SCENE UNDERSTANDING
DISCUSSION REPORT FROM 02.05.2018

TOMÁŠ ONDRUCH (11740257)

1. Motivation for topic selection

At the very start of the discussion, the opponent was interested in further motivation and presenter's background which lead to choosing the problems of "floor plan reconstruction" and "room layout reconstruction" in the general topic of "scene understanding". The student explained her interest and involvement in a project which aims at developing software for interior design and room layout. In the recent years, both computer programs and also tools for augumented reality have made increasing use of computer vision in such tasks.

2. Image understanding vs. Scene understanding and interpretation

Subsequently, the importance of distinguishing between terms "scene interpretation", "scene understanding" and "image understanding" was properly discussed. It was agreed that there is no unique interpretation of a scene. Humans interpret the scene based on their experience from reality, having knowledge about the functionality of objects depicted in the image. Both their appearance and arrangement in the image can't be arbitrary and need to *match* their purpose in real world. So the keypoint of the matter is really to move the visual information from the picture towards an interpretation in reality, where the objects have specific functionality. Then we can evaluate the interpretation by measuring the extent to which visual information from an image corresponds with our pre-information from the real world.

- *Image Understanding* tries to explain every region in the image by segmentation and subsequent annotation. In relation to the presentation from 18.04. on Benchmarking for Image Understanding, the use of synthetic images supplied with information about segmentation and tags was mentioned as an effective tool in the process of dataset creation.
- *Scene Understanding* aims at relating various pieces of the image together. The goal goes beyond pure tagging of an image; the visual information is understood on a deeper level.

3. Specific features of planar objects' recognition problem

It was observed that multiple algorithms for indoor object recognition face problems when identifying objects such as hanging pictures, windows or doors, which are coplanar with main room-layout elements (walls, floor, ceiling). Interesting remarks were made in relation to this issue:

- Considering rooms of residential objects, an entrance in the form of a door is mostly a neccessary element of a room. Thus, if a method is not able mark the entrance in a created room layout, the results will suffer from significant drawbacks.
- Although the factor of volume of planar objects is usually insignificant in room layout planning, the importance of openings in a wall such as windows or doors is crucial in problems such as emergency route planning or crisis management.
- In relation to aforementioned problems, it is neccessary to develop methods, which are able to check whether a key element (or a set of such elements) is successfuly identified in a scene or if it is missing.

4. Floor plans for navigation problems

The last section of discussion focused on floor plan construction methods. The initial general question "To what extent are the static floor plans sufficient enough for successful navigation / path finding tasks in a building (i.e. for a mobile robot equiped with computer vision tools)?" was followed by naming multiple examples from real world, where the orientation is problematic. The possible solutions were stated, as well:

- Distinctive landmarks considerably help in navigation. However, complex office buildings often evince uniform design of interior spaces. For a human newcomer, such environment can become a true maze. The architectural solution of such issues usually involves marking the floors with specific colours or indicating departments by numbers and letters.
- Involving tags (of floor sections / rooms / spatial points) into floor plans can also lead to performance enhancement of robots' navigation in buildings. An example of a "hungry" mobile robot on a search for electric plug was given.
- In the real world, the environment can evince changes in mobile objects' arrangement. Thus, in case of map based indoor robot navigation a computer vision system still should be supplied. This enables easy obstacle detection (+ possible recognition) and subsequent avoidance.
- Supplying additional information to an automatized navigation system may considerably increase its performace. An example of numerous locked doors in the corridors of PRIP department of Faculty of Informatics was mentioned to demonstrate, that a floor plan itself doesn't always give a solution when searching for a path in a building.

A final remark was made by prof. Kropatsch, regarding suitable use of reference notation in presentations - simple enumeration of literature is not recommended. Notation involing first author's name and the year of publication is prefered.

Tomáš Ondruch (11740257)