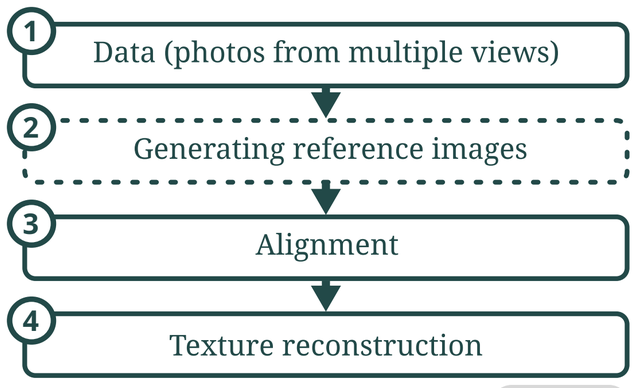
TEXTURE RECONSTRUCTION FROM MULTIPLE VIEWS

Author: *Bc. David Kedra (xkedra00@vutbr.cz)*
Supervisor: *Ing. Michal Hradiš, Ph.D.*

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF INFORMATION TECHNOLOGY

Excel@FIT 2025

A deep learning model for high-resolution **texture reconstruction** and image **enhancement** of flat surfaces using images captured from multiple viewpoints under varying lighting conditions. The views are aligned to a reference image (which contains the required texture colors) using **homography** and SEA-RAFT **optical flow**. The reconstruction is performed by **CNNs,** mainly based on **U-Net** architectures, and the models are trained primarily on synthetic data generated in Blender.



Scheme 1

Figure 1

1 Synthetic & real data

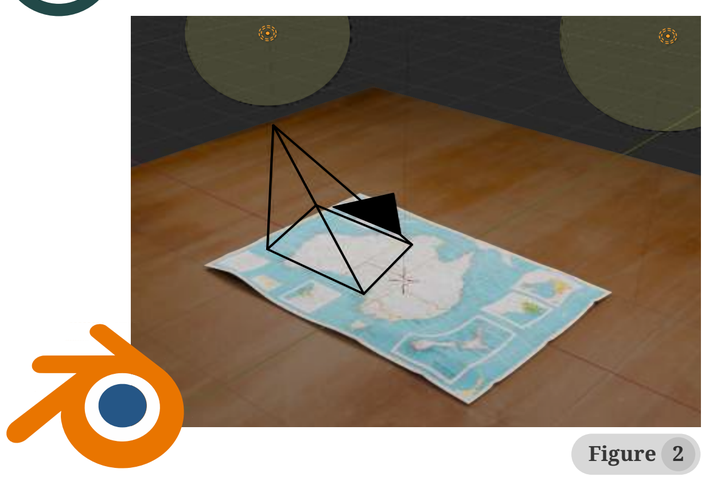


Figure 2

2 Generating low-resolution reference images using realistic degradation of GT textures

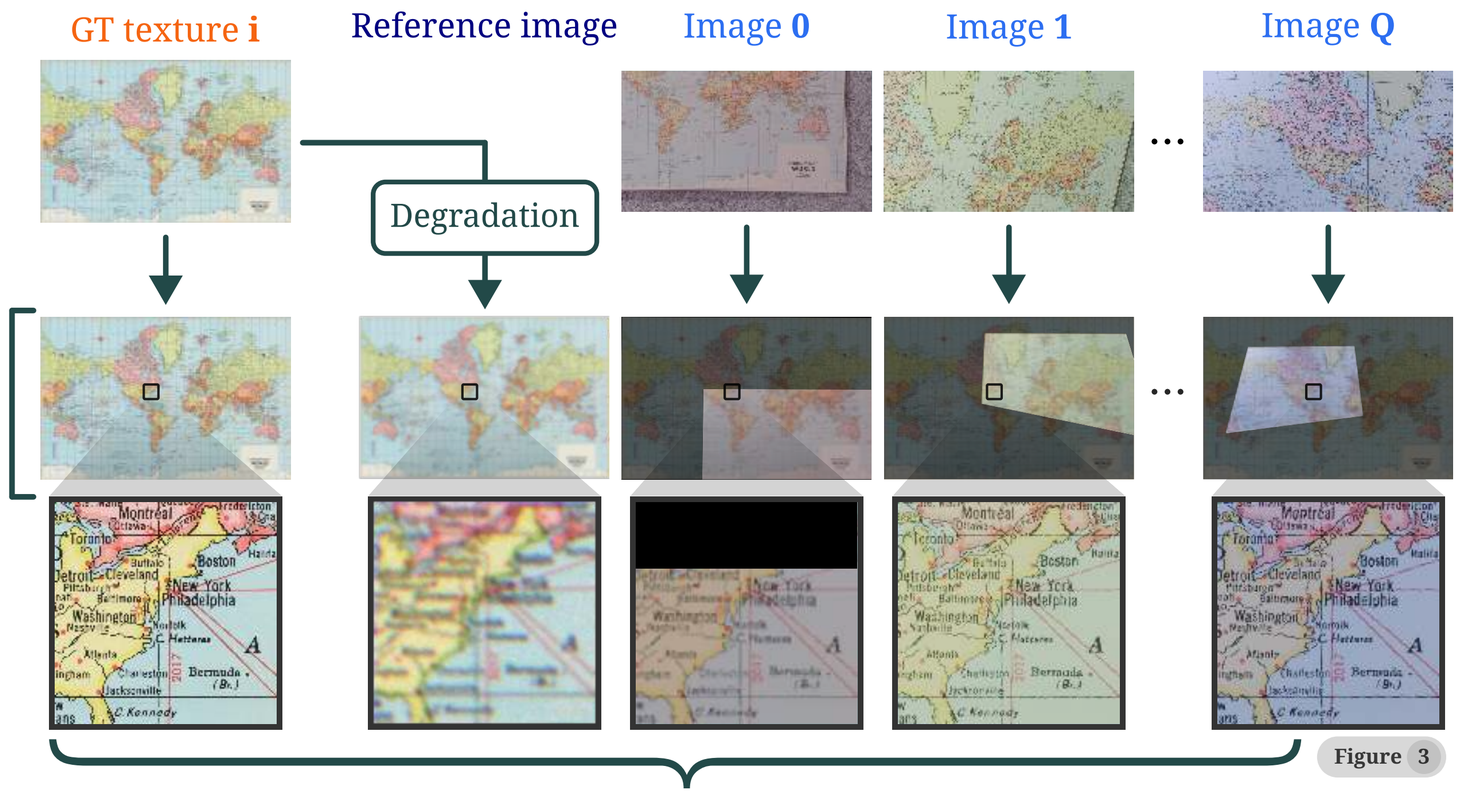


Figure 3

3 Accurate alignment to the reference view

- UV maps (synthetic images for training)
- homography + optical flow (real photos)

Splitting images into patches (512×512 pixels)

Each training sample consists of 1 GT texture patch, 1 reference view patch, and other N view patches selected from a set of Q aligned input images (views) at the specific position (for the project N was set as 5).

4 Texture reconstruction (reference image enhancement)

Table 1: The models are named by encoder-decoder types. MV_EDSR denotes a multi-view variant of EDSR.

	Scene A		Scene B	
	PSNR	SSIM	PSNR	SSIM
MV_EDSR	21.44	0.58	22.06	0.74
ResNet34_U-Net	21.83	0.60	22.02	0.74
MiT_U-Net	22.31	0.61	22.47	0.74
ResNet101_UNet++	22.01	0.61	22.76	0.75

Table 1

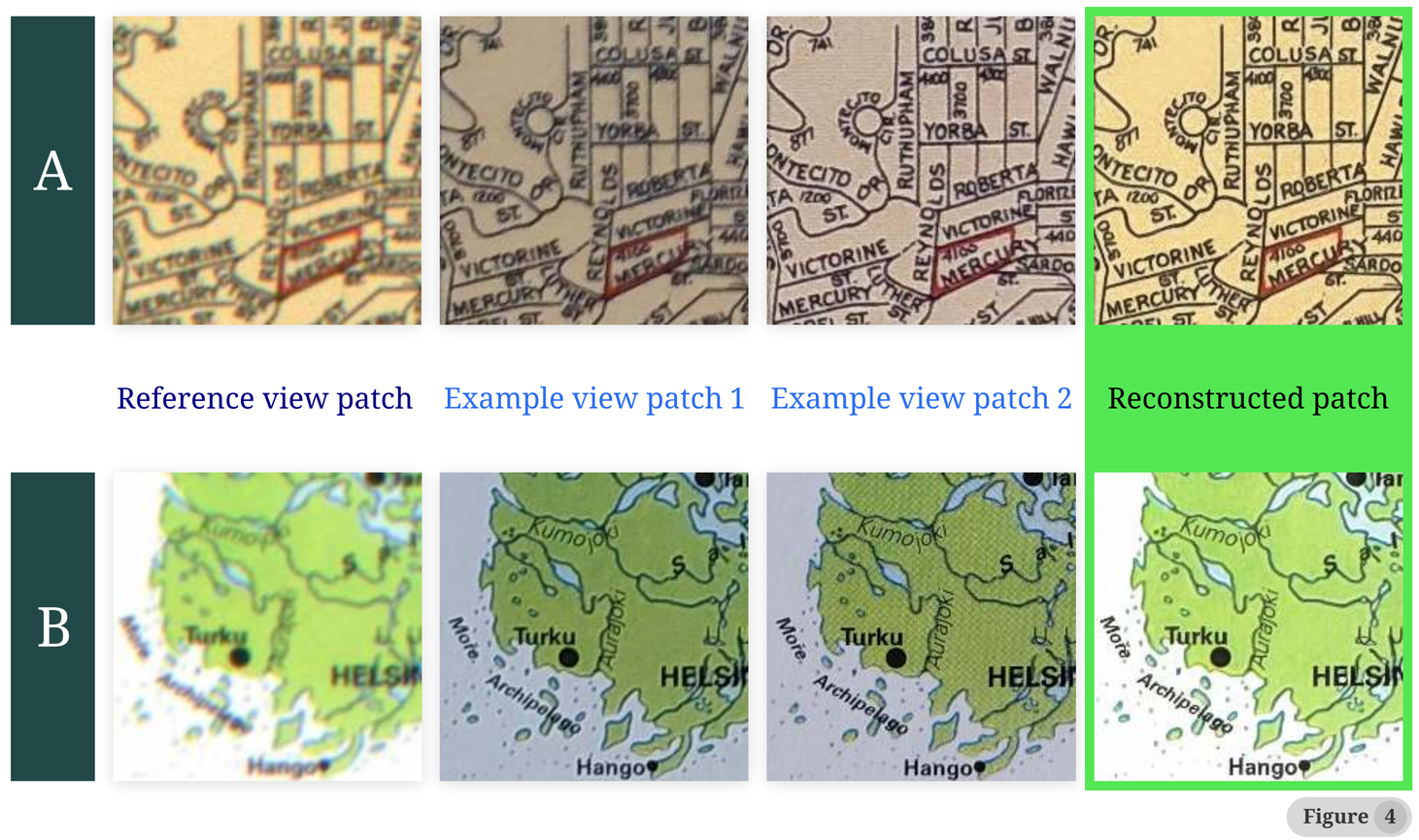


Figure 4