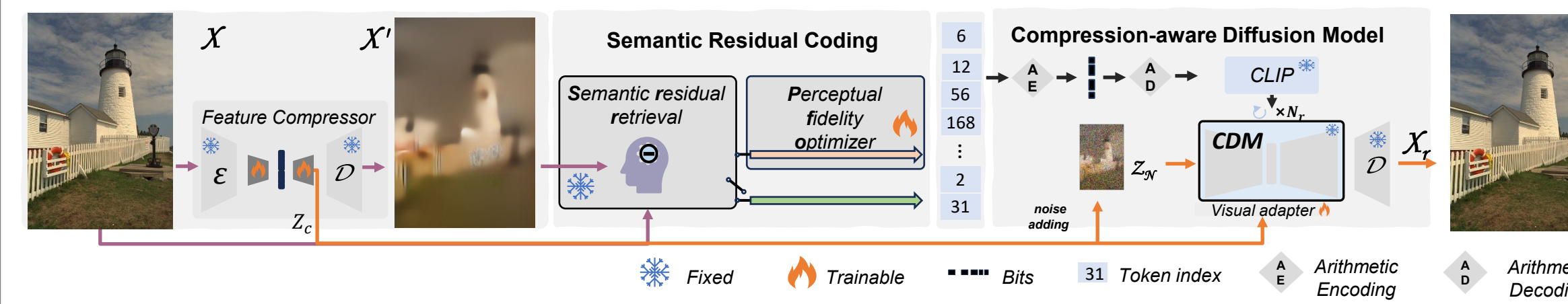
➤ GAN-based methods struggle with clear reconstruction at ultra lowrate compression, while diffusion-based methods lack realism-fidelity balance.



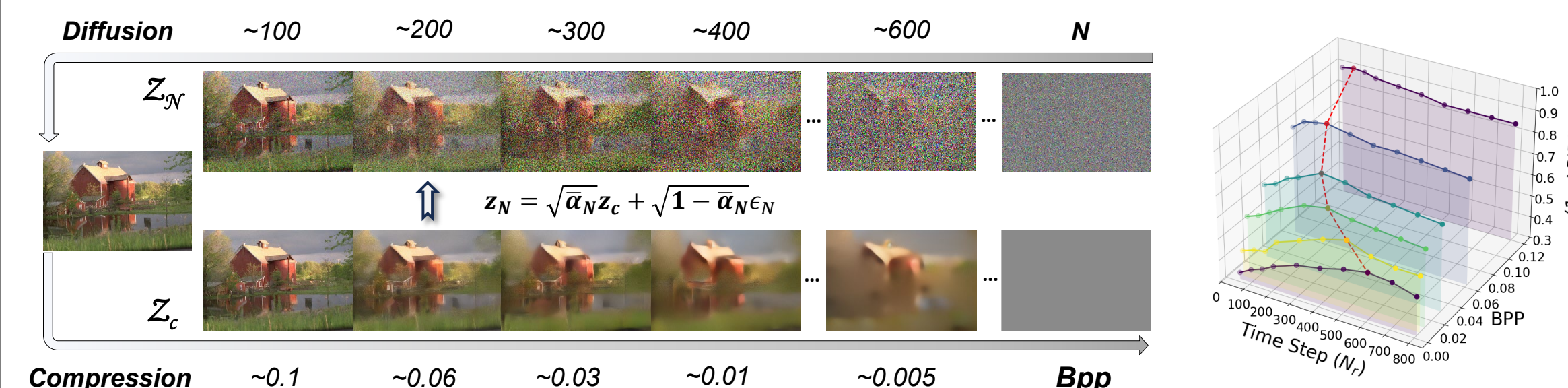
➤ **Problem:** Existing multimodal image compression often fragments semantic retrieval, latent coding, and generative models.

➤ **Solution:** Propose a joint framework that integrates residual signals throughout the compression pipeline.

Methodology



Compression-aware Diffusion Model (CDM)

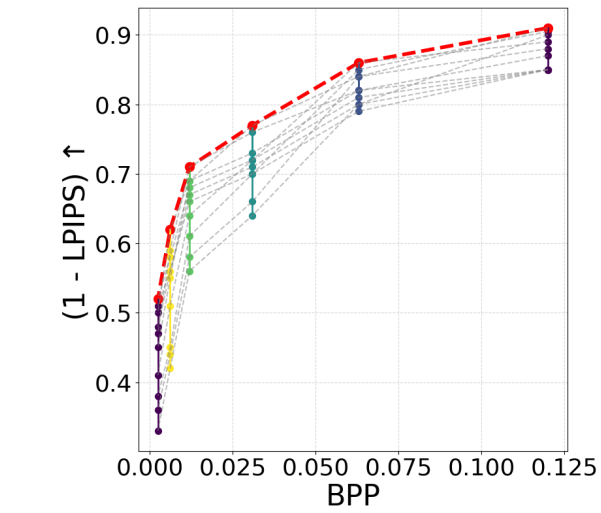


➤ **Forward:** $z_n = \sqrt{\alpha_n} z_0 + \sqrt{1 - \alpha_n} (\gamma_n \rho_{res} + \epsilon_n)$, $n \in \{1, \dots, N_r\}$

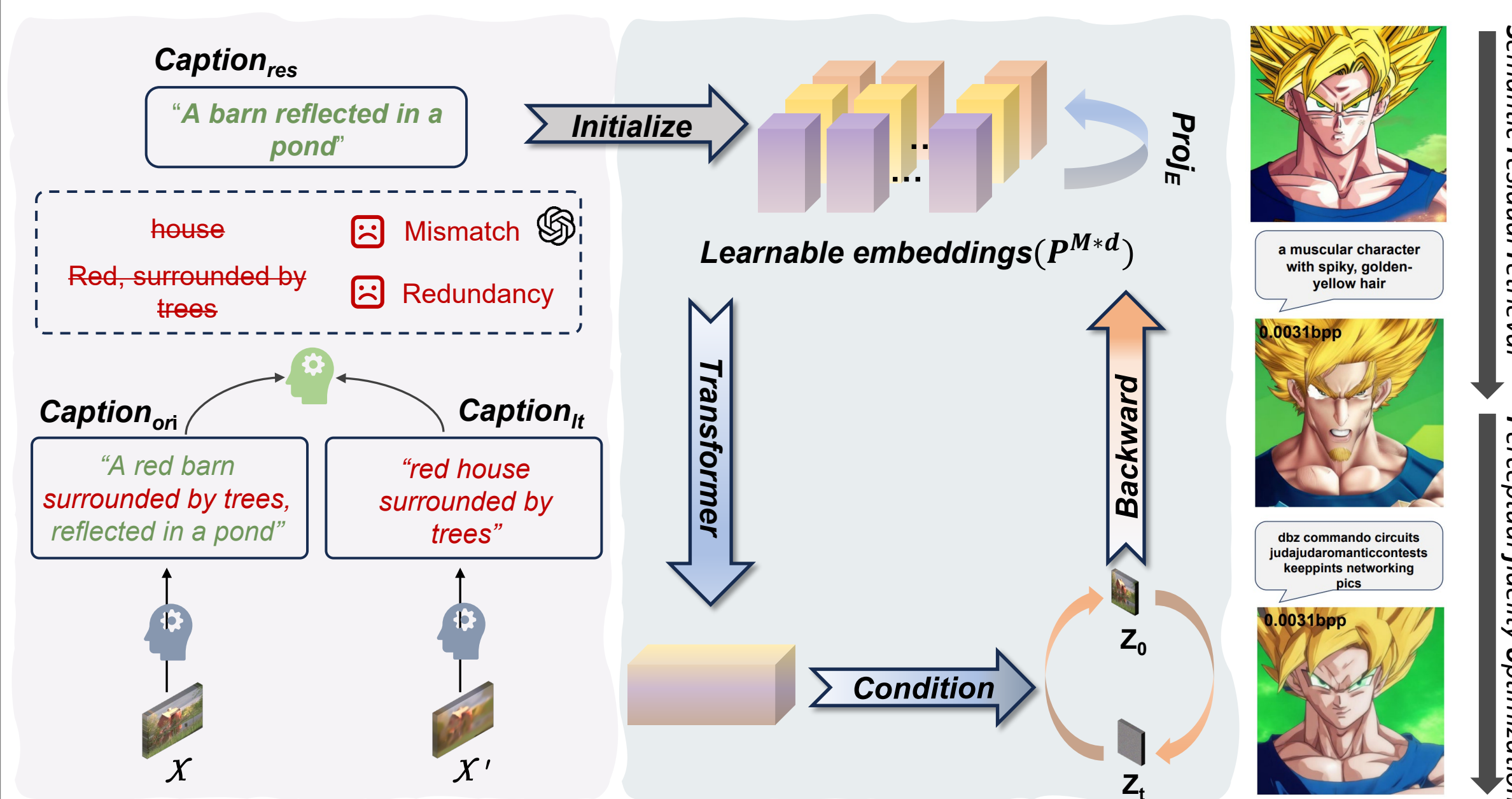
➤ **Backward:** $p_\theta(z_{n-1} | z_n, \tilde{z}_0, z_c) \sim q(z_{n-1} | z_n, z_0, z_c) \sim \mathcal{N}(z_{n-1}; \iota_n z_n + \zeta_n z_0, \sigma_n^2 \mathbf{I})$

➤ **Training:** $\mathcal{L} = \mathbb{E}_{z_0, c, z_c, n, \epsilon} \|\epsilon - \epsilon_\theta(z_n, c, z_c, t)\|^2 + \lambda_d \text{MSE}(x_0, \hat{x}_0) + \lambda_p \text{LPIPS}(x_0, \hat{x}_0) + \mathcal{L}_D + \lambda_R \mathcal{L}_R$

Feature Compressor

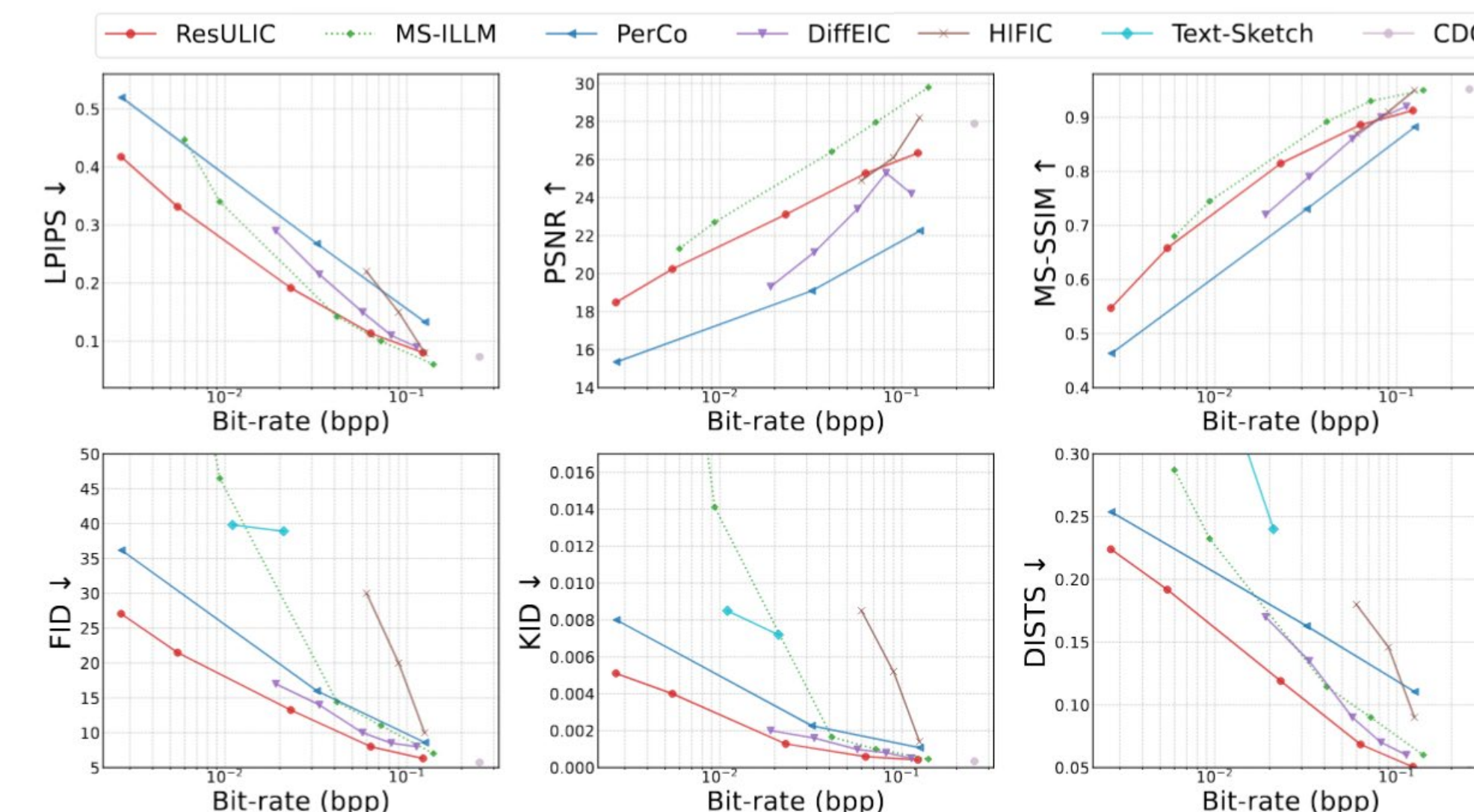
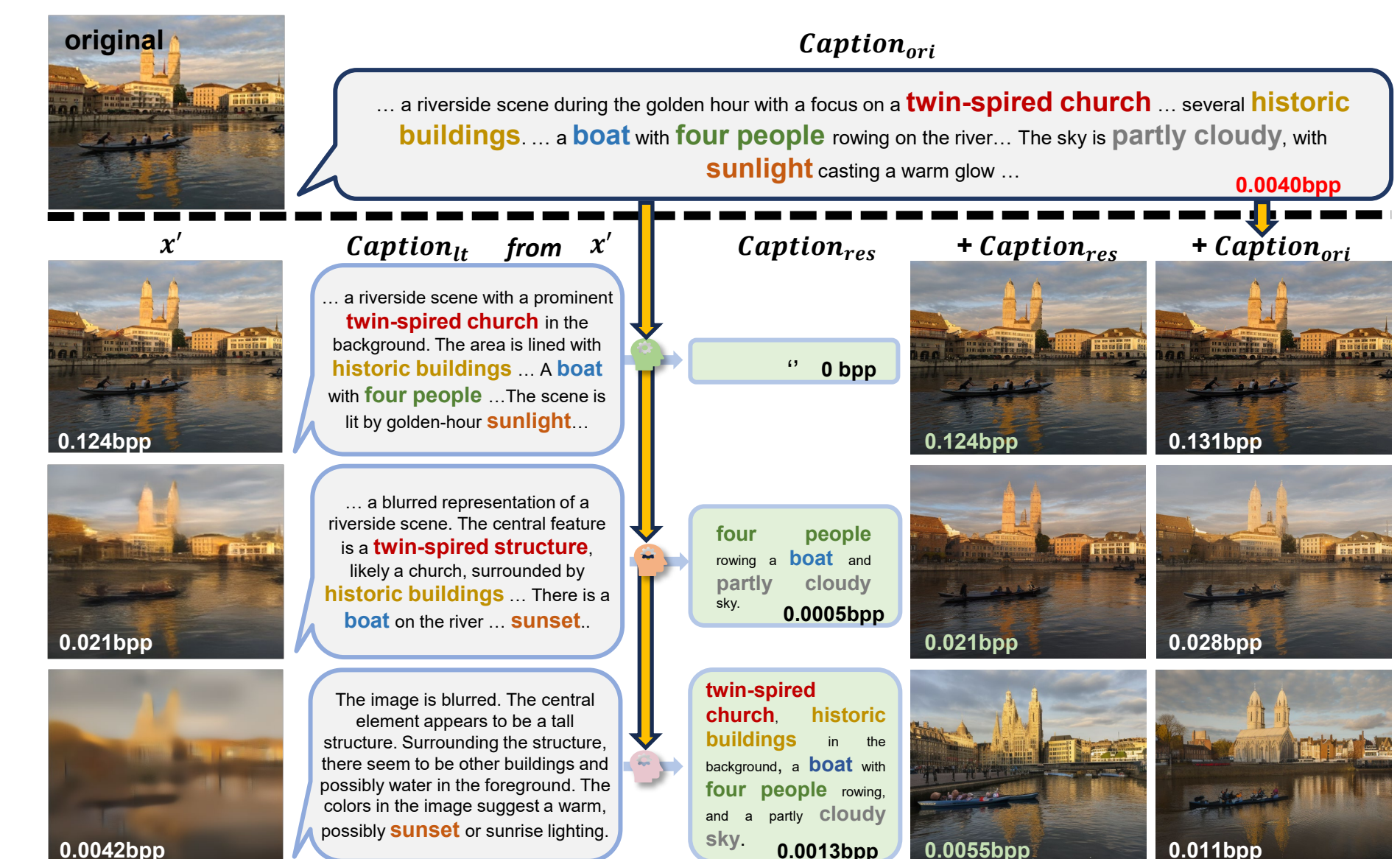


Semantic Residual Coding (SRC)



➤ **Optimization:** $\mathcal{L}_{pfo} = \lambda_l \mathbb{E}_{z_t, \epsilon} [\|\epsilon - \epsilon_\theta(z_n, n, z_c, E_{clip-c}(P'))\|_2^2] + \lambda_c \mathcal{L}_{aux}(P', x)$

Visualization



| Type | Methods | Perceptual Realism | | Perceptual Fidelity | | Distortion | |
|-----------------|-----------------|--------------------|-------|---------------------|-------|------------|---------|
| | | FID | KID | DISTS | LPIPS | PSNR | MS-SSIM |
| GAN based | MS-ILLM | 0 | 0 | 0 | 0 | 0 | 0 |
| | HiFiC | 65.0 | 154.7 | 39.0 | 75.7 | 122.3 | 46.2 |
| Diffusion based | DiffEIC | -28.4 | -21.9 | 1.7 | 9.5 | 392.3 | 133.8 |
| | PerCo | -17.9 | -25.7 | -0.77 | 113.4 | 1092.2 | 244.1 |
| | ResULIC w/o Pfo | -54.9 | -64.7 | -45.7 | -18.8 | 120.4 | 41.8 |
| | ResULIC | -62.9 | -68.5 | -51.6 | -26.9 | 120.3 | 37.2 |

