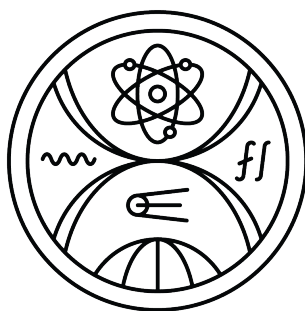


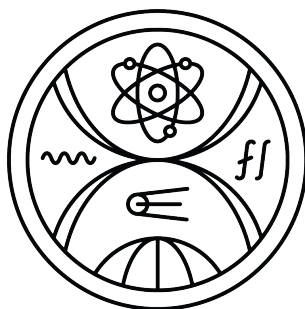
Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky



GENEROVANIE TEXTÚR POMOCOU DIFÚZNYCH MODELOV

Diplomová práca

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky



GENEROVANIE TEXTÚR POMOCOU DIFÚZNYCH MODELOV

Diplomová práca

Študijný program: Aplikovaná informatika
Študijný odbor: Aplikovaná informatika
Školiace pracovisko: Katedra aplikovanej informatiky
Školiteľ: Mgr. Andrej Mihálik, PhD.



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Robert Pagáč
Študijný program: aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský
Sekundárny jazyk: anglický

Názov: Generovanie textúr pomocou difúzných modelov
Texture Generation Using Diffusion Models

Anotácia: Generovanie textúr je kľúčovou súčasťou tvorby obsahu pre hry, filmy a virtuálnu realitu. Tradičné metódy (ručne vytvárané textúry alebo procedurálne algoritmy) sú často časovo náročné a vyžadujú odborné znalosti. Difúzne modely, schopné generovať realistické obrazy na základe textových popisov, ponúkajú možnosť automatizácie a zrýchlenia tohto procesu. Navyše môžu byť použité na tvorbu PBR textúr, čo otvára nové možnosti pre realistické renderovanie a materiálový dizajn.

Vedúci: Mgr. Andrej Mihálik, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.
Dátum zadania: 16.12.2024

Dátum schválenia: 16.12.2024

prof. RNDr. Roman Ďurikovič, PhD.
garant študijného programu

.....
študent

.....
vedúci práce

Prehlasujem, že som diplomovú prácu vypracoval samostatne s využitím získaných teoretických poznatkov a s použitím uvedenej literatúry.

Bratislava, 2026

.....
Bc. Robert Pagáč

Pod'akovanie

...

Abstrakt

...

Klíčové slova:...

Abstract

...

Keywords: ...

Obsah

1	Súčasný stav riešenej problematiky	2
1.1	Zahraničný výskum	2
1.2	Stav výskumu na Slovensku	2
2	Cieľ práce	3
3	Výskum	4
3.1	TexSliders: Difúzna úprava textúr v CLIP priestore	4
3.1.1	Motivácia a problém	4
3.1.2	Prehľad metódy	5
3.1.3	Vlastnosti a výhody	6
3.1.4	Vyhodnotenie a príklady	6
3.1.5	Obmedzenia a diskusia	6
3.1.6	Zhrnutie	7
3.2	Diffusion Texture Painting	7
3.2.1	Úvod	7
3.2.2	Motivácia a problém	8
3.2.3	Súvisiaca práca	8
3.2.4	Metóda	9
3.2.5	Záver	11
4	Metodika práce	12
4.1	Použité metódy	12
4.2	Postup práce	12
5	Návrh softvérového diela	13
5.1	Architektúra systému	13
6	Výsledky práce	14
7	Interpretácia výsledkov a diskusia	15

Zoznam obrázkov

3.1	Ukážka zlyhania tradičných metód pri zachovaní identity textúry [2] . .	5
3.2	Ukážka editácie textúr pomocou TexSliders [2]	7

Zoznam tabuliek

Terminológia

Motivácia

...

Kapitola 1

Súčasný stav riešenej problematiky

1.1 Zahraničný výskum

1.2 Stav výskumu na Slovensku

Kapitola 2

Ciel' práce

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. Všetky 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, obmedzenia a možnosti integrácie do existujúcich grafických postupových línií (pipeline).

3.1 TexSliders: Difúzna úprava textúr v CLIP priestore

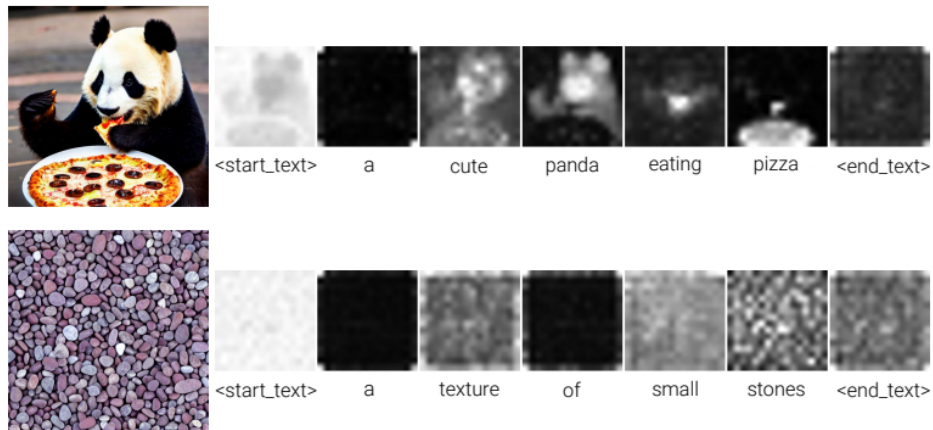
V tejto kapitole predstavím prístup, ktorý umožňuje editáciu textúr pomocou difúzných modelov. Cieľom metódy je vytvoriť intuitívny spôsob úprav textúr bez straty ich identity, pričom na to stačia iba jednoduché textové popisy, bez potreby opätovného trénovania modelu modelu.

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 obrazy s jasnou štruktúrou objektov, no v prípade textúr sú neefektívne, pretože textúry sa často nemajú viacero 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

realistických úprav. [2]



Obr. 3.1: Ukážka zlyhania tradičných metód pri zachovaní identity textúry [2]

3.1.2 Prehľad metódy

Hlavný princíp TexSliders je definícia smerov úprav textúr v embedding priestore, ktoré môžu byť aplikované ako *posuvníky*. Postup je nasledovný:

1. Pre každý textový prompt, pôvodný a cieľový vzhľad, sa generuje množina CLIP image embeddingov pomocou **texture diffusion prior**. Tento krok zabezpečuje, že embeddingy sa nachádzajú v subdoméne textúr.
2. Určí sa počiatočný smer úpravy d' ako rozdiel centroidov týchto embeddingov:

$$d'_j = \frac{1}{n_e} \left(\sum_k t_j^{(k)} - \sum_i o_j^{(i)} \right), \quad (3.1)$$

kde $o^{(i)}$ a $t^{(k)}$ sú embeddingy pôvodného a cieľového promptu a n_e je počet embeddingov.

3. Vyberú sa len relevantné dimenzie, ktoré prispievajú k požadovanej zmene, aby sa minimalizovala strata identity:

$$d_j = \begin{cases} d'_j & \text{ak } |\tilde{d}'_j| > \tau \cdot \text{std}_j \\ 0 & \text{inak} \end{cases} \quad (3.2)$$

kde τ je prahový parameter a std_j je štandardná odchýlka dimenzie.

4. Výsledný smer d sa aplikuje na embedding vstupnej textúry e_0 :

$$e_\alpha = e_0 + \alpha \cdot d, \quad (3.3)$$

kde α určuje intenzitu úpravy. Upravený embedding sa následne použije na generovanie finálnej textúry difúznym modelom. [2]

3.1.3 Vlastnosti a výhody

TexSliders umožňuje:

- vytvárať **textové posuvníky** pre ľubovoľné úpravy textúr v priebehu niekoľkých minút,
- zachovať **identitu textúry** aj pri extrapolovaní úprav,
- kombinovať viacero smerov úprav súčasne,
- použiť metódu na syntetické aj reálne textúry.

V porovnaní s existujúcimi metódami dosahuje TexSliders lepšiu presnosť editácie a zachovanie identity textúry, čo dokazuje aj kvantitatívna analýza v CLIP priestore. [2]

3.1.4 Vyhodnotenie a príklady

Na vyhodnotenie bola použitá testovacia sada pozostávajúca z rôznych typov materiálov ako sú drevo, kameň, kov, textil alebo papier a atribútov ako sú napríklad hladký, tvrdý alebo starý.

Výsledky ukazujú, že TexSliders dokáže:

- zvládnuť globálne úpravy celých textúr (napr. „small stones“ → „big stones“),
- zachovať vizuálne charakteristiky pôvodnej textúry,
- kombinovať viaceré smery úprav (napr. „scale“ + „mossiness“).

3.1.5 Obmedzenia a diskusia

Medzi hlavné obmedzenia patrí:

- citlivosť na niektoré koncepty v CLIP priestore (napr. orientácia vzorov),
- možná strata identity pri extrémnom extrapolovaní úprav,
- zameranie na farebné textúry; generovanie materiálových máp je možné, no limitované dostupnými dátami.

Budúca práca môže skúmať:

- rozšírenie na materiálové mapy ako normalové alebo výškové
- hlbšiu analýzu identity textúr v embedding priestore,
- integráciu do plnohodnotného používateľského rozhrania pre interaktívne a jednoduchšie editovanie. [2]



Obr. 3.2: Ukážka editácie textúr pomocou TexSliders [2]

3.1.6 Zhrnutie

TexSliders predstavuje jednoduchý a efektívny spôsob editácie textúr pomocou difúzných modelov a CLIP embeddingov. Umožňuje vytváranie **sliderov**, ktoré zachovávajú identitu textúry, podporujú kombináciu viacerých atribútov. Sú použiteľné na syntetické aj reálne textúry, bez potreby retrainingu modelu či anotovaných dát. Ako aj autori v článku spomínali momentálne nepodporujú

3.2 Diffusion Texture Painting

3.2.1 Úvod

Difúzne modely úplne zmenili spôsob, akým sa generujú obrázky, a v posledných rokoch sa stali základom pre množstvo aplikácií. Ich schopnosť generovať kvalitné a rozmanité výstupy z nich robí ideálny nástroj nielen pre tvorbu 2D obrázkov, ale aj pre nové interaktívne pracovné postupy. Tento článok predstavuje metódu *Diffusion Texture*

Painting, ktorá umožňuje umelecky maľovať realistické textúry priamo na povrch 3D objektov pomocou generatívnych difúzných modelov.

Kľúčová myšlienka je jednoduchá: používateľ poskytne malý vzor, napríklad kúsok trávy, dreva alebo látky, a systém z neho vytvorí štetec, ktorý následne dokáže interaktívne a v reálnom čase generovať textúru priamo na 3D sieť. Každý odtlačok počas ťahu štetcom je vygenerovaný difúznym modelom s ohľadom na už predchádzajúce namaľované oblasti, čím sa zabezpečí plynulé a bezšvové nadväzovanie textúry. [3]

3.2.2 Motivácia a problém

Autori riešia viacero problémov, ktoré doteraz brzdili realistickú maľbu textúr na 3D objekty:

- **Stabilná identita štetca** – difúzny model musí dodržiavať charakter textúry odvodený zo vstupného vzorku a neodbiehať k neželaným halucináciám.
- **generovanie na základe obrazových výrezoch (patch-based)** – každá pečiarka vzniká lokálne, ale musí dokonale zapadať do predošlých častí maľby.
- **Aplikácia v 3D priestore** – generovanie prebieha v lokálnom kamerovom priestore orientovanom podľa povrchu 3D siete a pohybu používateľa.
- **Reálne nové prechody textúr** – model dokáže vytvárať prechody medzi úplne inými textúrami (napr. tráva → kamene), čo tradičné textúrovacie systémy nedokázali.

Výsledkom je plne interaktívny systém, ktorý využíva generatívny model bez potreby dodatočného tréningu a funguje s rôznymi typmi textúr. [3]

3.2.3 Súvisiaca práca

Digitálne maľovanie

Tradičné 2D a 3D aplikácie na maľovanie ako Mari, Substance Painter alebo 3D Coat používajú štandardný pečiatkový prístup, kde sa obrázok štetca periodicky aplikuje na povrch. Tento proces síce funguje dobre pre jednoduché textúry, ale zlyháva pri komplexných vzorkách. Používateľ potom musí ručne riešiť prechody medzi rôznymi oblasťami, rotácie vzorov, či vyhladzovanie prechodov medzi jednotlivými časťami.

Predchádzajúce systémy založené na príkladoch pracovali s patch-based syntézou, kde sa kopírovali úseky textúry z jedného exemplára. Autori tohoto článku chceli ale tento prístup vylepšiť a namiesto kopírovania systém generuje nové variácie textúry pomocou difúzných modelov a dokáže dopĺňať realistické prechody bez špeciálneho označovania alebo manuálnych úprav. [3]

Syntéza textúr

Klasické metódy textúrovej syntézy ako napríklad Efros–Freeman alebo PatchMatch vytvárajú väčší obrázok spájaním výrezov z malej ukážky. Tieto metódy sú však obmedzené len na klonovanie zdrojového vzorca. Moderné neurónové prístupy podporujú väčšiu variabilitu, ale väčšinou fungujú len v uzavretej doméne dát, na ktorých boli trénované.

Naopak, difúzne modely s otvorenou doménou dokážu pracovať so širokým spektrom textúr a vytvárať inteligentné prechody aj medzi textúrami, ktoré nikdy nevideli pri tréningu. [3]

Textúrovanie 3D objektov

Existuje množstvo prístupov na automatické textúrovanie, vrátane metód založených na GANoch a novších difúzných systémoch schopných generovať celé materiálové mapy. Tento článok sa však špecializuje na situáciu, kde používateľ chce mať okamžitú kontrolu nad výsledkom a nechce sa spoliehať na plnú automatizáciu.

Zvolený prístup používa klasické UV mapovanie, keďže ide o najbežnejší spôsob textúrovania v herných a vizuálnych pipeline. Autori ukazujú, že metóda funguje aj pri komplikovaných typológiách alebo vysokých zakriveniach. [3]

3.2.4 Metóda

Požiadavky na generátor

Pre funkčné interaktívne maľovanie musí uzavretý generátor, teda difúzny model, spĺňať:

- **Podporu rekonštrukcia chýbajúcich častí obrazu** – model musí vedieť doplniť chýbajúce časti obrazu.
- **Kontrolu identity štetca** – musí rešpektovať vzor a generovať konzistentné výstupy.
- **Rýchlosť** – generovanie jednej pečiatky musí byť v interaktívnom čase.

Autori definujú generátor ako funkciu

$$G(I, b) \rightarrow I'$$

kde I je RGBA záplata, b je informácia o štetci a výstup I' je farebná záplata s doplnenými pixelmi. [3]

2D pačiatkové maľovanie

V 2D priestore prebieha maľovanie tak, že každých pár pixelov pohybu kurzora sa vytvorí nová pečiatka. Záplata sa následne vystrihne z plátna, pošle sa do generátora a výsledok sa naspäť zapíše do obrazu.

Kľúčové prvky, ktoré autori spomínali sú:

- prekrývanie pečiatok zabraňuje vzniku švov,
- maska umožňuje prepisovať aj existujúcu textúru,
- všetky pečiatky zdieľajú rovnaké parametre štetca. [3]

Maľovanie na 3D sieť

Pri práci v 3D sa generovanie záplat nevykonáva v perspektíve hlavnej kamery, ale z pohľadu **lokálnej kamery orientovanej podľa normály povrchu**. Toto rieši problémy s deformáciou pri jednoduchom projekčnom maľovaní.

Teda postup je nasledovný:

1. Raycastingom sa nájde bod na povrchu a jeho normála.
2. Vytvorí sa lokálna ortografická kamera orientovaná kolmo na povrch.
3. Vyrenderuje sa malý záplata textúry z lokálnej perspektívy.
4. Záplata sa pošle do difúzneho modelu na následné dopĺňanie obrazu.
5. Výsledok sa *back-projektuje* späť do UV textúry pomocou reverzného renderovania.

Použitím rovnakého renderovacieho jadra pre projekciu aj spätnú projekciu sa zabezpečí presné mapovanie farieb do UV priestoru bez ohľadu na topológiu 3D siete. [3]

Veľkosť štetca

Veľkosť štetca je riadená parametrom β , ktorý určuje šírku záberu lokálnej kamery. Rozlíšenie generátora je síce pevne dané, no autor uvádza, že s neustále rýchlejšími difúznymi modelmi by sa v budúcnosti mohlo meniť podľa priblíženia alebo detailov. [3]

3.2.5 Záver

Diffusion Texture Painting je prvý systém, ktorý využíva difúzne modely s otvorenou doménou na plne interaktívne maľovanie textúr na 3D objektoch. Ako bolo spomenuté umožňuje:

- generovať neobmedzené variácie textúr,
- plynulo prepájať úplne odlišné textúry a materiály,
- maľovať priamo na 3D sieť s konzistentným lokálnym priestorom,
- realizovať proces úpravy maľovania, ktorý bol doteraz nemožný.

Výsledky ďalej naznačujú, že generatívne modely majú obrovský potenciál nielen v tradičnom renderovaní, ale aj v interaktívnych umeleckých nástrojoch, kde je rýchlosť, konzistentnosť a kreatívna kontrola nevyhnutná. [3]

Kapitola 4

Metodika práce

4.1 Použité metody

4.2 Postup práce

Kapitola 5

Návrh softvérového diela

5.1 Architektúra systému

Kapitola 6

Výsledky práce

Kapitola 7

Interpretácia výsledkov a diskusia

Manuál

Literatúra

- [1] Tianshi Cao, Karsten Kreis, Sanja Fidler, Nicholas Sharp, and Kangxue Yin. Texfusion: Synthesizing 3d textures with text-guided image diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. Oral.
- [2] Julia Guerrero-Viu, Milos Hasan, Arthur Roullier, Midhun Harikumar, Yiwei Hu, Paul Guerrero, Diego Gutierrez, Belen Masia, and Valentin Deschaintre. Texsliders: Diffusion-based texture editing in clip space. In *ACM SIGGRAPH 2024 Conference Papers*, 2024.
- [3] Anita Hu, Nishkrit Desai, Hassan Abu Alhaija, Seung Wook Kim, and Maria Shugrina. Diffusion texture painting. In *ACM SIGGRAPH 2024 Conference Papers*, 2024.