GameCodec: Neural Cloud Gaming Video Codec

Cloud Gaming

Motivation
- Modern video games require powerful hardware to render frames in high quality
- Cloud gaming is a type of online gaming that aims to address this issue

Game Engine → Neural Encoder → Neural Decoder → Game Player

Challenges of video codecs in the cloud gaming setting
- Requires high quality frames under extremely low-latency constraints
- Rich textures and visual effect
- Extreme camera and object motions

Camera Motion Compensation
- Use camera pose and depth to enable egomotion compensation
- Use a separated Camera Motion AE to encode and transfer the camera motion

Object Motion Compensation
- Object Motion AE based on the P-frame flow AE in the Scale-Space Flow model
- Mean-Scale Hyperprior AE estimates the object motion scale-space vector field \( \mathbf{p}_{obj} \)

Rate-Distortion objective

\[
\mathcal{L}_{\text{RD}}(x) = \beta \cdot \mathcal{L}_A + L_D = \mathbb{E} \left[ \beta \cdot \| x - \hat{x} \|^2_2 - \log p_0(x) \right]
\]

Visualization of decomposed motion compensation

Results

Comparison to literature on TartanAir
+ 26.7% BD-rate savings compared to SSF
+ outperforms HEVC-SCC in low-rate
+ underperforms HEVC-SCC in high-rate

Generalization study on AirSim
+ maintains edge over SSF and HEVC-SCC
+ gap closes with HEVC-SCC on low-rate

Ablation study on TartanAir
+ shows importance of both object and camera motion AE
+ confirms result from SSF where scale-space warping improves performance