Diffusion-based Holistic Texture Rectification and Synthesis

Guoqing Hao[1,2], Satoshi Iizuka[1], Kensho Hara[2],
Edgar Simo-Serra[3], Hirokatsu Kataoka[2], Kazuhiro Fukui[1]

[1]University of Tsukuba

[2]AIST

[3]Waseda University

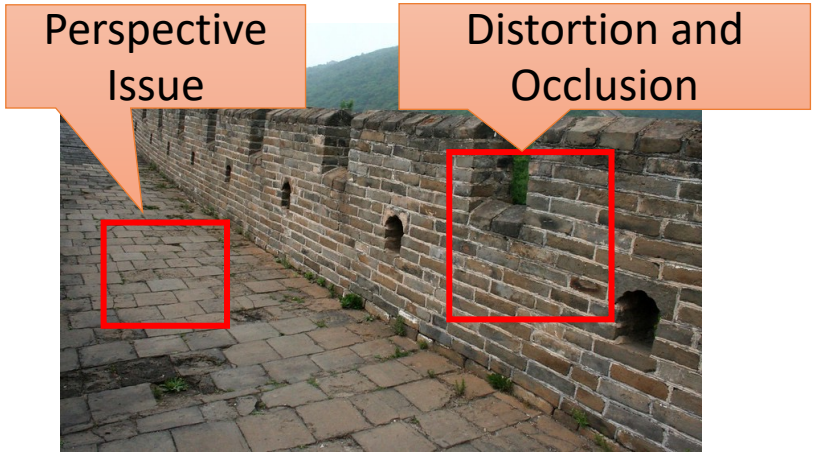
Background

- Texture represents visual characteristics of an object's surface
  - Application: 3D modeling and image editing
  - Derivation: hand-drawn images; real images
- Difficult to obtain planar texture
  - Distortions, perspective issues, and occlusions in real images

Original Image

Planar Texture

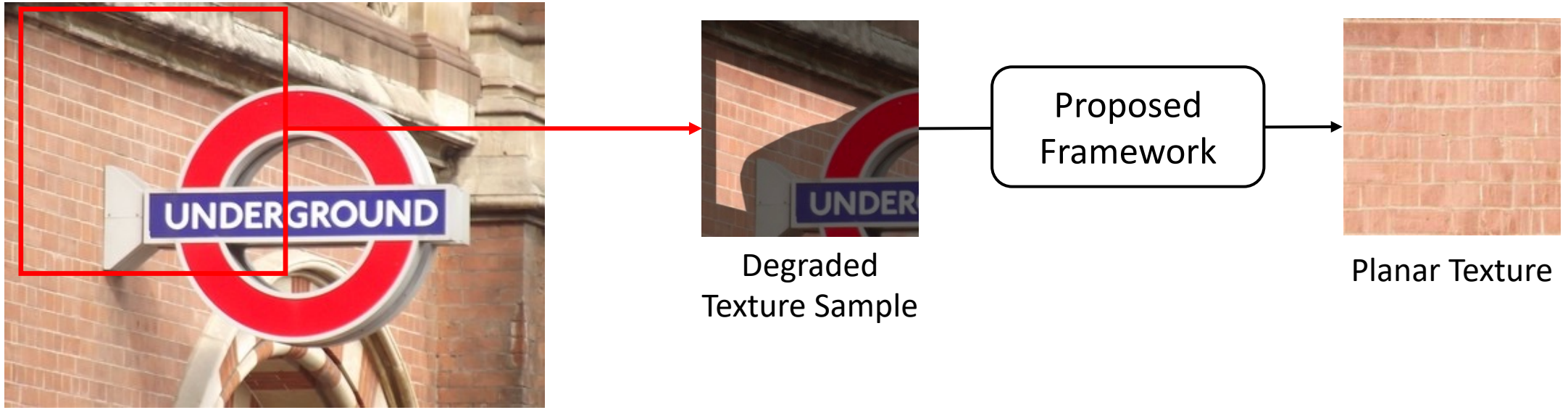
Edited Image



Real Image

Our Goal

- Obtaining planar textures from degraded textures in real images
  - By rectifying distortions, perspective issues and occlusions
  - Preserving visual characteristics
  - Extending applicability of texture synthesis



Related Work: Exemplar-based Texture Synthesis

- Synthesizing larger texture from small texture sample

  [Efros and Freeman 2001; Liu+ 2020; Mardani+ 2020]

- Texture samples must be square and undistorted

- Difficult to obtain such texture samples:
  - Perspective issues, geometric distortions, and occlusions
  - Easy to synthesize disappointing texture

Real Image

Texture Sample

Synthesized Texture

Related Work: Texture Scraping [Li+ 2022]

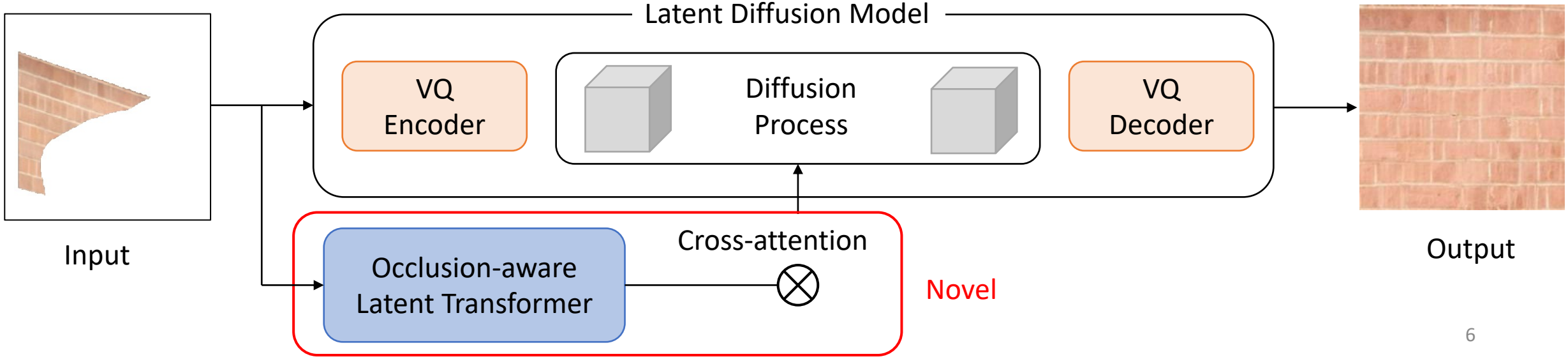
- Texture scraping from real images
  - Texture grouping with convolutional networks and graph networks
  - Texture synthesis by completing missing regions
- Handle occlusions but ignore geometric distortions
  - Generate unsatisfactory results

Input Image Grouping Texture

Input Image Grouping Texture

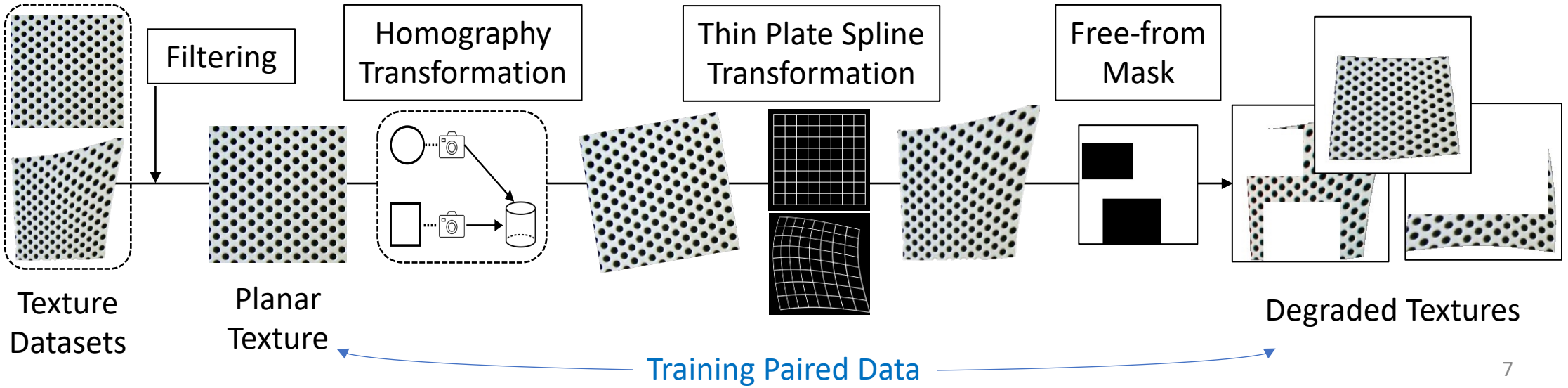
Methodology Overview

- First diffusion-based framework for rectifying distortions and occlusions in textures
- Occlusion-aware latent transformer
  - effectively compute guidance for the generation process
- Novel mechanism for synthetic training dataset construction

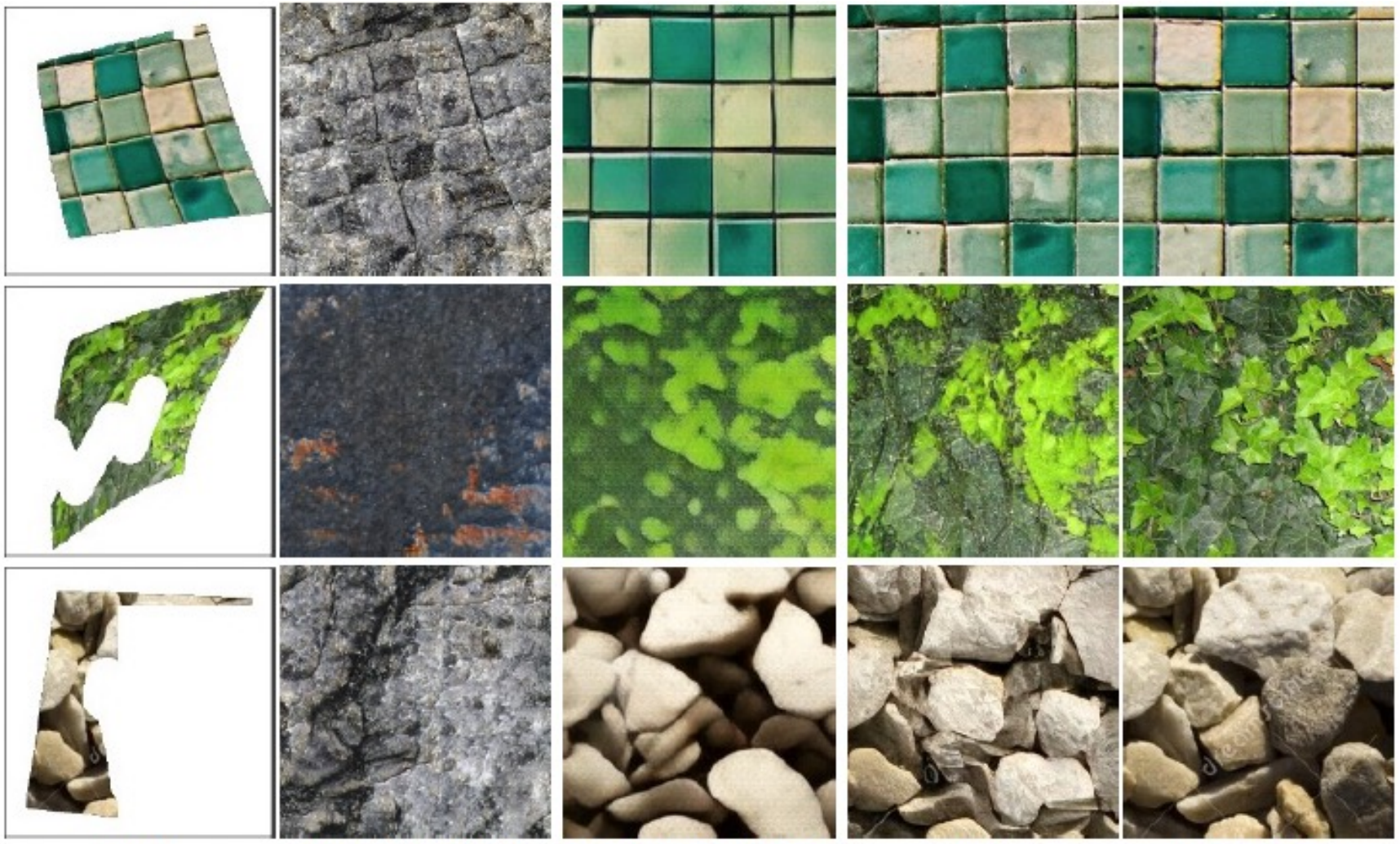


Synthetic Training Dataset Construction

- Collect 22,043 planar texture images from various sources
- Apply transformations and masking on planar textures:
  - Homography transformation; Thin Plate Spline transformation; Free-from mask
- Generate various paired texture degradation data



Qualitative Results on Synthetic Test Images



Input Texture | MAT [Li+ 2022] | VQGAN [Esser+ 2021] | Ours | Ground Truth

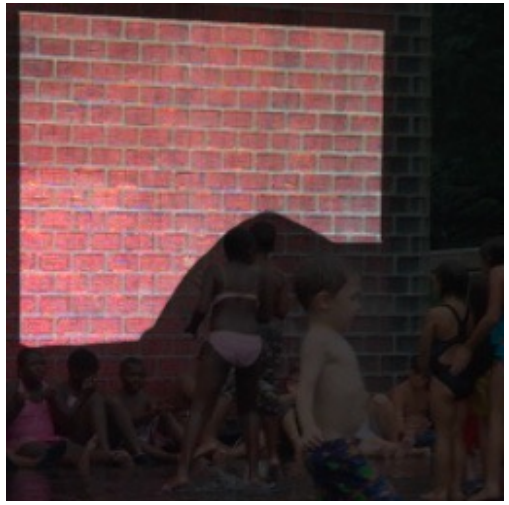
Qualitative Results on Real Images

Real Image	MAT [Li+ 2022]	VQGAN [Esser+ 2021]	Ours

Failure Cases

- Produce imperfect results under varying lighting conditions and extreme distortions
- Possible solutions:
  - Masking regions with significant lighting changes
  - Introducing more training data with extreme distortions

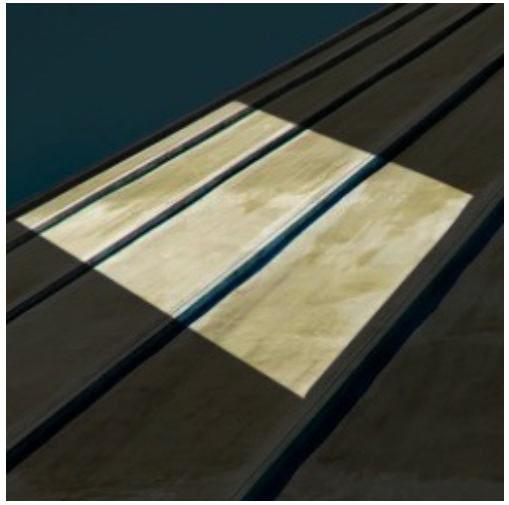
Varying Lighting Condition



Input Image

Rectified Texture

Extreme Distortion



Input Image

Rectified Texture

Conclusion

- The first framework for rectifying distortions and occlusions in textures
- Introduce a new occlusion-aware latent transformer
- Propose a novel mechanism for synthetic training dataset construction
- In-depth evaluation that demonstrates the superior performance of the proposed framework