Plan Overview

A Data Management Plan created using DMPonline.be

Title: Mega-Scale Structure-from-Motion in Extremely Challenging Environments

Creator: Seong Hun Lee

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Project abstract:

Imagine autonomous robots with onboard cameras freely navigating in forests, underwater, underground, Moon and Mars terrain. Imagine how many disasters we can prevent through monitoring, how many lives we can save through search & rescue operations, and how many secrets of the unexplored territories we can unveil. One of the key technologies that will make this happen is Structure-from-Motion (SfM). This fundamental problem in computer vision addresses the task of reconstructing camera poses and scene structure from a set of unordered images. Unfortunately, the deployment of the aforementioned applications is hindered by the current limitations of SfM. In this project, I propose a novel approach for SfM that is robust to extremely challenging environments, and at the same time, scalable to millions of images. The main challenges are that (1) the images collected in the wild are prone to a large fraction of outliers and (2) reconstructing millions of images is very time-consuming, especially when there are outliers. I propose to tackle these challenges through the following efforts: (1) Redesigning the pipeline in an innovative way, (2) enhancing some of the components significantly compared to the state of the art, and (3) thereby removing the need for the full bundle adjustment at the end of the pipeline. Specifically, this project will involve multiple novel contributions in relative pose estimation, rotation/translation averaging, and rotation/pose-only optimization.

ID: 211439

Last modified: 03-12-2024

Mega-Scale Structure-from-Motion in Extremely Challenging Environments

Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data) or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
Synthetic dataset	Synthetic dataset for testing SfM systems	N	D	N	txt	<1GB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, codebook.tsv etc. where this information is recorded).

I will add README.txt files where all necessary information is recorded.

Will a metadata standard be used to make it easier to find and reuse the data?
If so, please specify which metadata standard will be used.

If not, please specify which metadata will be created to make the data easier to find and reuse.

- Yes

DataCite

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Other (specify below)

Hard drive on my work laptop + GitHub repository

How will the data be backed up?

- Other (specify below)

GitHub repository

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data access will be made available only to me and my advisor.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

Negligible

Data Preservation after the end of the Research Project

Which data will be retained for 10 years (or longer, in agreement with other retention policies that are applicable) after the end of the project?

In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

Where will these data be archived (stored and curated for the long-term)?

- Other (specify below)

Public repository

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

NA

Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project?
Please explain per dataset or data type which data will be made available.

- Yes, as open data

If access is restricted, please specify who will be able to access the data and under what conditions.

NA

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

Please explain per dataset or data type where appropriate.

- No

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

GitHub

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- GNU GPL-3.0 (code)

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- No

What are the expected costs for data sharing? How will these costs be covered?

Negligible

Responsibilities

Who will manage data documentation and metadata during the research project?

Myself

Who will manage data storage and backup during the research project?

Myself

Who will manage data preservation and sharing?

Myself

Who will update and implement this DMP?

Myself