Object Recognition and Pose Estimation Algorithm using Point Pair Feature and Curve Set Feature

著者	李 明宇
学位授与機関	Tohoku University
URL	http://hdl.handle.net/10097/00127339

Object Recognition and Pose Estimation Algorithm using Point Pair Feature and Curve Set Feature (ポイントペア特徴量とカーブセット特徴量を用いた物体認識と姿勢推定アルゴリズム)

Mingyu Li

Abstract

Chapter 1　Introduction

Vision-based rigid object recognition and pose estimation has been researched for years because of the importance in robotic applications. The image from the vision sensor, which could be two dimensional or three dimensional, is usually called a scene. Given the CAD model of the target object, the algorithm is designed to recognize the target instances from the scene and estimate the 6 Degree-of-Freedom poses. This is generally performed by computing some features in the scene and matching with the model features. However, there are still problems that make it difficult to achieve a high performance.

The first one is the model. It is desirable that the algorithm is capable of handling arbitrary shape objects. However, some objects are not characteristic enough, for example, the industrial objects used in robotic bin picking. Thus the algorithm should be designed properly to recognize them effectively.　The second problem comes from the cluttered background. In many cases, there are other objects in addition to the target in the captured images and the algorithm needs to localize the target from them. On the one hand, this increases the computation time since it is necessary to search in a larger area. On the other hand, if some object is similar to the target, it might be mistaken for the target which causes recognition failure. The third challenge is to perform multi-model tasks. Instead of only recognizing and estimating one target, sometimes the algorithm is expected to recognize multiple targets and the performance might decrease with the increase of target number.

Therefore, in this thesis, I focus on three tasks: single-model pose estimation for bin picking, multi-model pose estimation for bin picking and object recognition and pose estimation for daily objects in cluttered scenes. It is assumed that the CAD model of the target is available and the deformation is very small. Though many state-of-the-art algorithms incorporate depth information with color information to achieve higher performance, I use only depth information to ensure our algorithm could be applied widely.

In the first task, point pair feature is utilized to propose pose candidates. The features between every two model points are precomputed and stored in a hash table. An efficient voting scheme is performed to match scene features to model features. Then a voxel-based robust pose verification method is proposed to select best poses from tens of thousands of candidates. A pipeline of selecting multiple instances fast is introduced.The method could detect multiple instances of the target object in bin picking with high recognition rate and accuracy.

To achieve multi-model tasks, point pair feature method is not proper since the detection time increases linearly with the model number. I propose a novel descriptor Curve Set Feature(CSF) to describe a point by the surface fluctuation around this point. Rotation Match Feature (RMF) is proposed to match CSF efficiently. The matching process combines the idea of matching in 2D space of origin PPF algorithm with nearest neighbor search. The curve set feature method is evaluated against a large number of bin picking scenes and it is proved that the performance on multi-model tasks is as good as that on single-model tasks.

To solve the problem of background clutter, a boundary-based preprocessing method is conducted to remove most background points. Then the search is performed on reserved points and the verification method is modified to implement on a large scene space. Boundary verification is introduced to improve the verification accuracy and the method is proved to be more accurate than state-of-the-art algorithms that use both depth information and color information.

Chapter 2 Single-Model Pose Estimation for Bin Picking

Bin picking system typically consists of three components: a sensor above the objects, a processor and the robot arm. The sensor captures the 2D or 3D images of the objects and sends to the processor. The processor computes the position and orientation (pose) of the objects and decides the picking point and picking path. Finally the robot arm picks up the object and places it to the specified position. In order to achieve a higher productivity, an accurate and efficient pose estimation algorithm is necessary to recognize the target instances and computes the 6 DoF poses.

For bin picking tasks, the objects are piled in the container and the position of the container is usually fixed. Therefore, I only need to focus on the fixed area. The difficult of bin picking tasks is that the targets are usually industrial objects that are usually consisted of primitives such as plane, sphere, cylinder. These primitives are relatively less discriminative to recognize compared to arbitrary shapes because they are lack of curvature changes.

In this chapter, I introduce our point pair feature-based pose estimation algorithm for bin picking. Normals are computed in a more precise way to improve the accuracy of the algorithm. Then the point pair features of scene points are computed and matched with model points to compute pose candidates. Different from the origin PPF algorithm, only a small part of the point pairs are computed to improve the efficiency. Pose verification method is performed to verify all the poses. A multiple selection method is used to select result poses without repetition. Finally Iterative Closest Point (ICP) algorithm is performed to improve precision.

Chapter 3 Multi-Model Pose Estimation for Bin Picking

In last chapter, I introduced our point pair feature-based pose estimation algorithm for bin picking tasks of single model. The algorithm presented good performance on a large number of bin picking scenes. But sometimes, different objects are piled together in the container, which I call multi-model tasks. To achieve this, a simple method is to save the features of all the models in the hash table. But the computation time would increase linearly with the model number because the point pair feature method stores all the features in the hash table. More features in the hash table causes more matching point pairs in the voting scheme, thus increases matching time.

In this chapter, I introduce the curve set feature-based pose estimation algorithm for bin picking tasks of multi-models. Different from some template matching algorithm whose matching time increases with model template number, the time of our algorithm seldom increases with model number and at the same time, the recognition rate for multi-model tasks does not decrease compared with single model tasks. Neither color information nor segmentation is necessary. An efficient boundary verification method is introduced to boost the pose verification method. Based on the pose estimation method, a picking point selection method is introduced to decide which object to pick up and how to grab it with our two-finger gripper without collision. Experiments on synthetic and real scenes show that our algorithm can estimate the poses of different objects with high recognition rate and efficiency. A success rate of 98.2% on robotic picking experiment was performed.

Chapter 4 Object Recognition and Pose Estimation for Daily Objects in Cluttered Scenes

In this chapter, I address the problem of recognizing and estimating the poses of objects from cluttered background. Different from the bin picking tasks, the target object could appear in any position in the scene, thus the searching area is much wider which causes increase in computation time. Besides, other objects sharing similar shape as the target makes it difficult to recognize with depth information accurately.

In this chapter, a point pair feature based pose estimation algorithm using depth information is proposed. To improve the efficiency of point pair feature approach, a boundary-based preprocessing method is proposed to remove background points and points belonging to foreground objects that are larger than the target. Then, the point pair feature approach is performed on remaining points to obtain possible poses. For objects that are difficult to recognize from some viewpoints, an additional hash table is built and a model template selection method is proposed. A fast and accurate pose verification method considering both point correspondence and boundary correspondence is introduced to grade the poses and select the best pose. Our algorithm is proved to be able to compete with state-of-the-art algorithms using RGB-D information on published datasets.

Chapter 5 Conclusions and Future Work

In this thesis, object recognition and pose estimation algorithm for rigid object using depth information is proposed. Corresponding algorithms are designed for different tasks.

In Chapter 2, point pair feature-based single-model pose estimation algorithm for bin picking is introduced. Considering the matching of industrial objects takes more time than daily objects, less features are utilized during the online stage. A voxel-based pose verification method is proposed. It is robust, fast and easy to code and transplant to other methods. By evaluating and grading a large number of poses efficiently, the accuracy and speed of the algorithm have shown a huge increase compared to the origin point pair feature method.

In Chapter 3, curve set feature method is proposed for handling multi-model tasks for bin picking. By matching the features with FLANN, the matching time and accuracy of multi-model tasks remain almost the same as single-model tasks. Besides, a boundary verification method is proposed to further improve the result of verification. A pose estimation-based picking point selection method is proposed on our two-finger gripper. The method is proved to be accurate and fast against a large number of synthetic and real scenes and a robotic picking experiment.

Chapter 4 presents our point pair feature based method for cluttered scene. Based on the relative position with neighboring boundary points on different directions, the background points are recognized and removed to save computation time. By implementing the verification method in an inverse way, the problem of memory and time increase in large space is solved. Our algorithm shows superior performance than state-of-the-art methods on two public datasets.