論 文 の 内 容 の 要 旨

論文題目：

A Visual Tracking Based Object Learning Framework for Interactive Robot Systems

（対話型ロボットシステムのための視覚追跡に基づく物体学習フレームワーク）

氏　　名　チャウドゥリー クリシュニール　チャンド

For humans, the most influential data for decision making are acquired through visual perception. Humans ability to generate contextual awareness of the habitat for understanding the surrounding is key for adaptation and survival. Humans can easily generalize prior knowledge accumulated from different environments to unseen environments, generate accurate semantics, navigate and manipulate objects without assistance or prior learning. The ultimate goal in robotic research is to equip robots with learning capabilities that can emulate the natural learning process similar to humans. The natural learning capabilities will allow robots to incrementally enhance their internal knowledge of the surroundings directly from the environment with minimal human effort. Robots integrated with such learning capabilities will be able to facilitate self-learning, which will not only enable them to perform required tasks more reliably, flexibly, and efficiently but will also allow easy adaption to dynamic and unstructured environments.

A robot can enhance its knowledge of the surrounding by learning the objects found in the environment. Providing prerequisites of the objects for learning in offline, limits robots ability to known object instances. This limitation is crucial as there is no provision for robots to perform autonomous learning of unknown objects. Regardless of how extensively a robot is trained, it is likely to encounter novel objects when operating in a human environment hence, autonomous capabilities is an important requirement. Learning novel objects without prior knowledge is difficult due to unknown object attributes.

This thesis proposes an interactive visual learning framework that enables robots

to incrementally enrich their internal knowledge of the surrounding through online interactions. The framework allows robots to autonomously annotate and accumulate models of unknown objects from observations of human-object interactions. Specifically, the robot learns novel object representations by watching objects manipulated by humans. The intuition is that the motion of the object can be useful for segmenting and annotating object regions, which can be used as pseudo ground-truth for object learning. However, this is a challenging task as handheld objects moves with the motion of the hand which results in hand region segmented as part of the object due similar motion distributions.

The thesis first describes a robust tracking and segmentation framework that tracks and segments class-agnostic handheld objects. It is an end-to-end learning based system that learns to segment handheld objects by removing the hand and background outliers from the object space. Until now, the handheld object modeling algorithms are based on handcrafted heuristics which lacks the ability to generalize to various object categories. The object tracking and segmentation are formulated as a template matching problem which is solved using a learning based system that uses prior known object template to perform comparison in deep feature space. The segmented regions of the handheld objects are accumulated as object appearance models and used for learning the object. The experiments demonstrate the effectiveness of the method both on rigid and deformable objects.

Next, the thesis presents a segmentation approach for segmenting and discovering unknown objects from the environment. The method is a bottom-up segmentation approach that over-segments the scene into small clusters and groups the clusters to generate hypothesis of the objects. Unlike previous methods on bottom-up segmentation that uses handcrafted similarity function for aggregating cluster into objects, the proposed system learns a similarity function directly from the data. This allows the method to capture complex and non-linear relationships which are difficult to define manually. In robotic applications, the robots are expected to interact and manipulate objects of various geometrical structures. Manipulation of unknown objects is a difficult problem. To allows robots to manipulate unknown objects, a method for detecting object affordance is described. It describes part affordances of the object that explains to the robot the functions of the object.