



Bachelorarbeit

# **Realistic MVS dataset**

Eberhard Karls Universität Tübingen  
Mathematisch-Naturwissenschaftliche Fakultät  
Wilhelm-Schickard-Institut für Informatik  
Computergrafik  
Peter Trost, [peter.trost@student.uni-tuebingen.de](mailto:peter.trost@student.uni-tuebingen.de), 2019

Bearbeitungszeitraum: von-bis

Betreuer/Gutachter: Prof. Dr. Hendrik Lensch, Universität Tübingen  
Zweitgutachter: Prof. Dr. Max Mustermann, Universität Tübingen



# Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe und dass alle Stellen, die dem Wortlaut oder dem Sinne nach anderen Werken entnommen sind, durch Angaben von Quellen als Entlehnung kenntlich gemacht worden sind. Diese Bachelorarbeit wurde in gleicher oder ähnlicher Form in keinem anderen Studiengang als Prüfungsleistung vorgelegt.

---

Peter Trost (Matrikelnummer 4039682), April 20, 2019



# Abstract

Template



# Acknowledgments

If you have someone to Acknowledge ;)





# Contents

|   |           |
|---|-----------|
| <b>1. Introduction</b>  | <b>11</b> |
| <b>2. Related Work</b>  | <b>13</b> |
| 2.1. Synthetically rendered datasets . . . . .  | 13        |
| 2.1.1. A naturalistic open source movie for optical flow evaluation .   | 13        |
| 2.1.2. Playing for data: Ground truth from computer games . . . . .   | 15        |
| 2.2. Overview . . . . .   | 15        |
| 2.2.1. The synthia dataset: A large collection of synthetic images for<br>semantic segmentation of urban scenes . . . . . | 17        |
| 2.2.2. SyB3R: A Realistic Synthetic Benchmark for 3D Reconstruction<br>from Images . . . . .                              | 18        |
| 2.3. Problem Statement . . . . .  | 19        |
| <b>3. Conclusion</b>  | <b>21</b> |
| <b>A. Blub</b>  | <b>23</b> |



# 1. Introduction

What is this all about?

Cite like this: [AFS<sup>+</sup>11]



## 2. Related Work

### 2.1. Synthetically rendered datasets

#### 2.1.1. A naturalistic open source movie for optical flow evaluation

[BWSB12]

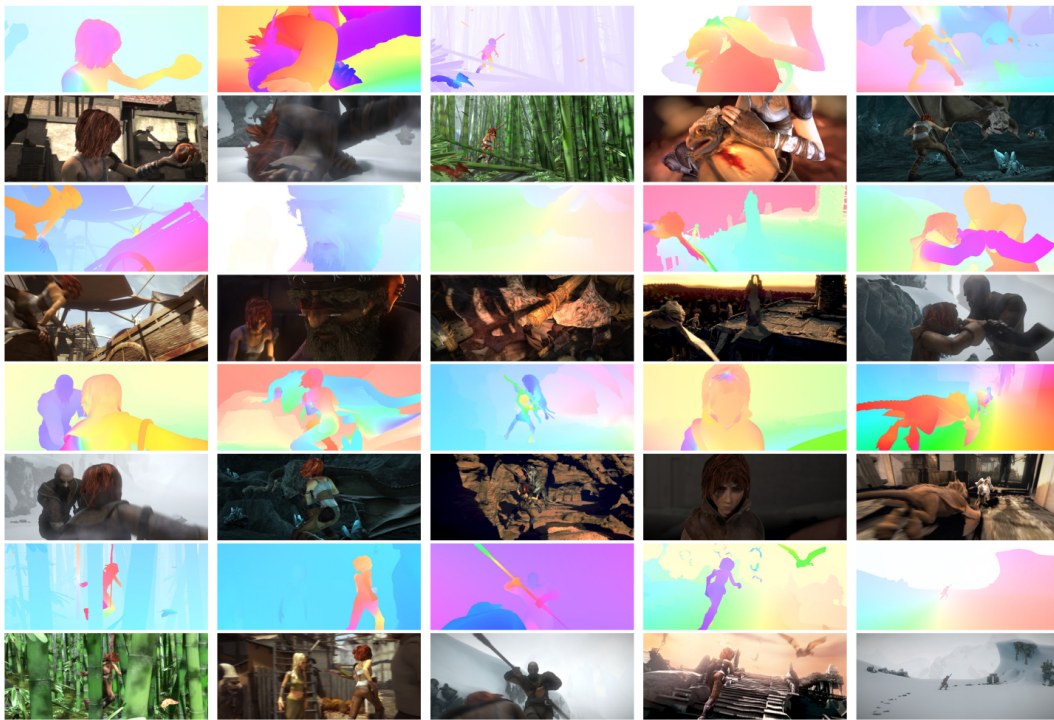


Figure 2.1.: Example images (ground truth flow in uneven rows and RGB in even ones) from the Sintel dataset

#### Overview

In this paper the authors provide a dataset for optical flow estimation derived from the open source 3D animated short film Sintel **TODO: cite Sintel: <https://durian.blender.org/>.** The dataset contains long sequences, large motions, specular reflections, motion blur, defocus blur, atmospheric effects and more. Its scenes are rendered in varying

## Chapter 2. Related Work

complexity through the source graphics data provided by the authors of the film. Because of this aforementioned variety the dataset can be used to improve optical flow methods.

### **Render passes**

As mentioned above the dataset contains image sequences rendered in the following varying complexity:

- Albedo Pass: Flat and unshaded. Surfaces exhibit constant albedo over time
- Clean Pass: Illumination including smooth shading and specular reflections adds realism
- Final Pass: Full rendering with all effects including blur due to camera depth of field and motion, and atmospheric effects.

### **Main aspects**

The main aspects of the Sintel dataset are the following:

It contains varying and more challenging (for existing methods) scenes than older datasets. Sequences are 50 frames long and are provided with 49 ground truth flow fields which are a measure of changes in position for objects in the scene from frame to frame. Some frames include motions of well over 100 pixels. There are 1628 frames with 564 for testing and 1064 for training. The Sintel dataset contains sequences having real-world challenges like lighting variations, shadows, complex materials, reflections and more.

### **Meta**

The authors modified Blender's internal motion blur pipeline to give accurate motion vectors at each pixel which provide ground truth optical flow maps. Although the clips are selected so that optical flow is realistic, one still has to be cautious when training and evaluating algorithms that strongly rely on real-world laws of physics. The images are saved as 8-bit PNG files and the clips have a framerate of 24 fps.

### 2.1.2. Playing for data: Ground truth from computer games

[RVRK16]

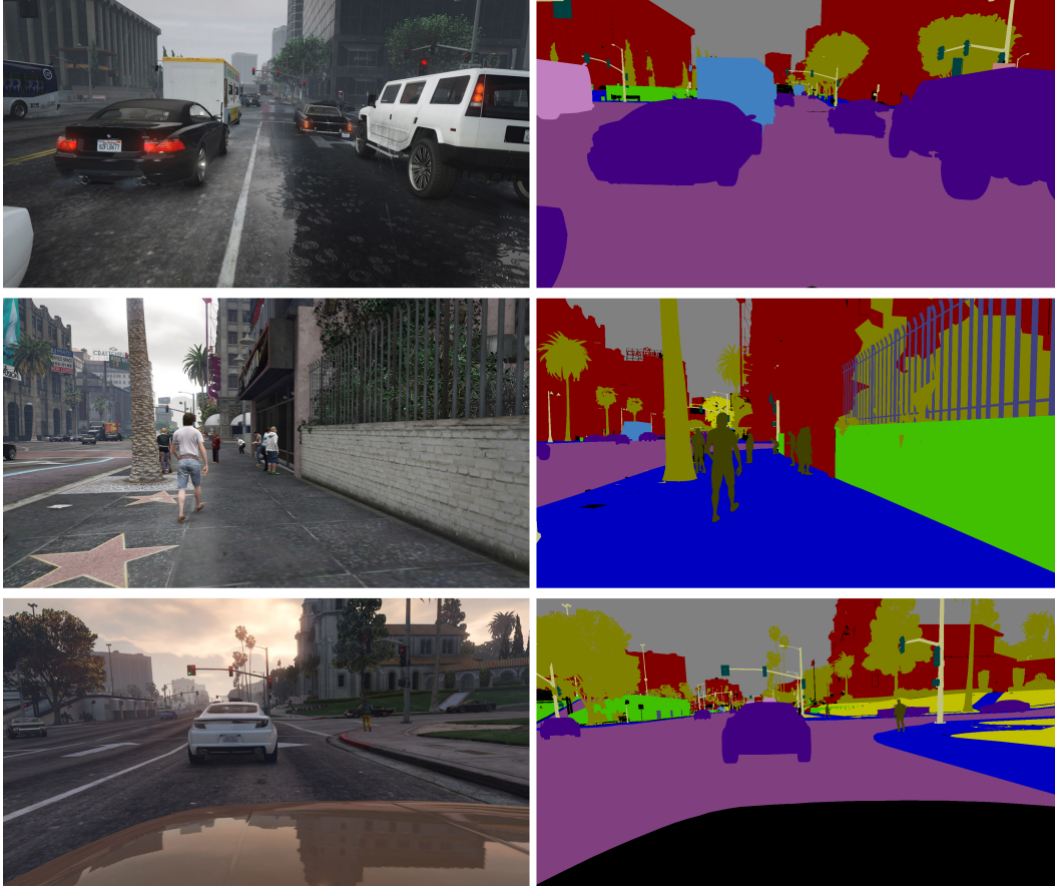


Figure 2.2.: Example images (RGB on the left and semantic segmentation right) from the "Playing for data"-dataset

## 2.2. Overview

The main aspect of this paper is to get pixel-accurate ground truth of synthetic data and therefore be able to label objects in the images accurately and efficiently. The authors use detouring (i.e. injecting a wrapper between the game and the operating system) to record, modify and reproduce rendering commands from the game Grand Theft Auto 5 (GTA5). They retrieve the distinct rendering resources (geometry, textures, shaders) which they hash in order to create object signatures. These signatures are then used to label the objects pixel-accurately. The signatures then enable them to propagate these labels across time and instances that share distinctive resources. The dataset of the paper contains 25,000 images from GTA5

## Chapter 2. Related Work

with pixel-level semantic segmentation ground-truth. Labeling the data took 49 hours which is 3 orders of magnitude faster than other semantic segmentation datasets with similar annotation density). This is achieved through the object signatures: When an object is labeled in a given image this label is propagated to every image that contains this object using the object signatures.



**2.2.1. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes**

[RSM<sup>+</sup>16]

### **2.2.2. SyB3R: A Realistic Synthetic Benchmark for 3D Reconstruction from Images**

[LHH16]

## 2.3. Problem Statement

**TODO:** what you have to do here :)



## 3. Conclusion

To conclude...



## **A. Blub**





# Bibliography

- [AFS<sup>+</sup>11] Sameer Agarwal, Yasutaka Furukawa, Noah Snavely, Ian Simon, Brian Curless, Steven M. Seitz, and Richard Szeliski. Building rome in a day. *Commun. ACM*, 54(10):105–112, October 2011.
- [BWSB12] D. J. Butler, J. Wulff, G. B. Stanley, and M. J. Black. A naturalistic open source movie for optical flow evaluation. In *European Conf. on Computer Vision (ECCV)*, Part IV, LNCS 7577, pages 611–625. Springer-Verlag, October 2012.
- [LHH16] Andreas Ley, Ronny Hänsch, and Olaf Hellwich. *SyB3R: A Realistic Synthetic Benchmark for 3D Reconstruction from Images*, pages 236–251. Springer International Publishing, 2016.
- [RSM<sup>+</sup>16] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio Lopez. The SYNTHIA Dataset: A large collection of synthetic images for semantic segmentation of urban scenes. 2016.
- [RVRK16] Stephan R. Richter, Vibhav Vineet, Stefan Roth, and Vladlen Koltun. Playing for data: Ground truth from computer games. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *European Conference on Computer Vision (ECCV)*, volume 9906 of LNCS, pages 102–118. Springer International Publishing, 2016.