First interim report

Yana Kryshchuk, Yaroslav Borys, Danylo Sahaidak

Research area

Right now we are surrounded by millions of smart devices: smartphones, smartwatches, etc. If we consider looking at our phones more accurately, we will see that they have at least 2 cameras (not talking about front and back one). Having this allows usage of stereo vision: blur background, help in face recognition. We apply it in different fields: from measuring parallax in astronomy to creating three-dimensional maps of objects. Also, stereo vision can be used in self-driving cars as the source of data about the outer world as well as parktronic systems which let people make the process of parking their car much easier.

Stereo vision is close to human sensations and has a bunch of advantages. For example, when a person loses sight in one eye, it is almost impossible for him/her to grasp depth perception at the same level as when they had sight in both eyes. But the stereoscopic vision gives the ability to determine the depths of the objects on the image, herewith we do not lose the information about colours.

In the modern world, stereo vision enables aviators, surgeons, and car drivers to perform their functions at a high degree of accuracy that would be quite challenging otherwise, no matter if we are talking about technology or a human eye.

Our aim

Our aim is to learn new methods and approaches in the field of linear algebra. What is more, we want to revise what we have already learnt on the course. Our most important goal is to successfully implement our primary idea, which is to distinguish different objects based on using different linear algebra methods.

Also, during the process, we want to learn how to build a stereo map of a picture from the dataset. It is essential for us to familiarize with the algorithms that are applied to this approach.

Besides we would like to try to use the result of the research and implementation of our project for personal purposes which are interesting to us and are beneficial for others. For example, attaching our system to radio cars(POC course).

Different approaches we've considered

We had a bunch of methods to consider which one would suit the best. Among them, we have chosen 3 ones. The first methodology is Scharstein's, which is about matching images by comparing their gradient fields.

The second one is the SSDMF algorithm. There are no exact explanations on how to implement it and all researches state it`s poor quality.

The third one is based on Conditional Random Fields. We have found some researches and are studying them.

We don't know if those methods are easy to implement, and it's possible that we will change some approaches.

Pipeline

1. Read all necessary literature and investigate different approaches to this task, download the datasets, compare them, start the research.
2. Started to consider a few alternatives simultaneously.
3. Perform implementation of several methods at one time in parallel, determining along the way which methodology is more accurate and better in general.
4. Determine the best option and focus on its developing. Other alternatives we will carry out as much as possible, considering our abilities, time and other resources.
5. Write python language implementation for our method.
6. If everything goes well - develop another method and make a comparison with the previous one.
7. Present what we have achieved.
8. If results are satisfactory on the dataset, in future we would like to test our implementation on real data, obtained by taking pictures of the object from different positions.

Testing

Right now we will test everything on free datasets, which can be found here: http://vision.middlebury.edu/stereo/data/. There are different types of data: autogenerated, real etc. Each dataset consists of 2 views taken under several different illuminations and exposures. Some datasets have more than 2 views. Here is a short description of a data group: “These 33 datasets were created by Nera Nesic, Porter Westling, Xi Wang, York Kitajima, Greg Krathwohl, and Daniel Scharstein at Middlebury College during 2011-2013, and refined with Heiko Hirschmüller at the DLR Germany during 2014.“ The datasets are also grouped by years.

The datasets include depth maps, so we have all the instruments to evaluate our results. For instance, to test a two-view stereo algorithm, the two reference views im2 (left) and im6 (right) should be used. Ground-truth disparities with quarter-pixel accuracy are provided for these two views.

Our plan: challenges and visions

The biggest challenge is to use our time wisely. As people are separated due to COVID-19 pandemy and everyone stays in his/her home it becomes even harder to achieve this. What is more, we are unable to influence each other in the way we used to and it is harder to meet and work together.

Talking about challenges which may appear from the technical side of work on our project, it is possible that we will not be able to find a perfect method to achieve our goal, for instance, it may be easy for realization but performance may be slow and vice versa.

In addition, there is always a chance that we will not be able to implement all of the selected methods, so as a result, we will have one working method, but we will not be able to compare it with others.