THE 67TH INTERNATIONAL

OPEN READINGS

CONFERENCE FOR STUDENTS OF PHYSICS AND NATURAL SCIENCES

BOOK OF ABSTRACTS | 2024

Vilnius University

VILNIUS UNIVERSITY PRESS

Editors:

Martynas Keršys
Rimantas Naina
Vincentas Adomaitis
Emilijus Maskvytis

Cover and Interior Design:

Goda Grybauskaitė

Vilnius University Press
9 Saulėtekio Av., III Building, LT-10222 Vilnius
info@leidykla.vu.lt, www.leidykla.vu.lt/en/
www.knygynas.vu.lt, www.journals.vu.lt

Bibliographic information is available
on the Lithuanian Integral Library Information System (LIBIS) portal www.ibiblioteka.lt
ISBN 978-609-07-1051-7 (PDF)

CREATING 3D MODELS OF NATURAL OBJECTS FROM 2D IMAGES USING MACHINE LEARNING

Andrius Jedik[1], Donatas Narbutis[1]

[1]Faculty of Physics, Vilnius University
andrius.jedik@ff.stud.vu.lt

This work explores approaches to the structure-from-motion problem [1] by employing well-established methods to reconstruct 3D objects from various selected yet limited 2D image sets. We utilize synthetically generated views (Fig. 1a) of textured meshes rendered with PyTorch3D [2] with variable image quality, as well as photographs of objects from the natural world (Fig. 1b).

Reconstructing 3D meshes from any given image set requires knowledge of the camera positions from which the shots were taken. Neither of the selected approaches to create datasets (synthetic or natural) did not involve a way of logging camera positions. To locate those positions, COLMAP package [3] was used to select intrinsic features of images, mark them as points within the pictures, cross-match them, and find a global solution (Fig. 1c). However, achieving a successful reconstruction of any given scene is not straightforward. Therefore, this work investigates the minimal quantity and quality of images required in various controlled synthetically generated datasets to ensure a robust reconstruction. Using the extracted camera view positions, we employ the Neural Radiance Field method (NeRF) to synthesize novel views and generate a point cloud much more detailed than that obtained by COLMAP.

The model used in this study is similar to the one presented in the original paper [4] introducing the concept of NeRFs. Such model utilizes a volumetric method to project rays out of the picture, interacting with pixel data, to reconstruct object
's density in 3D space. Synthetically generated models with recovered camera positions were employed to recreate 3D meshes (Fig. 1d) and explore the results achievable with limited quality and quantity image sets. Finally, natural world images were used to create acceptable quality meshes and 3D scene reconstructions.

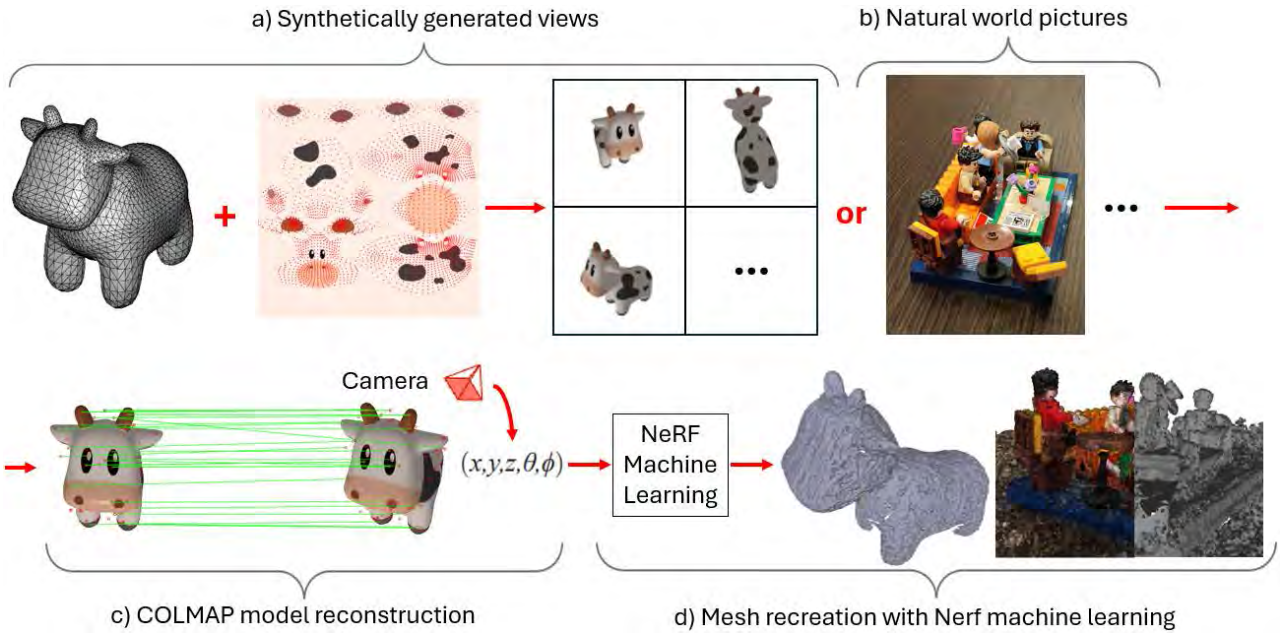

Fig. 1. Simplified workflow of 3D mesh reconstruction. a) Using Pytorch3D selected picture texture is transformed around mesh vertice map to create 3D object mesh and a set of views. b) Selected natural world object is photograped. c) Using COLMAP before mentioned sets of images are reconstructed and their camera view positions are extracted. d) Using NeRF machine learning and given coordinates of a picture set camera views, 3D meshes are generated.

[1] Hartley, R. I., Zisserman, A. (2004). A Brief Overview of Structure from Motion. International Journal of Computer Vision.
[2] Ravi, N. et al. (2020). Accelerating 3D Deep Learning with PyTorch3D. arXiv:2007.08501.
[3] Schönberger, J. L., Frahm, J.-M. (2016). Structure-from-Motion Revisited. In Conference on Computer Vision and Pattern Recognition (CVPR).
[4] Mildenhall, B. et al. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In ECCV.