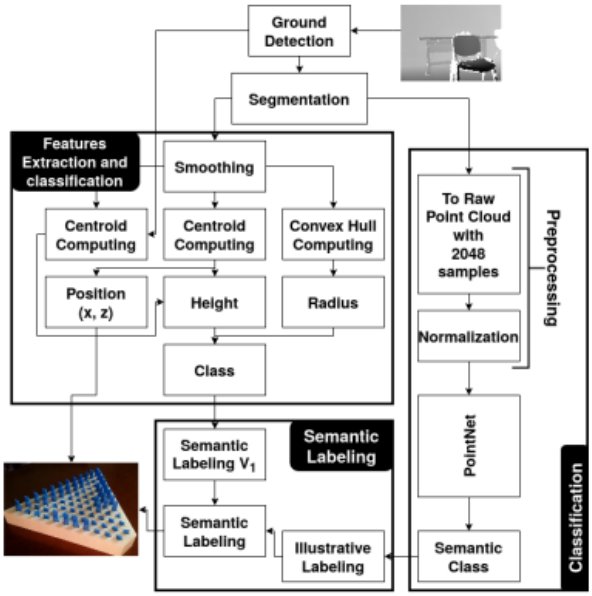
A tactile device for Scene understanding

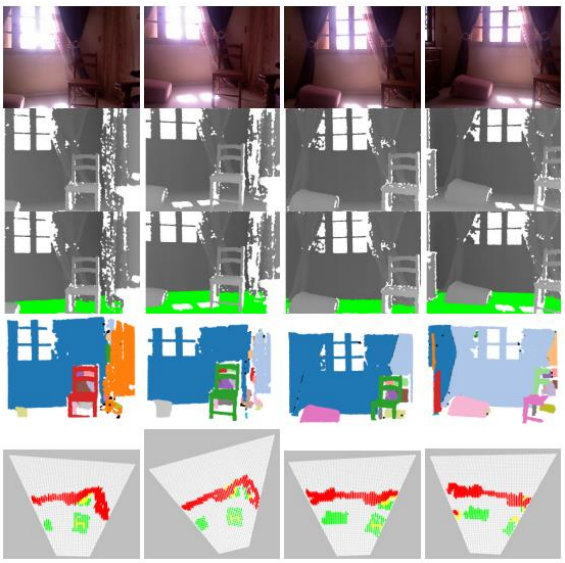
Chayma Zatout. PhD , Supervisor: Prof. Slimane LARABI, USTHB

Introduction

In daily life, people continuously interact with their surroundings in various ways, such as sitting on a chair, walking on a sidewalk, and navigating. To accomplish these tasks, humans execute a set of actions: first, they acquire information about the scene they are in, then they understand it and act accordingly. While obvious, it is important to state that understanding a scene is a fundamental step before executing most tasks, and it involves retrieving the scene geometry, detecting and locating objects, and identifying their shapes and relationships.



System overview



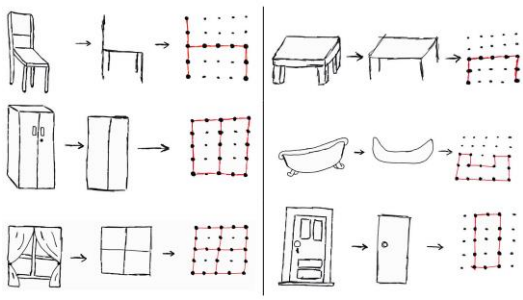
The execution of the system's different steps.

The method

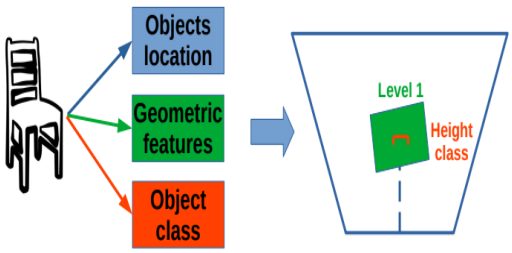
My thesis, entitled "Semantic Scene Labeling using RGB-D Data for Human-Scene Interaction", focuses on automating scene understanding, also known as scene semantic labeling. The main aim of my thesis is to build a computer-based system that can understand scenes for human-scene interaction using RGB-D data. The proposed system is based on point cloud processing computed from RGB-D data and includes four main steps. Firstly, the system detects the ground and extracts the occupied space that represents the objects in the captured scene.

A tactile device for Scene understanding

Secondly, the 3D representation of the occupied space is segmented using a clustering algorithm to extract the different objects constructing the scene. Thirdly, each object is classified using a deep neural network, and other object characteristics are extracted directly from the object's point cloud representation. Finally, the computed classes and features are used to generate semantic labels, providing a comprehensive understanding of the scene.



Semantic labels.



Mapping principle.

The results

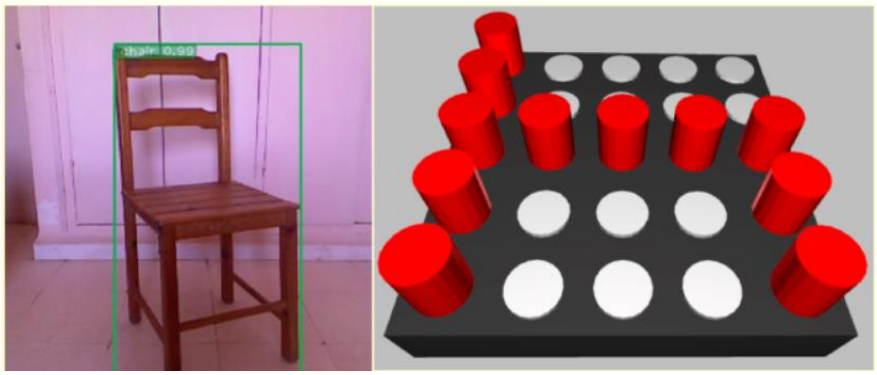
You may be wondering why we need to build a system to understand scenes when humans can do it effortlessly. The fact is that scene understanding is a critical component of many systems, particularly aids systems for people with visual impairments. In my thesis, the second and main focus was on developing a visually impaired aid system that uses semantic labels generated from scene understanding. As mentioned earlier, the system generates semantic labels that provide a comprehensive description of the scene. These labels are mapped onto a device called BASISR, which is designed to be understood by touch. The use of touch as a mode of understanding as well the deisgn of BASISR can be the subject of a whole article, but for now, we will only touch on the topic briefly. The semantic labels we propose are inspired by the Braille system and Kanji, the Japanese writing system. They are derived from the objects' shapes in the real world and represented in cells with 25 raised dots.

Once the geometric features and classes of the objects in the scene are computed, we generate their semantic labels on the BASISR device at the objects' locations. To represent the object's shape, the convex hull of the segment is mapped. To represent the object's class, the corresponding label is mapped onto the center. To represent the object's height, the label height corresponds to the object's actual height. By exploring the BASISR surface, visually impaired individuals can understand the objects' locations in the scene, their nature, and their height. This information can be used to accomplish various tasks such as navigation and searching for objects.

A tactile device for Scene understanding

Conclusion

This article has provided a brief overview of the focus of my thesis, which aims to develop a system for semantic scene labeling using RGB-D data for human-scene interaction, with a particular emphasis on developing a visually impaired aid system. The development of such a system holds great promise for improving the quality of life of individuals with visual impairments and underscores the importance of scene understanding in a variety of applications.



Left: The recognized object, Right: Its coding on synthetized 5x5 pins

The prototype of the Device designed and realized by the Master Students on Electronic of Embedded Systems at USTHB University: BOUCHTOUT Khadidja and SAADAT Nadia