Interactive Perception based on Gaussian Process Classification Applied to Household Object Recognition & Sorting

Aamir Khan[1], Gerardo Aragon-Camarasa[1], J. Paul Siebert[1]

I. INTRODUCTION

We present an ongoing investigation into an interactive perception model for sorting objects observed in a workcell scene into bins based on Gaussian Process (GP) classification, which is capable of recognising objects categories from point cloud data. In our approach, FPFH features are extracted from point clouds to describe the local 3D shape of objects and a Bag-of-Words coding method is used to obtain an object-level vocabulary representation. Multi-class Gaussian Process classification is employed to provide a probability estimate of the identity of the object and serves a key role in the interactive perception cycle – modelling perception confidence. We present results from semi-autonomous object sorting experiments in which the proposed GP based interactive sorting approach outperforms random sorting by up to 30% when applied to scenes comprising configurations of household objects. Furthermore, we briefly discuss the autonomous interactive perception approach based on GP for sorting household objects.

It is essential for service robots to have the ability to recognise objects in their immediate vicinity when working in dynamically evolving human environments. Ideally, these robots should be capable of detecting and classifying objects within their environment and then interacting with these objects without need for supervision. In [1], it is argued that physical interaction further augments perceptual processing beyond that which can be achieved by invoking deliberate pose changes since interactive perception allows the perception module to acquire potentially more information about the object(s) in the surrounding environment. It also has the the potential to reduce the complexity of the observed scene, in this case encoded within a heuristic that directs the robot to grasp an object under investigation, separate it from other surrounding objects prior to investigating it further.

In this paper, we present an interactive perception system that is able to sort everyday household objects into bins, each bin allocated to a specific object category, through direct visual observation (typically when objects are not occluded), and then by means of active object manipulation should objects be occluded or "difficult to recognise". Our proposed framework does not require prior knowledge about the environment or scene. We also present a visually assisted object sorting system which is capable of segmenting a set of household objects lying directly on the robot's workspace table and categorising those objects into their respective object classes (ie.g. juice bottles, mugs, etc.). The system has been pre-trained on a subset of these object instances, while a proportion of the objects we investigated have not been used to pre-train the system. A high-level model of our classification model is presented in Figure 1.

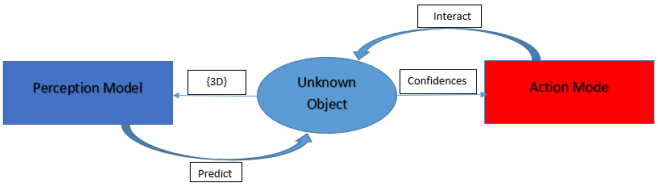

Fig. 1. Our Gaussian Process Classification based Interactive Perception Model (GP-IPM).

Our system is portable, invariant to 6 DOF pose changes and operates close to real-time. The pipeline comprises the following: object segmentation, visual representation, classification, semantic visualisation and finally in our current experiments, a human operator responsible for removing the object from the scene following its correct classification by the system. Our Gaussian Process classification based Interactive Perception Model (GP-IPM) has been cross-validated by comparing the categorisations results obtained when using an SVM multi-class classifier to ground truth.

The operating scenario we have adopted comprises a visual search task that attempts to locate and identify an object, such as a bottle or pen, which can potentially be partially occluded. If an initial search fails to locate an object with a prescribed minimum level of confidence, this object is then re-positioned. Based on the above scenario, we present results obtained from experiments where the object are sorted by the probabilities of the class predictions obtained from the GP based multi-class classifier. Our GP classification based interactive perception model makes the following contributions to the state-of-the-art in object sorting:

1) We present the first example of research to adapt non-parametric multi-class probabilistic classification (via Gaussian Processes) to the household object recognition problem.
2) We demonstrate that the proposed GP-IPM approach applied to a semi autonomous sorting task yields substantially improved performance over non-interactive

[1]Aamir Khan is with the School of Computing Science, University of Glasgow, G12 8QQ, UK a.khan.4.@research.gla.ac.uk
[1]Gerardo Aragon-Camarasa is with the School of Computing Science, University of Glasgow, G12 8QQ, UK gerardo.aragoncamarasa@glasgow.ac.uk
[1]J. Paul Siebert is with the School of Computing Science, University of Glasgow, G12 8QQ, UK paul.siebert@glasgow.ac.uk

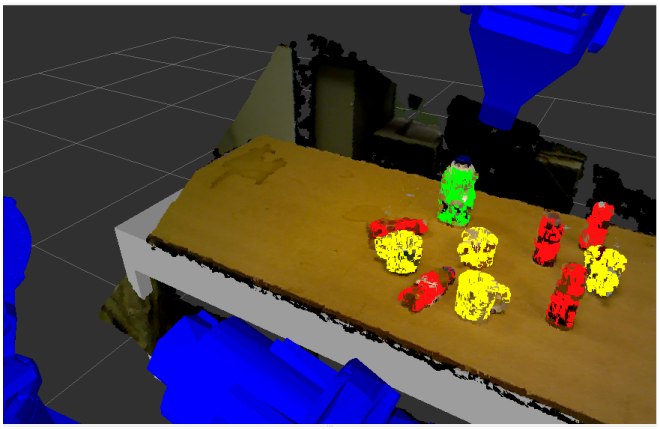

Fig. 2. Robot recognising object categories according to colours.

TABLE I

TABLE SHOWING TRUE POSITIVE FOR RANDOM AND GP-IPM

Scene Complexity	Method of Sorting	True Positives					Total
		Juice Bottles	Milk Cartons	Bowls	Mugs	Juice Boxes	
Uncluttered	Random	5/5	1/1	-	4/4	-	9/9
	GP-IPM	5/5	1/1	-	4/4	-	9/9
Cluttered	Random	2/4	1/2	2/3	2/3	1/2	8/14
							57%
	GP-IPM	4/4	2/2	2/3	3/3	1/2	12/14
							86%

alternatives.

II. RESULTS

We report an experiment are designed to demonstrate the validity of our Gaussian Process based perception model, i.e. confirm that it is capable of improving recognition rates when aided by object interaction. We carried out experiments for five household object categories i.e Juice Bottles, Milk Packs, Bowls, Mugs and Juice Boxes. We arranged scenes based on combinations of known and unknown object instances, of 5 different object classes. Objects were placed in arbitrary poses and locations as shown in figure 2. In figure 2, a red overlay depicts juice bottles, green for milk boxes, blue for bowls and yellow for mugs. Our training dataset consists of the above 5 object classes – we created this dataset by capturing point clouds of each object at angular intervals of approximately 20 degrees.

In this section, our GP based perception model is evaluated for sorting objects into their respective bins. Note that the objects are placed directly on the table. An input point cloud is obtained from the ASUS Xtion pro camera from a single view and this point cloud is then passed on to our pipeline.

To demonstrate the viability of our GP based interactive perception model for improving recognition rates for object sorting, we constructed two scenes; (1) comprising simple unoccluded objects, and (2) mutually occluding simple and challenging objects (such as bowls).

After all the objects have successfully been recognised and classified accordingly, objects were manually placed into their respective bins, Table I shows the success rate for the objects placed into the respective bin guided by our proposed GP-IPM. For complex scenes, containing objects having challenging shapes and occlusions, our proposed method outperforms random sorting by a factor of 30%.

III. CONCLUSIONS AND FUTURE WORK

We have described an interactive perception based object sorting system, incorporating depth sensing, 3D feature extraction, object representation and classification. Our visual perception system achieves near real-time object category identification with associated perception confidence estimates, which serves to establish whether an observation is sufficient to make a classification or to determine if it is necessary to interact with the object to achieve a sufficiently reliable classification of the instance contained in the observation.

We have also established a dataset of household objects used in the evaluation in this paper, and the experimental results show that our proposed perception pipeline achieves an 80% recognition rate for 5 object categories. As the autonomous robot manipulation phase of our work is currently ongoing, this paper focuses on the visual perception components of our system and we present sorting results obtained using a human operator to manipulate the objects. From these semi-autonomous sorting experiments, the proposed Gaussian Process based interactive sorting system outperforms random sorting by upto 30% in terms of sorting accuracy. We also observe that our interactive perception strategy not only mitigates segmentation failures prevalent in single-shot sorting, but also increases perception performance in terms of recognition rate, thereby facilitating the sorting decision.

We are currently extending our robot testbed facilities to incorporate the visual perception approach described above with autonomous manipulation skills (successfully demonstrtated in a number several visually-guided manipulations tasks [3], [2]). Our longer-term goal is to, achieve fully autonomous category-based visually-guided household object sorting. To facilitiate this aim, we propose to expand our dataset to include more object categories comprising a greater variety of object instances. We also propose to extend the system to segment objects in order to learn their appearance by interaction and to capture their visual representation from different views in a manner that mimics human observation of unfamilaiar objects.

REFERENCES

[1] J. J. Gibson, The senses considered as perceptual systems. Boston: Houghton Mifflin, 1966.

[2] Sun, Li and Rogers, Simon and Aragon-Camarasa, Gerardo and Siebert, J Paul, Recognising the clothing categories from free-configuration using Gaussian-Process-based interactive perception, 2016 IEEE International Conference on Robotics and Automation (ICRA).

[3] Li Sun, and Gerardo, Aragon-Camarasa, and Simon, Rogers and J. Paul Siebert, Accurate Garment Surface Analysis using an Active Stereo Robot Head with Application to Dual-Arm Flattening, 2015 IEEE International Conference on Robotics and Automation (ICRA). note = Accessed: 2016-07-15