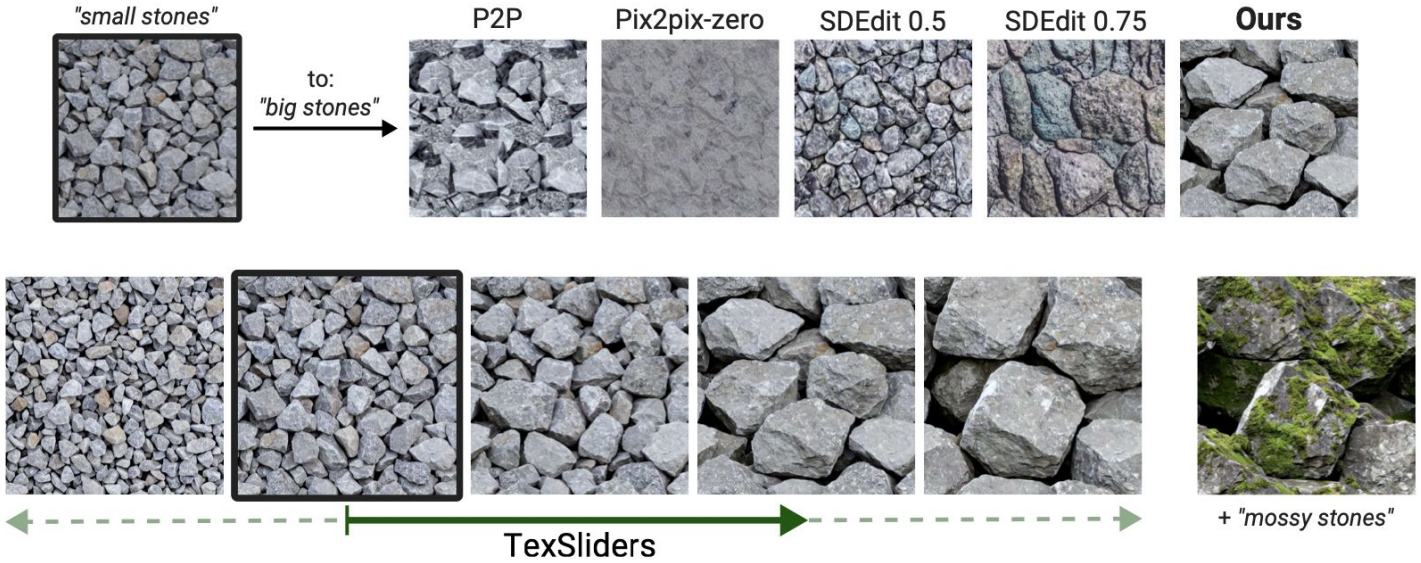
TexSliders for Background Texture Editing for Arbitrary Images

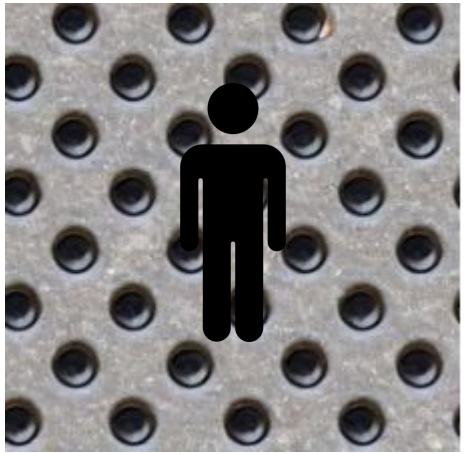
Team 4
Chanryeol Lee, Chanhyuk Lee

Review: Proposal

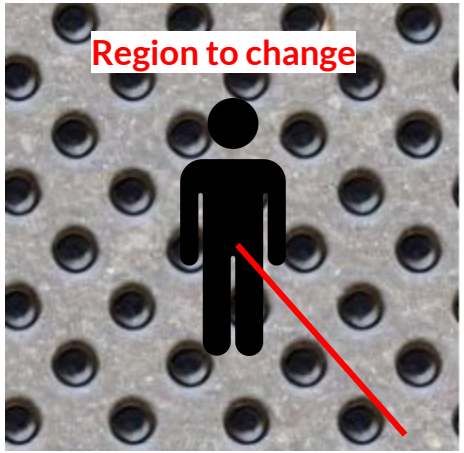
TexSliders: Diffusion-Based Texture Editing in CLIP Space
J. Guerrero-Viu et al., SIGGRAPH 2024



Review: Proposal



Metal



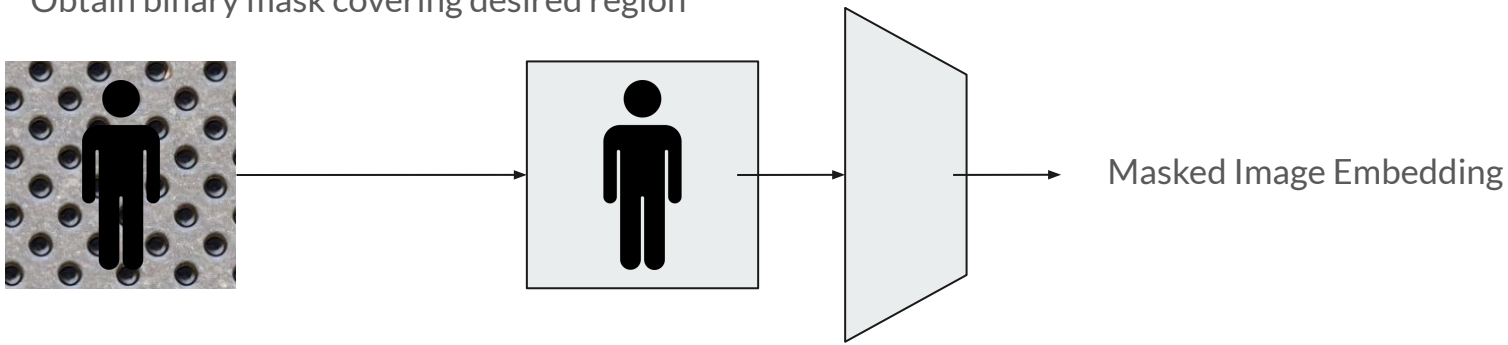
Region to preserve

Rusty metal

- Input: arbitrary image, desired region (mask), text prompt
- Output: image with modified texture of desired region

Review: Proposal

Obtain binary mask covering desired region



Masked Image Embedding

Encoder (CNN)

CLIP - Pretrained ResNet50

However....

Re: Request for Codebase Access - TexSliders: Diffusion-Based Texture Editing in CLIP Space

Julia Guerrero Viu <juliagviu@unizar.es> 11.11 22:00 [209.85.218.47, United States of America]
받는 사람: 이찬혁

Hi,

Glad you are interested in our work!

Unfortunately, I cannot share the code of this work as it was developed during an internship at Adobe, and the code is private for the company.

Problems

1. The original paper uses the diffusion prior model specially trained only with texture images.
   - Using open-source diffusion model for prior largely harmed the performance.

2. Datasets were strictly generated from Adobe stocks
   - Hard to collect datasets for our project.

Thus, we concluded that it’s hard to maintain this topic without access to the private resources.

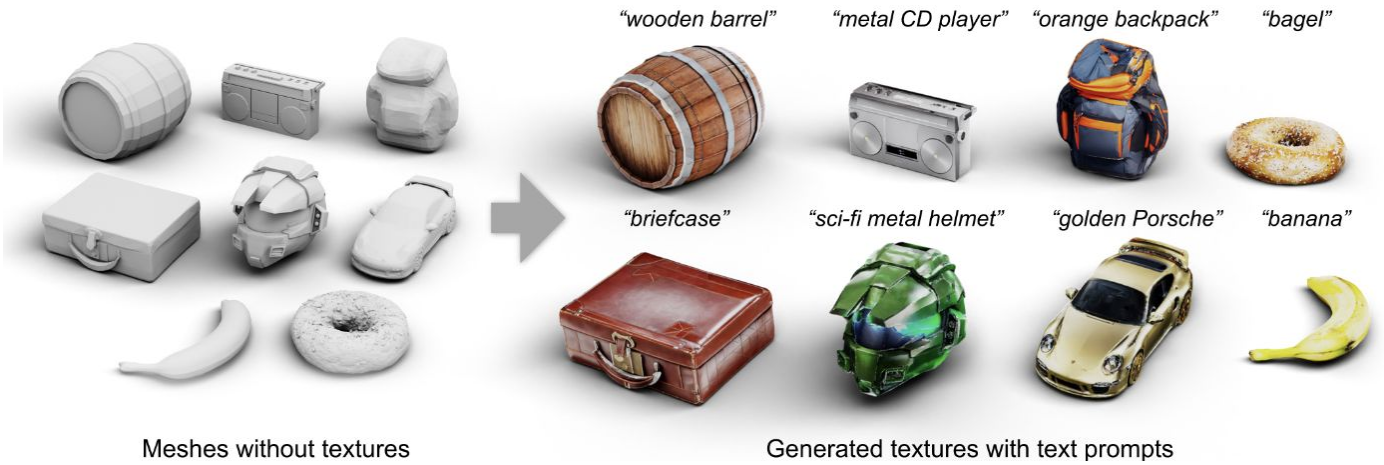
So……

Autonomous View Selection for Improved Text2tex Refinement

Team 4
Chanryeol Lee, Chanhyuk Lee

Review: Text2Tex

Generate high-quality textures **based on given 3D meshes** from the given text prompts!



"Text2tex: Text-driven texture synthesis via diffusion models." CVPR 2023

Review: Text2Tex

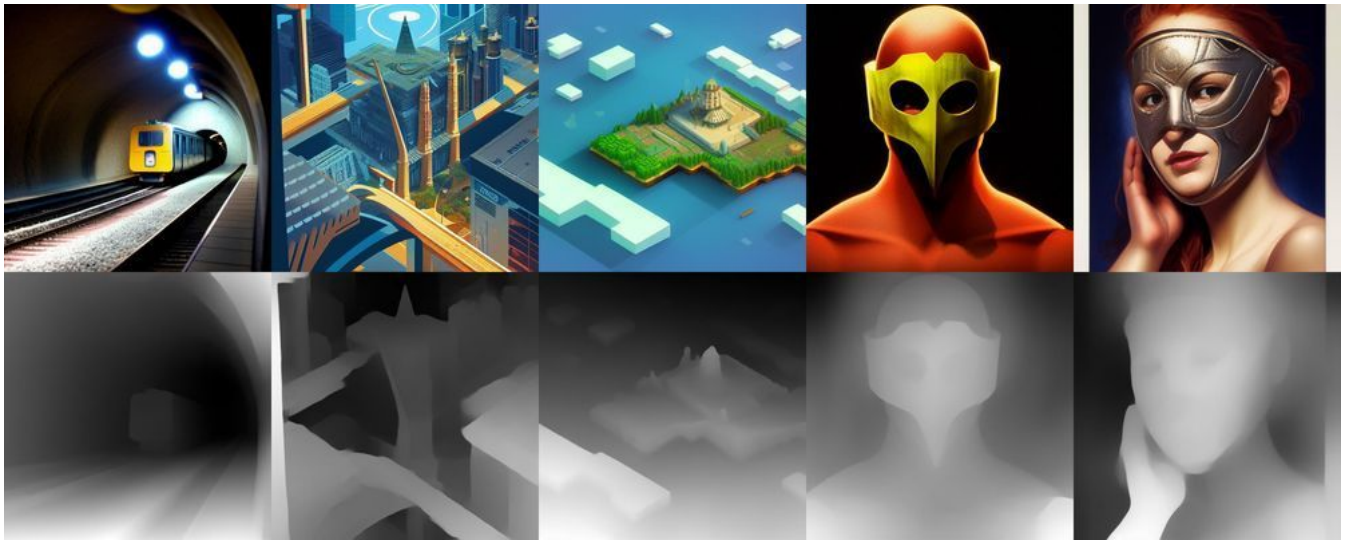
Goal: Non-Optimization Approach & Minimize Human Effort in Text-to 3D

1) Use **pretrained** text-to-image diffusion model, without **additional training**

2) Dynamic View Partitioning: **Generate-then-Refine**

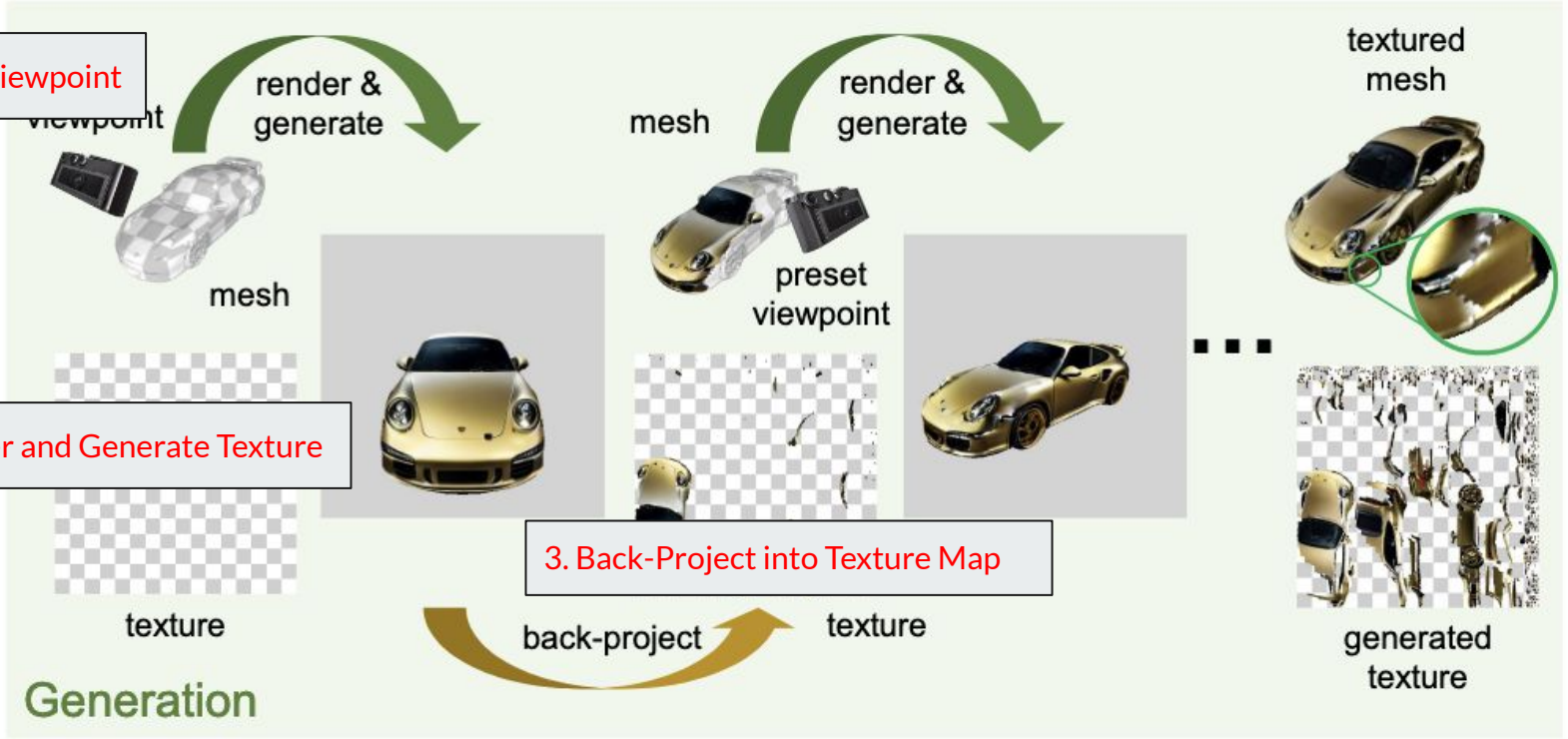
Pretrained Model: Depth2Image

From Depth information of pixels, generate appropriate image aware of 3d geometry



Depth-to-Image, https://huggingface.co/docs/diffusers/api/pipelines/stable_diffusion/depth2img

Progressive Texture Generation

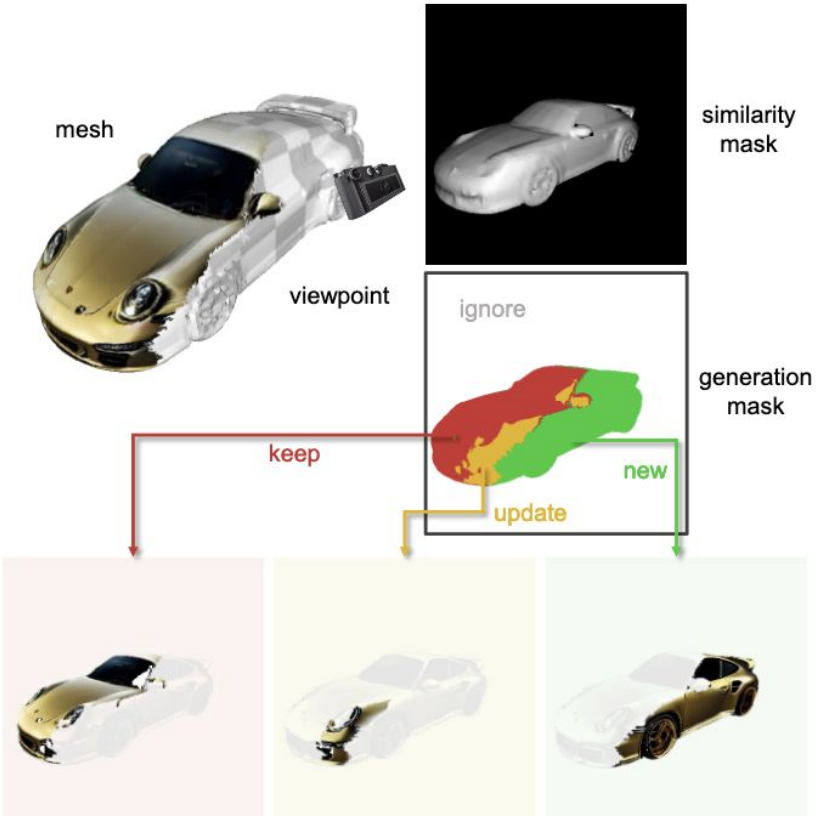


Dynamic View Partitioning

Generate Similarity Mask based on **Surface Normal**

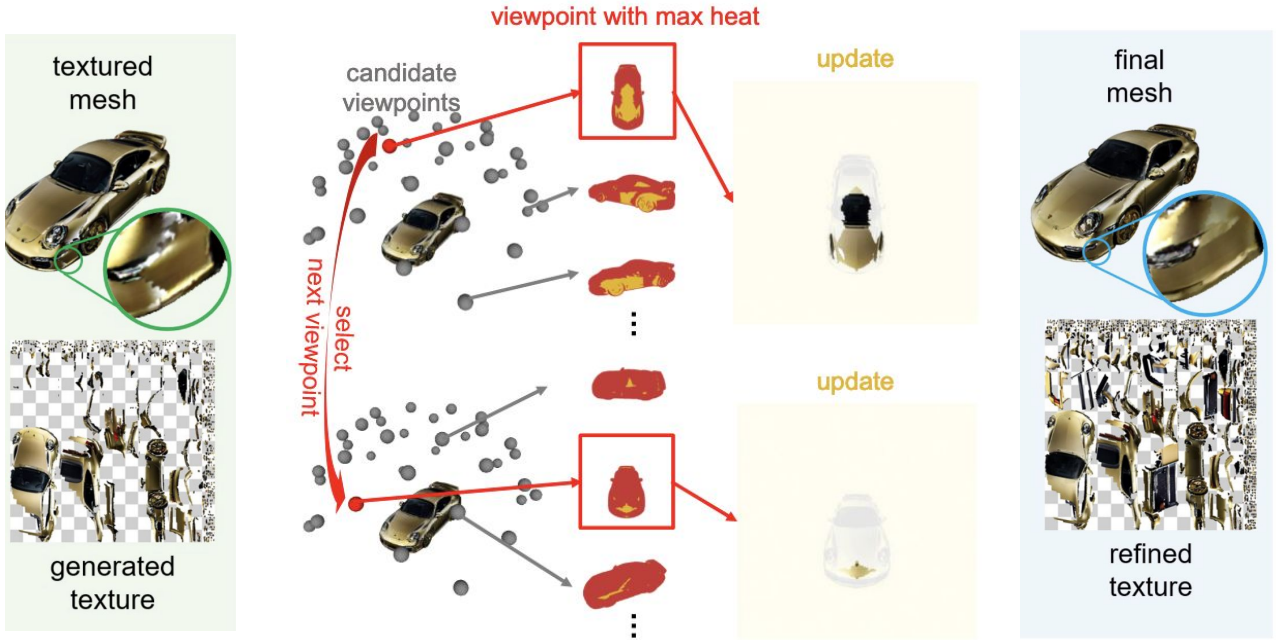
Based on similarity mask,
Partition generation mask into 4 types

1) **New**: not yet textured
2) **Update**: high similarity score (perpendicular to view angle)
3) **Keep**: low similarity score
4) **Ignore**: background



Texture Refinement with Automatic Viewpoint Selection

Refine the generated textures from **automatically chosen** additional viewpoints



Running Results of Original Code

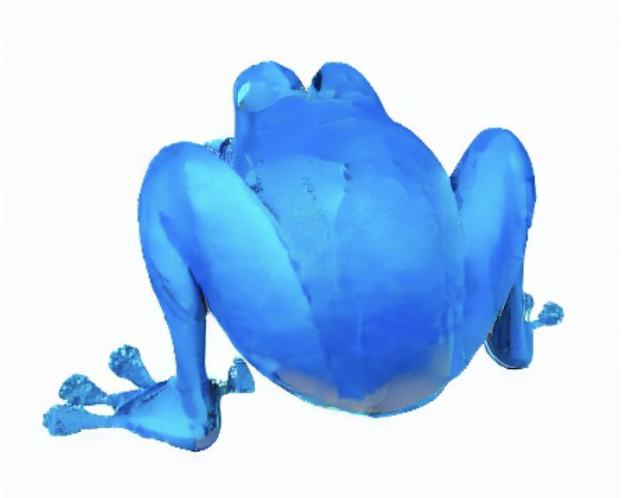
Prompt: Orange Backpack

Running Results of Original Code

Prompt: Banana

Bad Case

Prompt: Blue Frog



Another Failure Case

- Inconsistency between opposite views

Prompt: Yellow School Bus

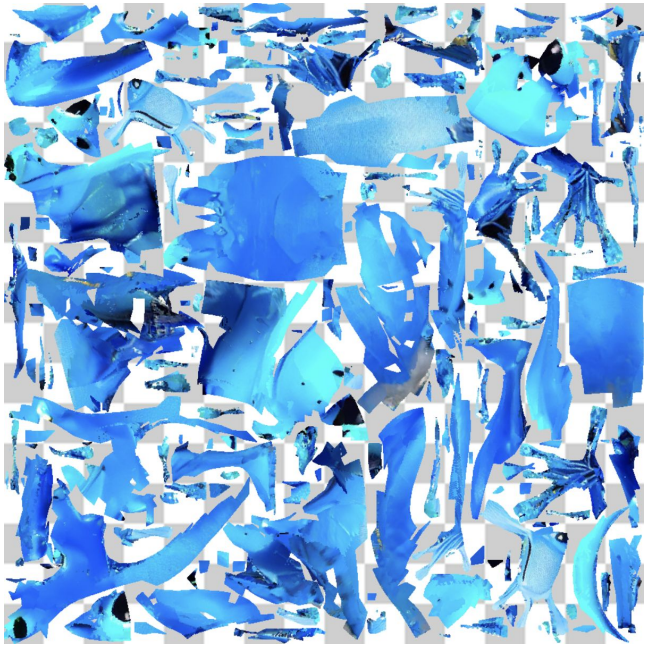
Another Failure Case

- Inconsistency between opposite views

Prompt: Cup

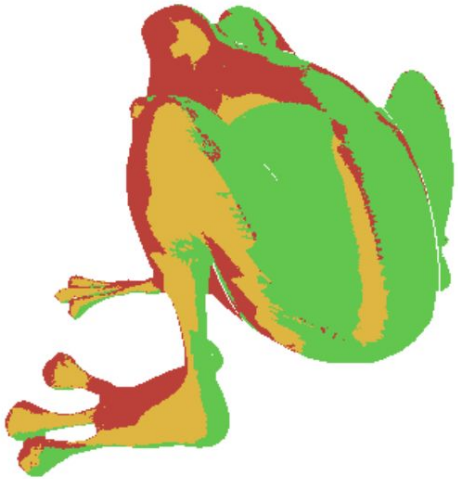
Reasons For Failure

Prompt: Blue Frog (Failure Case)



Reasons For Failure

Prompt: Blue Frog (Failure Case)

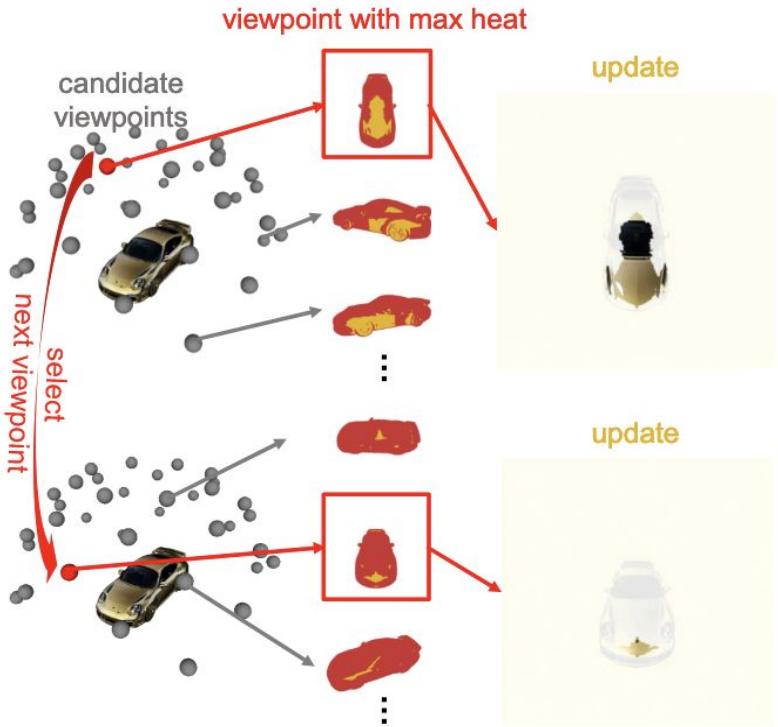


Partitioning masks are wider, smoother

Failure Case

- **Meshes with Complex Geometry**
    - Generates highly partitioned texture maps
- **Meshes with angled geometry**
    - Model cannot connect the different sides of texture
- **Cannot Handle Models with hole inside**
    - Unreachable views cannot be generated

Approaches



- Increase viewpoints for refining
  - Inference time increase
- Refine the correct viewpoints for regions with artifacts!

Previous Approach

View Heat.

- **View Heat Defines the weight for the regions that need update.**

$$\text{view\_heat} = \frac{1}{N} \sum_{i=1}^{N} w_i \cdot 1_{\{\text{quad\_mask}[i]=c_i\}}$$

Limitations:

- **Heuristic Approach: use pre-defined weights for each pixel**
- **Estimate the region for update, induces error**

Approach

We defined new scores for calculating the total score.

$$\text{combined\_score} = \alpha \cdot \text{view\_heat} + \beta \cdot \text{seam\_score} + \gamma \cdot \text{stretching\_score} + \delta \cdot \text{update\_coverage}$$

1) **Seam Score: Difference of RGB values in texture edges**
2) **Stretching Score: Measure the distortion of generated texture**
3) **Update Coverage: Area that the selected viewpoint could cover**

Previous Approach

1) Seam Score: Difference of RGB values in texture edges

$$\text{seam\_score} = \frac{1}{E} \sum_{(u,v)\in\mathcal{E}} \|\text{color}(u) - \text{color}(v)\|_2$$

Calculate the norm of texture rgb value with total edges of texture.

Previous Approach

2) Stretching Score: Measure the distortion of generated texture

$$\text{stretching\_score} = \frac{1}{F} \sum_{t \in \mathcal{T}} \left| \frac{\text{Area}_{\text{UV}}(t)}{\text{Area}_{\text{3D}}(t)} - 1 \right|$$

Measure the area ratio of actual 3d model and the uv map texture.

Previous Approach

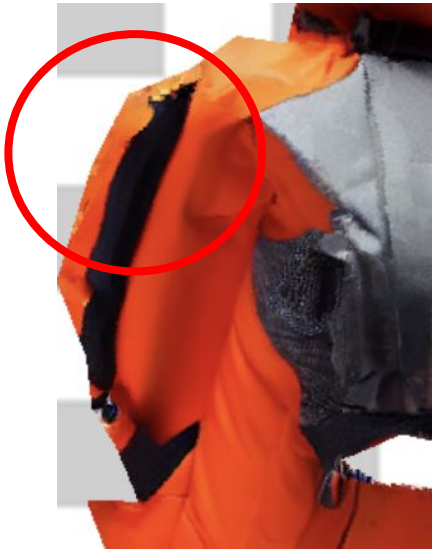
3) Update Coverage: Area that the selected viewpoint could cover

$$\text{update\_coverage} = \frac{1}{N} \sum_{i=1}^{N} 1_{\{\text{quad\_mask}[i]=1\}}$$

Results

Texture map - texture of the side with distortion is more detailed

Before



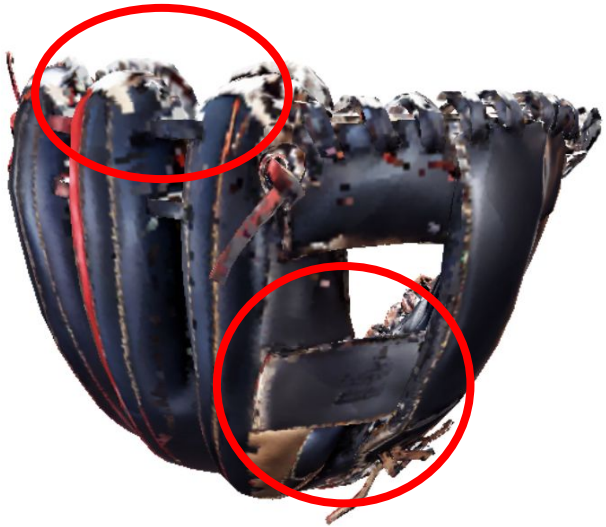
After

$[\alpha, \beta, \gamma, \delta]$ = [1.0, 0.2, -0.3, 0.3]

Results

Rendering - Advantage on coverage issues

Before



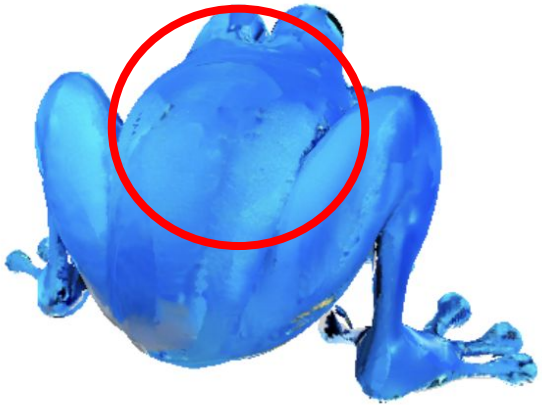
After

$[\alpha, \beta, \gamma, \delta]$ = [1.0, 0.2, -0.3, 0.3]

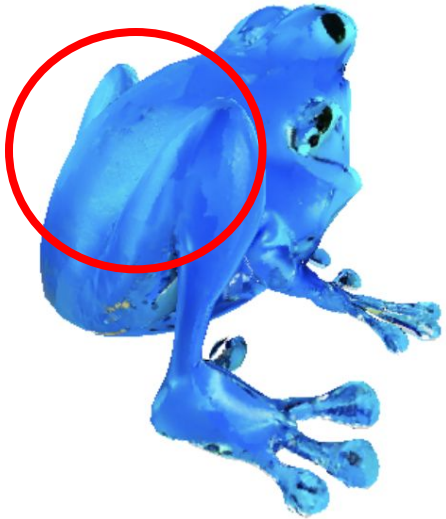
Results

Rendering - natural color connection generated in different viewpoints

Before



After



$[\alpha, \beta, \gamma, \delta]$ = [1.0, 0.2, -0.3, 0.3]

Conclusion & Limitation

Conclusion

- Text2Tex model has some failure cases in particular meshes or materials
- View-heat based refinement is not enough to address the problems
- Adding metrics can benefit the texture refinement scheme without more viewpoints

Limitation

- Failed to find meaningful qualitative metric to describe the refinement
- Still hard to refine difficult meshes with complex surface

Roles

Chanryeol Lee : Design experiments

Evaluate outputs

Presentation

Chanhyeok Lee : Implement score metric

Evaluate outputs

Material preparation

Thank you !