

Projektový seminár (2)

Robert Pagáč

Školiteľ: Mgr. Andrej Mihálik, PhD.

Úvod

Téma: Tvorba textúr pomocou difúzných modelov

Splnené ciele:

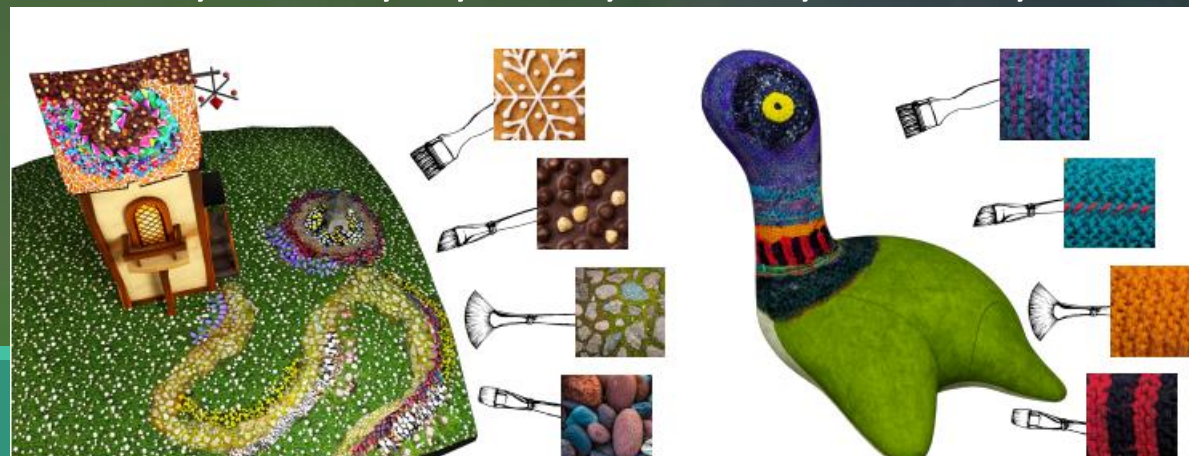
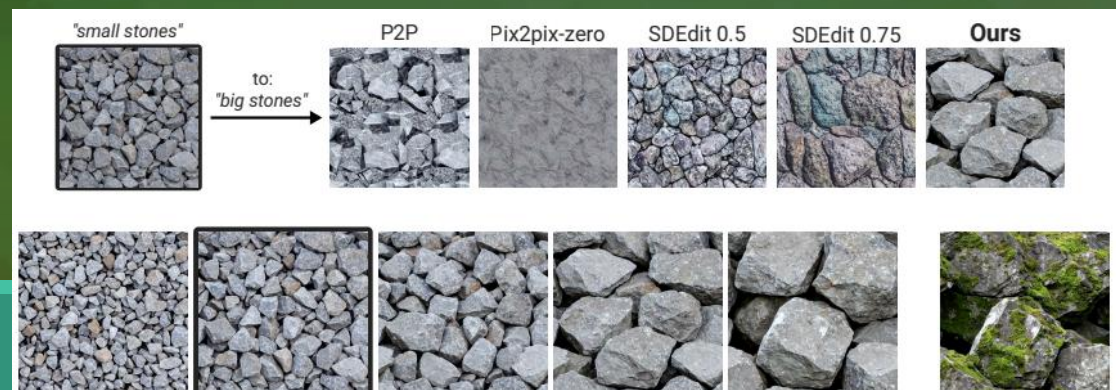
- Naštudovanie 3 vedeckých článkov
- Implementácia generovania promptu pomocou už dostupných difúzných modelov
- GitHub repozitár k diplomovej práci
- Kostra LaTeX diplomovej práce s dvoma vedeckými článkami

Ciele práce

- Preskúmať možnosti generovania a editovania textúr pomocou difúzných modelov.
- Implementovať nástroj využívajúci Stable Diffusion.
- Porovnať rôzne prístupy a techniky editácie textúr.
- Zhodnotiť kvalitu generovaných textúr.
- Aplikovať vygenerované textúry na 3D objekty.

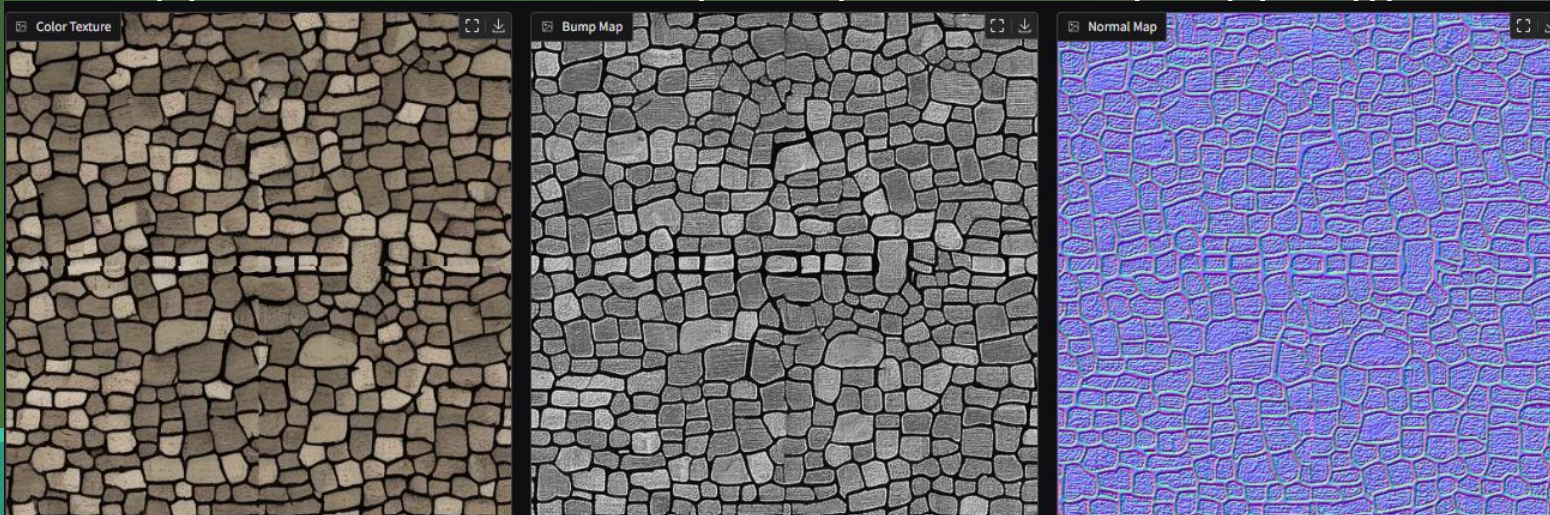
Východiská práce

- **TexSliders:** Tento článok predstavuje metódu na úpravu textúr v CLIP priestore pomocou dvoch textových promptov bez potreby trénovania modelu, čím umožňuje plynulé a kontrolované editovanie vlastností textúr.
- **Diffusion Texture Painting:** Článok opisuje nový spôsob interaktívneho maľovania realistických textúr na 3D modely pomocou difúzných modelov, ktoré generujú plynulé a variabilné „štetce“ z malých ukážok textúr.
- **TexFusion:** Tento článok popisuje metódu na generovanie konzistentných 3D textúr pre celé 3D modely pomocou textovo riadených difúzných modelov. TexFusion využíva viacnásobné pohľady a optimalizáciu v UV priestore, aby zabezpečil rovnaký vizuálny štýl, farby a detaily na všetkých častiach objektu bez švov a artefaktov.



Návrh riešenia

- Architektúra systému: využitie *StableDiffusionXLPipeline* na generovanie textúr a frameworku BentoML na nasadenie modelu. API rozhranie cez BentoML – generovanie obrázkov z textového promptu.
- Webové používateľské rozhranie: jednoduché UI umožňujúce zadávanie promptu, náhľad výsledkov a ovládanie parametrov generovania.
- Export výstupov: generovanie a sťahovanie PNG textúr.
- Doplnkové mapy: automatická tvorba výškovej a normálovej mapy z vygenerovanej textúry.



Implementácia

- využitie Stable Diffusion XL bez nutnosti ďalšieho trénovania, optimalizovaný na generovanie realistických textúr.
- **Inferenčná pipeline:**
textový prompt → syntetická textúra
prompt + vstupný obrázok → upravená alebo odvodená textúra
- generovanie výškovej a normálovej mapy z výslednej textúry pomocou následného spracovania,
- pipeline je volaná cez BentoML endpoint a vizualizovaná vo webovom rozhraní.

```
!pip install gradio diffusers transformers accelerate safetensors torch --quiet

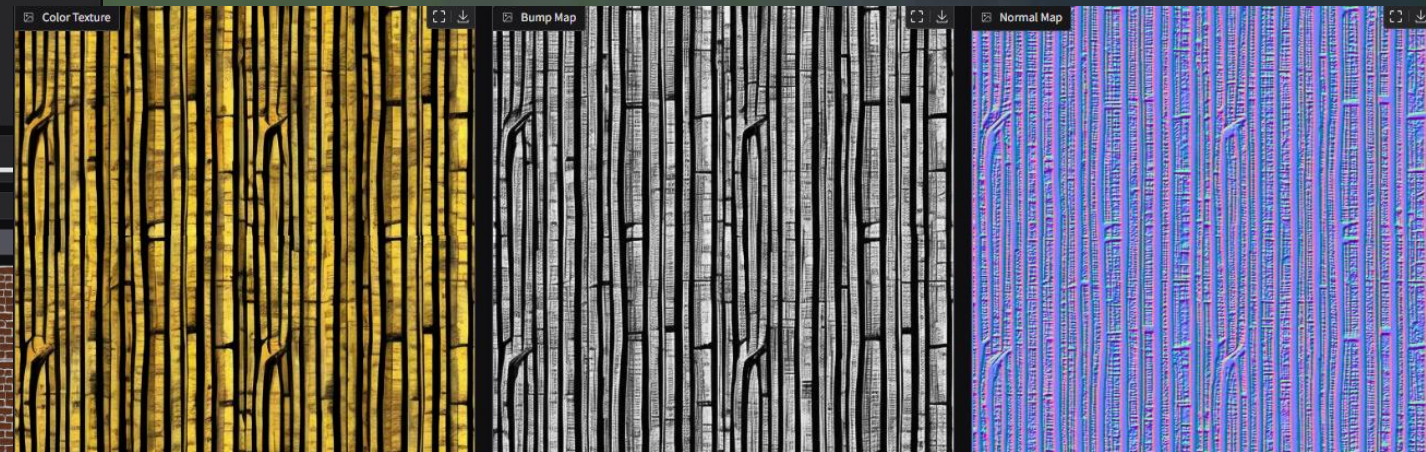
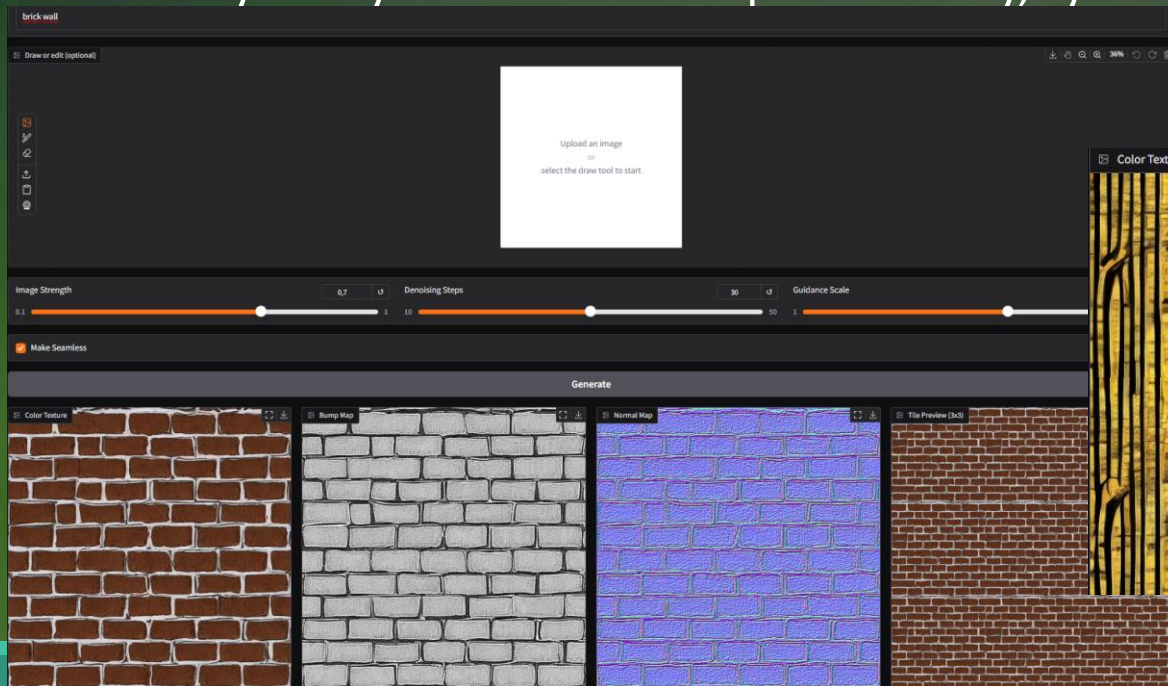
import gradio as gr
from diffusers import StableDiffusionPipeline, StableDiffusionImg2ImgPipeline
import torch
from PIL import Image, ImageChops, ImageFilter, ImageOps
import io, base64, numpy as np

# --- model ---
model_id = "dream-textures/texture-diffusion"
pipe_txt = StableDiffusionPipeline.from_pretrained(model_id, torch_dtype=torch.float16).to("cuda")
pipe_img = StableDiffusionImg2ImgPipeline.from_pretrained(model_id, torch_dtype=torch.float16).to("cuda")

# --- utils ---
def prepare_image(img):
    if img is None:
        return None
    if isinstance(img, dict):
        if "image" in img:
            b64_data = img["image"].split(",")[1]
            img_bytes = base64.b64decode(b64_data)
            img = Image.open(io.BytesIO(img_bytes))
        elif "composite" in img:
            img = img["composite"]
        else:
            return None
    if img is None:
        return None
    if img.mode == "RGBA":
        bg = Image.new("RGB", img.size, (255, 255, 255))
        bg.paste(img, mask=img.split()[3])
```

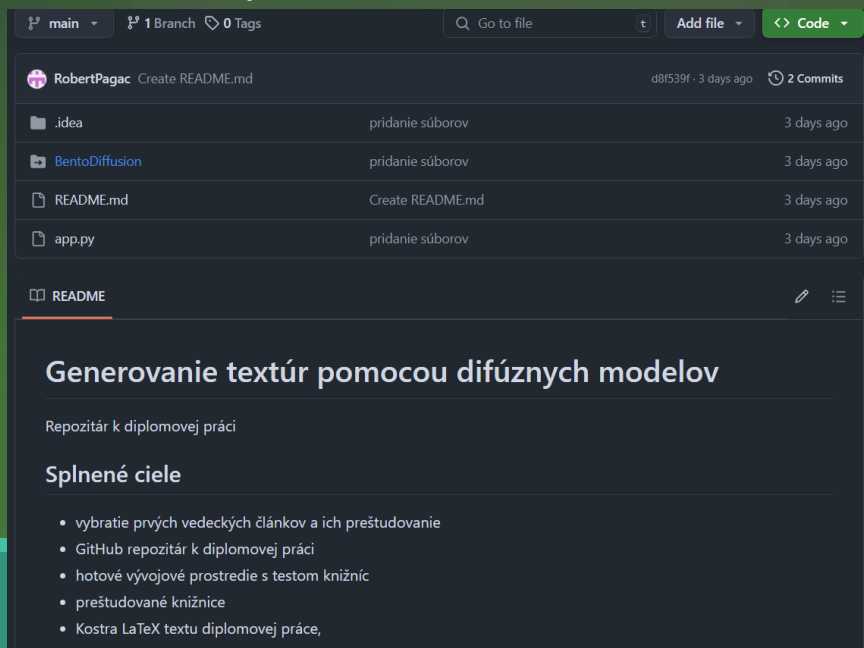

Vyhodnotenie

- Navrhnuté riešenie bolo testované na viacerých textových a img2img promptoch pomocou modelu Stable Diffusion XL.
- Textúry aj doplnkové mapy (bump/normal) sa generovali v rozlíšení 512×512.
- Výsledky boli hodnotené podľa kvality, rýchlosti generovania a toho, ako dobre sa dali použiť



Záver

- Naštudovanie 3 vedeckých článkov
- Implementácia generovania textúr, výškových a normálových máp
- Aktualizovaný GitHub repozitár k diplomovej práci s kódom, článkami a aktuálnou verziou práce
- Prvé strany diplomovej práce, rozpracovaná kapitola výskum kde sú zhrnuté preštudované vedecké články



Kapitola 3

Výskum

V tejto kapitole sú popísané aktuálne prístupy k generovaniu a úprave textúr pomocou difúzných modelov ako aj generovanie doplnkových máp k textúram. Vietky články riešia podobnú oblasť no každý iným spôsobom. Tieto články po kombinujú teóriu generovania textúr s praktickými metódami úprav a tvorby doplnkových máp, ako sú výškové a normálové mapy, čo umožňuje lepšie pochopenie a aplikáciu difúzných modelov. V nasledujúcich častiach sa podrobnejšie rozoberú jednotlivé prístupy, ich silné stránky, ohnadenia a možnosti integrácie do existujúcich grafických postupových línii (pipeline).

3.1 TexSliders: Difúzna úprava textúr v CLIP priestore

V tejto kapitole predstavím prístup, ktorý umožňuje editáciu textúr pomocou difúzných modelov. Cieľom metódy je vytvoriť intuitívny spôsob úprav textúr bez straty ich identity, pričom na to stačia iba jednoduché textové popisy, bez potreby opätovného tréningu modelu.

3.1.1 Motivácia a problém

Tradičné metódy úpravy obrazu, ako sú Prompt-to-Prompt, Pix2Pix-Zero alebo SDEdit, používajú cross-attention mapy na zachovanie identity obrázku počas editácie. Tieto metódy fungujú dobre pre obrázky s jasnou štruktúrou objektov, no v prípade textúr sú neefektívne, pretože textúry sa často nemajú viaceru rôznych objektov, s ktorých sa skladajú. Ako môžeme vidieť na obrázku 3.1 attention mapy nedokážu potom správne zachytiť ich vizuálnu identitu.

TexSliders sa preto sústreďuje na manipuláciu v obrázkovom embedding priestore CLIP modelu, čo umožňuje presnejšie zachovanie identity textúry a generovanie

Ďakujem za pozornosť