Jan Svensson elt12jsv@student.lu.se
Jonatan Atles elt13jat@student.lu.se

Expected start date: September 1st 2018
Expected end date: January 21st 2019

Object detection in AR

Background

We were contacted by Jayway to work on a master thesis proposal. The project involves developing an AR application that makes use of object detection and object recognition which utilizes machine learning strategies. Since we both are very interested in image analysis, graphics, machine learning and computer science, this project suits us very well because of our interests and competences.

Jayway gave us the opportunity to choose amongst the current AR technologies that exists today, to develop an application for one of these. The planned application is a digital manual for assembling furnitures. It works by aiming the camera at the pieces and then the application graphically demonstrates how to assemble them. This specific idea for an application might change during the project, but the general ideas will remain the same.

This project touches on a few different technologies that can be good to know about beforehand. These are Virtual Reality (VR), Augmented Reality (AR) and Object Recognition.

Virtual Reality

This is a technology that enables a user to enter into a whole new world, a virtual world. Here, the user can experience different things and feel as if he/she was really there. This is done, usually, by wearing a headset with goggles. These goggles render a 3D environment which the user observe by turning their heads and moving around.

Augmented Reality

Much like Virtual Reality, Augmented Reality combines the VR experience with the real world. This essentially means that virtual objects are rendered in the real world, as if they were really there. This is mostly done using smartphones, pads and specific AR devices, such as the HoloLens. Examples of applications that utilizes AR today are Pokemon Go and IKEA Place.

Object detection and recognition

Object detection is when a computer can detect an object from an image, such as detecting a chair in a room but without knowing that it is a chair. Doing this can be done by convolutional methods such as edge detection, corner detection, SIFT etc.

For recognition it is a bit more complicated. Here Machine Learning is necessary to accomplish the task. There are many types of methods to machine learning. Some of these are Support Vector Machine (SVM), Linear regression, K-nearest-neighbour or Neural networks.

The most usual approach is using for image recognition is a convolutional network which is a kind of neural network. The advantages with this method are the power of neural networks and limiting the input nodes. Therefore this is the method that we are going to use for our application.

AR and Object recognition together

The combination of these image related fields seem very reasonable and we believe that these technologies will become a standard in everyday life. In our application the AR will be used to detect planes in the real world and rendering instructions for the user. Object recognition will be used to locate and identify the different parts of the furniture. Together, they will form a complete solution.

Problem formulation

For this project there are a few obstacles that will have to be dealt with. Here we list a couple of them:

1. Today there exists a couple of different tools for AR frameworks. Thus, the first challenge will be to find a tool that suits our needs.
2. For training a convolutional network lots of data is needed. This data can sometimes be collected from an image library. These usually only contain ordinary objects such as cars, faces, dogs, etc. For special applications, however, the data will need to be collected manually and can create a lot of labour. Although, data can be artificially created sometimes, it is still a problem.
3. When our application detects an object, we want it to know about that object and its position even when the camera image moves a little. How do the application know that objects from two different images are the same object?
4. How many objects can be recognized at the same time, and how many times a second can this be done? What are the performance effects from this?
5. How do we deal with inaccuracy?

Delimitations

- The app will only work for one furniture. However, it can be trained for more products in the future.
- The app will only work for one furniture at a time.

Goals

Investigate different AR technologies for pros and cons and what they can be used for.

Combine Object detection, Object identification with Augmented Reality.

Come up with a possible prototype application with this technology.

Method

We will need a variety of tools which Jayway will provide. These tools consists of the following:

- Computers
- Work stations
- Platforms for AR