

DiffQRCoder: Diffusion-based Aesthetic QR Code Generation with Scanning Robustness Guided Iterative Refinement



Jia-Wei Liao^{1,2}



Winston Wang^{2,*}



Tzu-Sian Wang^{2,*}



Li-Xuan Peng^{2,*}



Ju-Hsuan Weng^{1,2}



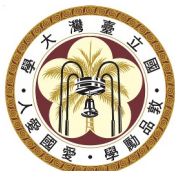
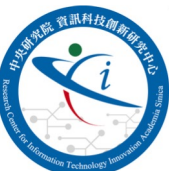
Cheng-Fu Chou¹



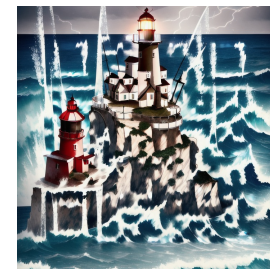
Jun-Cheng Chen²

¹ National Taiwan University,

² Research Center for Information Technology Innovation, Academia Sinica



[Project page](#)



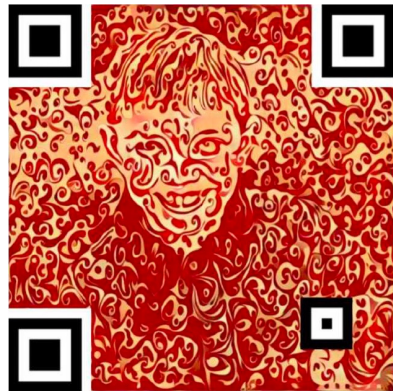
Aesthetic QR Code

Traditional Methods

Qart



ArtCoder



Q-Art Code



Generative-based Method

DiffQRCode (Ours)



Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.



Cherry blossom festival, pink petals floating in the air, traditional lanterns, peaceful river, people in kimonos, sunny day.

QR Code + Prompt + Diffusion Model → Next-Generation Aesthetic QR Code

Motivation & Challenge

Most Diffusion-based aesthetic QR code generation struggle to balance scannability and aesthetics.

- QR Code AI Art and QR Diffusion produce better scanning robust QR codes but are visually less appealing.
- QRBTf could generate visually appealing QR codes, however, they lack scanning robustness.

QR Code AI Art



QR Diffusion



QRBTf



DiffQRCode (Ours)

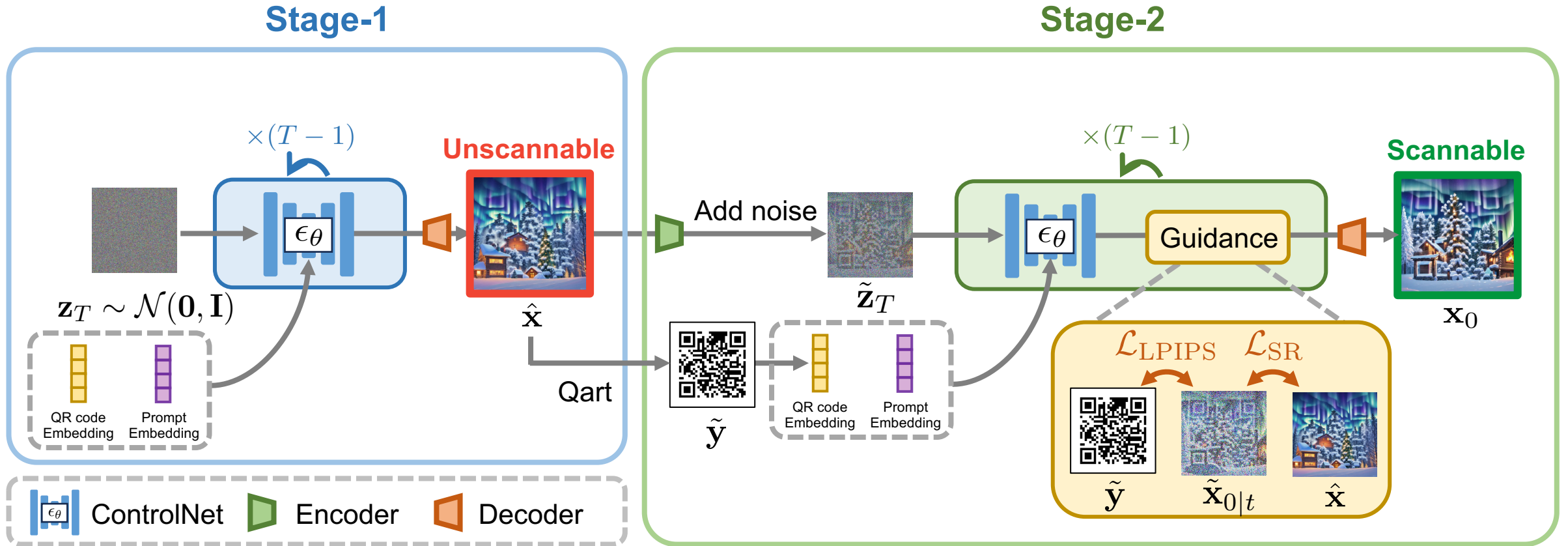


Green: scannable, **Red:** unscannable

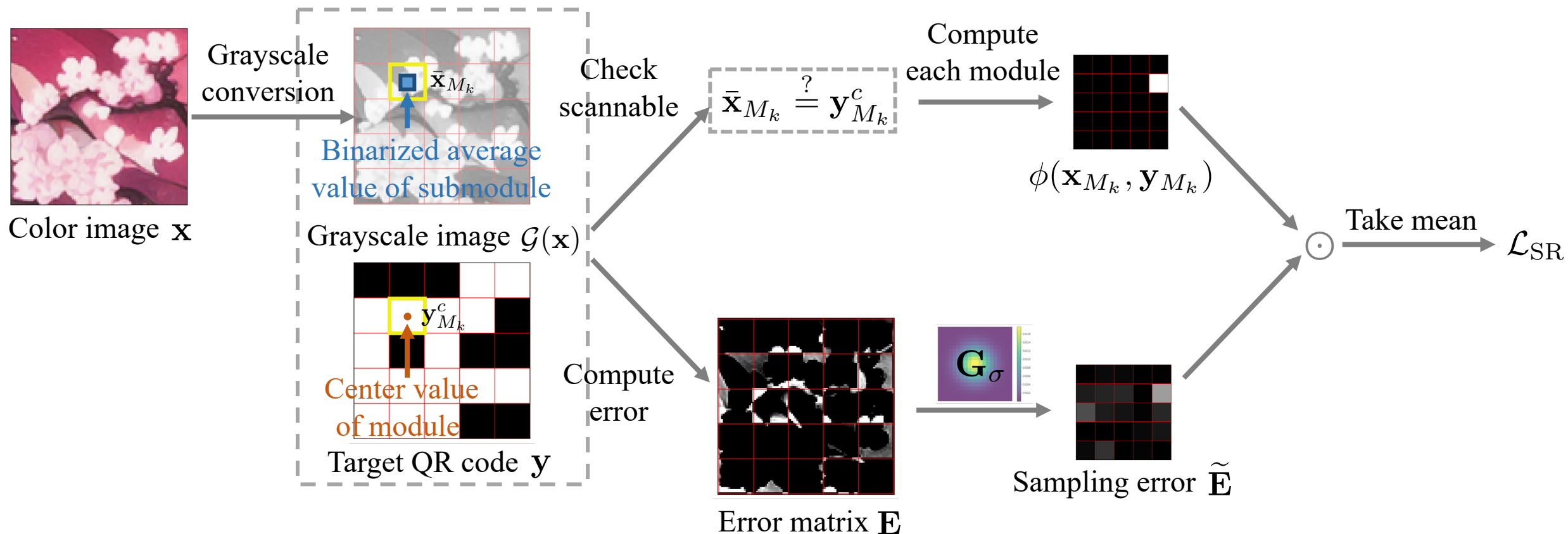
Contribution

- We propose a two-stage iterative refinement framework with **Scanning Robust Perceptual Guidance (SRPG)** to create scanning-robust, visually appealing QR codes without training.
- We develop **Scanning Robust Manifold Projected Gradient Descent (SR-MPGD)**, enhancing the Scanning Success Rate through latent space optimization.
- Our pipeline improves SSR from 60% to nearly 100% compared to ControlNet-only methods, maintaining aesthetics as validated by user evaluations.

Two-stage Iterative Refinement Pipeline



Scanning Robust Perceptual Guidance (SRPG)



$$\hat{\epsilon}_t = \epsilon_\theta(\tilde{\mathbf{z}}_t, t, \mathbf{e}_p, \mathbf{e}_{\text{code}}) + \sqrt{1 - \bar{\alpha}_t} \nabla_{\tilde{\mathbf{z}}_t} [\lambda_1 \mathcal{L}_{SR}(\tilde{\mathbf{x}}_{0|t}, \tilde{\mathbf{y}}) + \lambda_2 \mathcal{L}_{LPIPS}(\tilde{\mathbf{x}}_{0|t}, \hat{\mathbf{x}})]$$

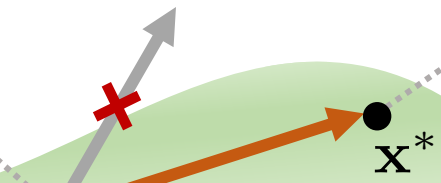
SR-MPGD Post Processing

$$\min_{\mathbf{x} \in \mathcal{M}} \mathcal{L}_{\text{SR}}(\mathbf{x}, \mathbf{y})$$

Unscannable



Gradient Descent



Scannable

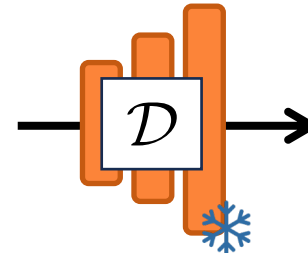


Manifold Projected
Gradient Descent

Image manifold \mathcal{M}

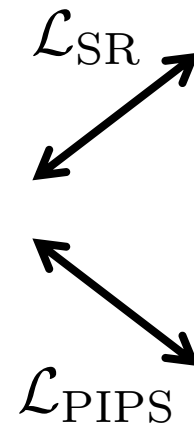


\mathbf{z}



$\mathcal{D}(\mathbf{z})$

\mathcal{L}_{SR}



\mathbf{y}

$\mathcal{L}_{\text{PIPS}}$



\mathbf{x}_0

$$\mathbf{z} \leftarrow \mathbf{z} - \gamma \nabla_{\mathbf{z}} [\mathcal{L}_{\text{SR}}(\mathcal{D}(\mathbf{z}), \mathbf{y}) + \lambda \mathcal{L}_{\text{LPIPS}}(\mathcal{D}(\mathbf{z}), \mathbf{x}_0)]$$

Qualitative Comparisons

Prompt

Winter wonderland, fresh snowfall, evergreen trees, cozy log cabin, smoke rising from chimney, aurora borealis in night sky.

QR Code AI Art



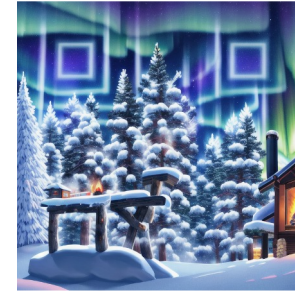
QR Diffusion



QRBTf

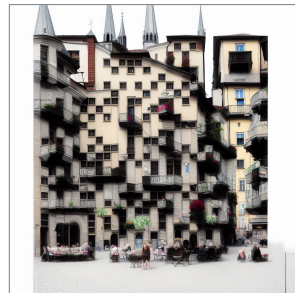


DiffQRCode (Ours)



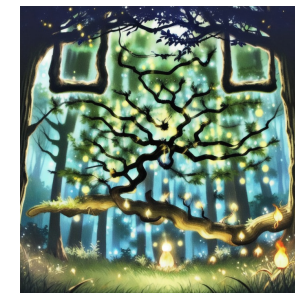
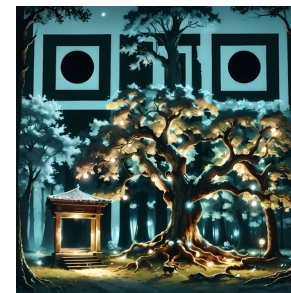
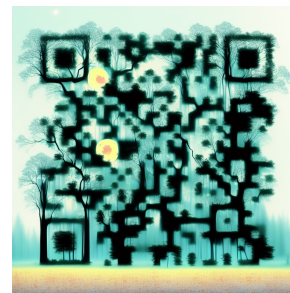
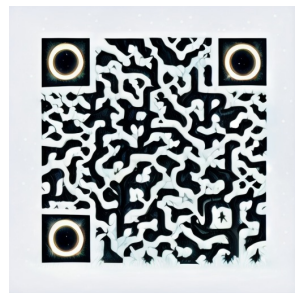
(a) Encoded message: Thanks reviews!

Old European town square, cobblestone streets, café terraces, flowering balconies, gothic cathedral, bustling morning.



(b) Encoded message: I think, therefore I am!

Forest clearing at night, fireflies, full moon, ancient oak tree, soft grass, mystical ambiance.



(c) Encoded message: <https://www.google.com.tw/>

Quantitative Results (I)

- **SSR:** Utilize qr-verify to assess the scanning success rate
- **CLIP-aes.:** Utilize CLIP aesthetic predictor to quantify the aesthetic
- **CLIP-score:** Utilize CLIP to quantify the text-image alignment
- **Avg-rank:** Perform user subjective aesthetic preference study

Method	SSR \uparrow	CLIP-aes. \uparrow	CLIP-score \uparrow	Avg-rank \downarrow
QR Code AI Art [13]	90%	5.7003	0.2341	2.71
QR Diffusion [15]	<u>96%</u>	5.5150	0.2780	3.18
QRBTf [18]	56%	7.0156	0.3033	1.86
DiffQRCoder (Ours)	99%	<u>6.8233</u>	<u>0.2992</u>	<u>2.25</u>

Quantitative Results (II)

- Scannability of different **rotated angles**

Degree	0°	15°	30°	45°
SSR ↑	100%	100%	100%	97%

- Scannability of different **QR code error correction levels**

Level	L (7%)	M (15%)	Q (25%)	H (30%)
SSR ↑	96%	100%	100%	100%

- Scannability of different **encoded messages**

Message	SSR ↑
I think, therefore I am.	97%
You are the apple of my eye.	100%
https://www.google.com.tw/	100%
https://www.wikipedia.org/	97%

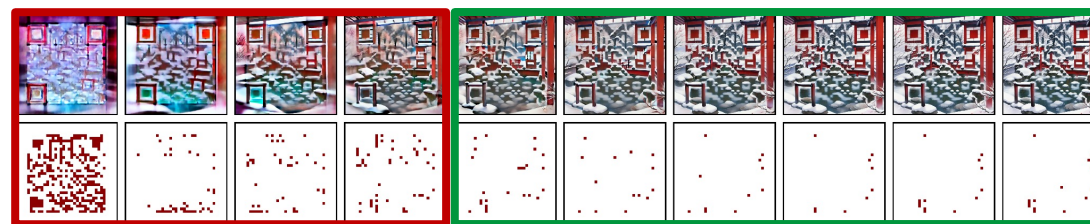
Ablation Study

Stage	λ_1	λ_2	SR-MPGD	CLIP-aes. \uparrow	SSR \uparrow
Stage-1-only	-	-		7.0661	60%
Two-stage	400	0		6.7860	86%
Two-stage	500	0		6.7259	88%
Two-stage	600	0		6.7183	94%
Two-stage	1000	0		6.5667	93%
Two-stage	400	0	✓	6.7567	98%
Two-stage	500	0	✓	6.7097	100%
Two-stage	600	0	✓	6.7002	99%
Two-stage	1000	0	✓	6.5629	99%
Two-stage	500	2		6.8600	90%
Two-stage	500	3		6.8744	89%
Two-stage	500	5		6.8357	89%
Two-stage	500	10		6.8409	88%
Two-stage	500	2	✓	6.8204	98%
Two-stage	500	3	✓	6.8233	99%
Two-stage	500	5	✓	6.7779	100%
Two-stage	500	10	✓	6.8040	97%

Table: Ablations for our proposed pipeline.

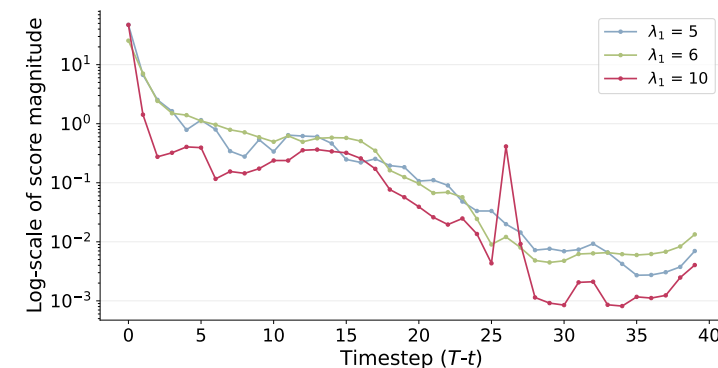
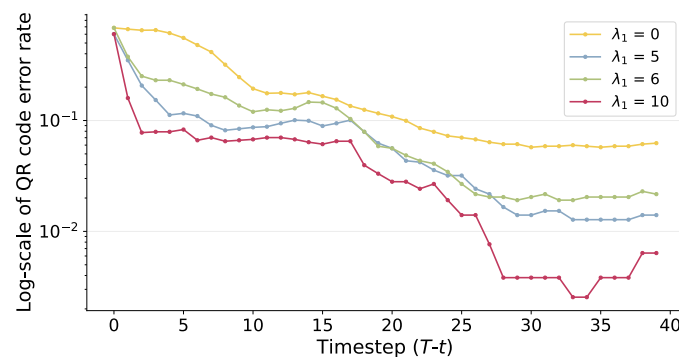
Error Analysis

$t = 40$ → $t = 0$



**Timestep 40 ~ 25
(Unscannable)**

**Timestep 25 ~ 1
(Scannable)**



Takeaway

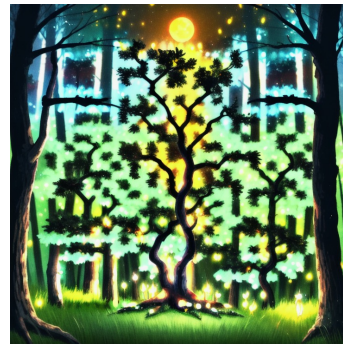
- We can add control to diffusion models via customized deterministic loss function without relying on pre-trained models or adapting additional modules.
- By breaking down the QR code scanning process and underlying mechanisms, we can design a differentiable loss function that serves as a gradient source for diffusion model guidance.
- Furthermore, leveraging VAE for latent optimization ensures improved visual quality while maintaining scannability.

Thanks for listening!

Project Page



Paper



Code

