Grasp Exploration for 3D Object Shape Representation Using Probabilistic Map

Diego R. Faria, Ricardo Martins, Jorge Dias

Grasp Exploration for 3D Object Shape Representation Using Probabilistic Map

Diego R. Faria, Ricardo Martins and Jorge Dias

Institute of Systems and Robotics,
Department of Electrical Engineering and Computers, University of Coimbra – Polo II,
3030-290 Coimbra, Portugal
{diego, rmartins, jorge}@isr.uc.pt

Abstract. In this work it is shown the representation of 3D object shape acquired from grasp exploration. Electromagnetic motion tracking sensors are used on the fingers for object contour following to acquire the 3D points to represent its shape using a probabilistic volumetric map. It is used the object referential for its representation. For that, it is found the center of mass of the 3D object through the moments to define its referential. The occupancy of each individual voxel in the map is assumed to be independent from the other voxels occupancy. The posteriori achieved from Bayes' rule is the probability distribution on the occupations percentage for each voxel. The probabilistic map in a Cartesian system is converted to the spherical coordinate system for visualization with more details on its surface.

Keywords: Grasp exploration; Probabilistic map; Object shape representation.

1 Introduction

Applications of service robots will require advanced capabilities of grasping objects and skills that allow a robot to recognize the object also through the grasp exploration. Human uses the hand for recognizing some objects properties such as size, texture, etc. The skills of grasp exploration to acquire objects properties are important also in the robotic field to complement others sensors such as vision and laser to obtain more information of an object. The in-hand exploration strategies performed by humans motivated the development of analogous approaches to promote the exploration of surfaces of objects by robotic hands. These exploratory procedures are used to created internal representations of objects in order to proceed to its identification and integrating its 3D spatial and physical characteristics distribution to control the interaction with it.

In this work, it is generated the representation of the shape of a 3D object extracted from the data acquired during the human hand exploration of the object (by contour following). The representation of the object shape is built using a probabilistic volumetric map. The referential associated to the representation of the shape of the object is located in the estimated center of the mass of the object.

2 Contribution to Technological Innovation

Nowadays, there is a tendency to move the development of robotic manipulators from simple grippers towards more natural human inspired articulated robotic hands, with integrated multimodal sensing technologies. This new generation of autonomous dexterous robotic hands is playing an important contribution to the evolution of the robotic research paradigms. The field of robotics is moving to the development of new systems adapted to be introduced in new environments far from the traditional robotic applications in the industrial production lines. There is an emerging trend in the development of methods and strategies to endow these robotic platforms with the ability to autonomously explore these environments, proceed to the characterization of objects, use them as tools or interact with them. This work intends to contribute to the development of autonomous dexterous robotic hands by developing methods to extract a representation of the manipulated object in order to estimate the object shape and size. These methods are developed using contact points acquired from demonstrations performed by humans during the in-hand exploration. The developed methodology will be then transferred to robotic dexterous hands in order to estimate the shape and size of a manipulated object, from contact points acquired by a robotic hand through contour following. Achieving the 3D object shape estimation, it is possible to determine, through its geometrical properties, the best region of the object to perform the grasp. Using this knowledge (object representation) it is possible to endow a robot to grasp different types of objects including unknown objects.

3 Related Work

Researches about human perception concerning haptic exploration disclose that contour following is an ordinary way of "exploratory procedures" that people use for determining the geometry of an object [1]. Many approaches have been proposed regarding robot haptic exploration for object recognition and object shape representation. In [2] the authors present haptic object recognition using a dexterous robot hand with a manipulator arm. Through the hand contact by enclosing the objects at predefined positions was possible evaluate joint angles and force readings. The object shape recovery was performed using sparse contact points from hand. The point clouds were fitted to superquadric models defined by a set of parameters describing shape and pose. In [3] the authors developed a framework for haptic exploration by contour following that can be used with robot or human hand. The object shape contour is acquired by a human hand using a data glove where the human operator visually guides the contact sensor along the contours of the object by the index fingertips. It is used an extended superquadric function for primitive modeling which can represent a variety of cubical and spherical geometries. They address modeling of basic superquadric shapes. The authors have used stereo vision for wrist tracking by particle filter. They have used a marker bracelet attached to the wrist of the subject with the data glove for its tracking. The proposed work in [4], the authors have used superquadric functions for shape recovery from haptic exploration with multi-fingered robot hands using fingertip tactile sensors. They have applied a hybrid minimization method utilizing a genetic algorithm by considering the contact normal information to recover superquadric primitives from synthetic exploration data. In a previous work [5] we developed a method to extract basic shapes (primitives) from data acquired from grasp exploration. Gaussian Mixture Models was used for points clustering and outlier removal. After a learning phase, the object shape was classified through Bayesian techniques. It was used least square minimization of a distance metric to find the shape orientation and scale for its representation.

4 Research Contribution and Innovation

In this section is described the methodology adopted for object representation using probabilistic map. Mapping techniques as occupancy grid [6], [7] has been used tin robotics field to describe an environment of a mobile robot. Two-Dimensional grid has been very used for static indoor mapping [8]. The idea is to verify the probability of each cell to be full or empty after the sensors observation. In [8] is described the standard grid mapping algorithm that is a version of Bayes filter. This filter is used to calculate the posterior over occupancy of each cell. Probabilistic volumetric maps are also useful as presented in [9] for data fusion (visual and auditory perception). The main idea of using the probabilistic map is due to: reach a simpler way of static object reconstruction and representation and the uncertainty of sensor noise due to real world (the sensor probability model depends on the characteristics of the sensor and the object being sensed). In the next subsections are presented the detailed content about the employed methodology for object representation through grasp exploration by contour following.

4.1.1 Probabilistic Volumetric Map for Object Representation

This work presents a model for grasp exploration, but it also can be used for data fusion. Our intention is to acquire the 3D shape of the object so that instead of work in an egocentric way, we will work with object-centric representation in a spherical coordinate system to reach more resolution on the object surface. The volumetric map is updated along the exploration in discrete intervals. The proposed methodology will be used for data fusion (multimodal perception) in a future work, e.g. using vision and touch. The occupancy of each individual voxel is assumed to be independent from the other voxels' occupancy and thus O_C is a set of independent random variables: $C \in M$ - index a cell on the Map; O_C - probability value describing the occupancy of the cell C; Z_{grasp} - grasp exploration measurement that influences the cell ***C*.** It represents 5 sensors which each one returns the 3D position of a movement; $P(O_C)$ - Probability distribution of preliminary knowledge of coverage value describing the occupancy of the cell C**,** initially is an uniform distribution; $P(Z_{grasp} \mid O_C)$ - probability distribution corresponding to the set of measurement Z_{grasp} that influences the cell C taken from the grasp exploration data. This distribution is taken from occupancy model. In case of data fusion from different sensors, it is necessary to declare the variables that represent the distribution of the measurement, e.g. working also with vision, then $P(Z_{vision} \mid O_C)$ represents the sensor model. In case of different sensors the joint distribution decomposition of the relevant variables shows the dependency assumptions according to Bayes' rule as follows:

$$P(Oc\, Z_{vision}\, Z_{grasp}) = P(Oc \mid C)P(Z_{vision} \mid Oc)P(Z_{grasp} \mid Oc) \quad \mathbf{(1)}$$

The *posteriori* is the probability distribution on the occupation's percentage $P(O_C \mid Z_{vision}\, Z_{grasp})$ for each voxel:

$$P(Oc \mid Z_{vision}\, Z_{grasp}) = \alpha P(Oc \mid Z_{vision})P(Oc \mid Z_{grasp}) \quad \mathbf{(2)}$$

Since this work is just for grasp sensor model, it is possible to simplify the estimation model of occupancy, so that the *posteriori* in this case is the probability distribution on the occupation's percentage $P(O_C \mid Z_{grasp})$ for each voxel. The occupation's probability is given by:

$$P(O_C \mid Z_{grasp}) = \frac{P(Z_{grasp} \mid O_C)P(O_C)}{P(Z_{grasp} \mid O_C)P(O_C) + (1 - P(Z_{grasp} \mid O_C))(1 - P(O_C))} \quad \mathbf{(3)}$$

where $P(O_C \mid Z_{grasp})$ is the posteriori value; $P(Z_{grasp} \mid O_C)$ is acquired by the sensor measurement (likelihood); $P(O_C)$ is the *priori* information (at the beginning it is a uniform distribution representing the state full or empty) and subsequently the last *posteriori* becomes the *prior* for the next computation.

4.1.2 Grasp Exploration Occupancy Model

For the data acquisition (grasp exploration), we are using the Polhemus Liberty system [10]. It is used one magnetic sensor on the fingers (thumb, index and middle) to acquire the shape of an object by contour following. Each sensor return the 3D coordinates based on Polhemus Liberty referential. The frame rate of each sensor was configured to be up to 15Hz. During the data acquisition, it is defined a workspace ($35x35x35cm^3$) in the experimental area to place the object for mapping. This space is subdivided in $0.5x0.5x0.5cm^3$ cells. During the displacement of each sensor in the workspace area, each one gives the information about its 3D position and it is possible to identify in which cell that the measurement is inserted. Due to the size of each cell relatively to the standard deviation of the magnetic tracking sensors measurements (until 0.2 cm for linear movements of 10cm), it is defined inside each cell a 3D isotropic Gaussian probability distribution, $\boldsymbol{P(Z_{grasp} \mid Oc)}$, centred at the cell central point with standard deviation 0.2 and mean value equal to the cell central point coordinates of the cell. The probability of a measurement belongs to that cell is given by the equation (4). The values are normalized in order to consider that the probability value assigned to a point located at the centre of the cell is equal to 1.

$$P(Z_{grasp} \mid O_C) = \frac{1}{(2\Pi)^{3/2} \mid \Sigma \mid^{1/2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \, \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right) \quad \mathbf{(4)}$$

In (4) $|\Sigma|$ represents the determinant of Σ (sensor noise variation). It can also represent a scalar value. Due to the normalization used, the equation (4) takes the form:

$$P(Z_{grasp} | O_C) = \exp\left(-\left(\frac{(x-u_x)^2 + (y-u_y)^2 + (z-u_z)^2}{2\sigma^2}\right)\right) \quad \textbf{(5)}$$

4.1.3 Object Referential & Representation in Spherical Coordinate System

To have a better representation of the object and to reach more details on its surface we are working with object-centric representation. To find the object referential and represent it centred in a sphere, it is computed the center of mass of the object. As long as the center of mass of the object is determined, it is possible to define the object referential. The center of mass of the points' distributions that represents the object is found along the grasp exploration in the workspace. While the data is acquired during the exploration, the center of mass the object is changing due to the amount of data and the object surface growing (Fig. 1(a)). The axes of the tracker device referential are kept, so that we just need to translate the points in the tracker referential to the object referential. Through the centroid, it is possible to define the radius of a sphere is the center of mass of the object, thus it is possible to center the object inside a sphere for its volumetric representation as shown in Fig.1 (b). To compute the center of mass of a 3D object shape we compute the discrete moments. The moments are a measure of the spatial distribution of the mass of a shape. The centroid is reached through some steps, computing the zero moment (summation of the voxels); first moments for *x*, *y* and *z* and then it is computed the centroid for each axis. The moments are computed in an iterative way, at some discrete intervals, after a new exploration the center of mass of the object is updated.

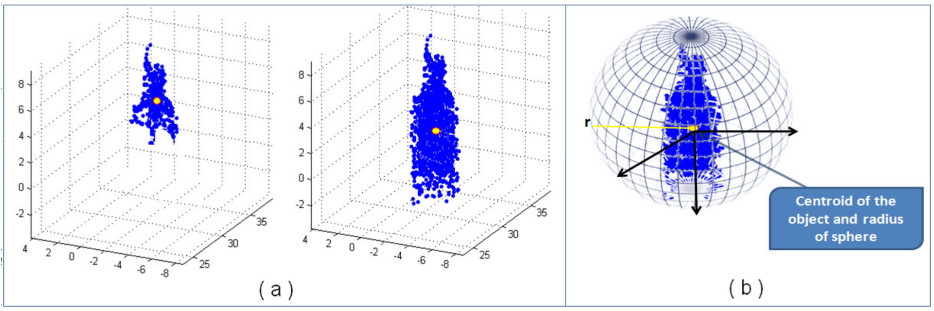
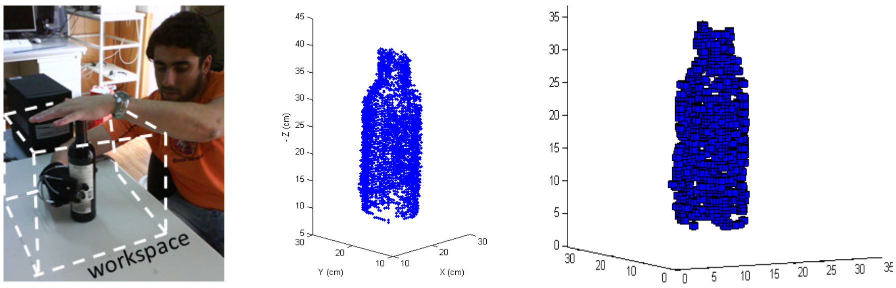


Fig. 1. (a) The center of mass of the object changes along the grasp exploration; (b) Illustration of the object representation in spherical coordinate system.

4.1.4 Experimental Results

The magnetic sensors were attached to a glove to perform the object contour following to acquire the contact points. For this experiment, it was decided to use just the thumb, index and middle fingers to explore the object surface due to be enough to cover the object shape through the movements around it. The setup for the experiments is composed of a wooden table, without any metallic parts, since the electromagnetic tracker is sensitive to nearby ferromagnetic materials. The rigid 3D object is fixed on the tabletop in a defined workspace for all experiments having the object in the same area. The magnetic tracker emitter unit that determines the frame of reference for the motion tracking system is placed on the tabletop at ~30cm near to the object. A workspace of 35x35x35cm^3 was defined on the table for the object mapping. The referential of the workspace was defined to be parallel to the sensor referential and then a simple translation is enough to convert the 3D points on the sensors referential to the workspace referential. Fig. 2 (a) shows the experimental setup area. Each voxel of the volumetric map was defined to represent an area 0.5x 0.5x0.5cm^3 due to the good precision of the sensor position resolution at 30cm range (less then 1mm). The chosen exploratory procedure is performed during 90 seconds. The volumetric map is updated with the new sensors measurements every 15 seconds. The contour following movements have no pattern, that is, it is not necessary to follow some pre-defined rules such as making movement just in one direction. The subject is allowed to do movements in any direction, but for better representation it is suggested to perform horizontal and vertical movements around the object until to contour all surface. The chosen object for this experiment was a bottle of wine. Fig. 2 (b) and (c) shows the results achieved in the end of the exploration.



(a) (b) (c)

Fig. 2.(a) Experimental area for the in-hand exploration; (b) raw data; (c) representation of the object shape (just voxels with probability higher than 0.7).

Fig.3 shows other views of the object centered in the estimated center of mass.

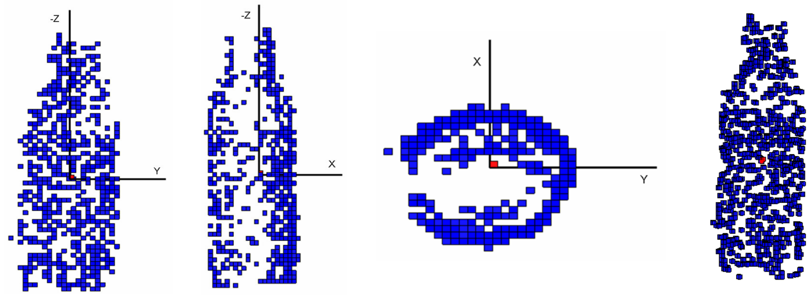


Fig. 3. Representation of the occupied cells with probability higher then 0.7 (object centered in the estimated center of mass).

5 Discussion of the Results

This approach for shape representation has a limitation: the rigid object need to be fixed in the specified workspace. If the object abruptly moves many times during the exploration, the reconstruction of the object can be less precise. The idea of this methodology is to do not represent the points with low probability so that the reach-to-grasp movement before the object contour following starts are not represented because the hand trajectory reaching the object have very few points and that points are not repeated (i.e. do not pass in the same cell twice) to increase the probability of the cells that the movement has occupied. The application ignores all points outside of the workspace where the object is placed. The adopted method uses a unique global map, where all the fingertips data are represented e.g. if one or more sensors passed through the same cell the probability increases. Other method could be applied; instead of building a unique map for all sensors, it could be built a local map and then fusing them into a global map through a Bayesian filter. The exploration was done with the sensors on right hand. During the trials, the subject that was performing the exploration usually hold with the left hand the top of the object to guarantee that it does not move, thus it is possible to reach better results, but at least a small variation of the object position can happen. After analyzing the exploration performance of some subjects, it was possible to see that the object shape influences the exploratory procedure, e.g. the bottle of wine was more explored in the middle due to the cylindrical shape. The top and bottom of the explored object is poorer in the representation due to the subject explore less those parts. The retrieved object attributes such as shape and size were satisfactory. The raw data, i.e., fingers movements inside the workspace were 6169 points. We have considered all cells with probability higher than 0.7 to represents the object shape. After computing the probabilistic volumetric map, the object shape has occupied 723 cells (probability >= 0.7). The use of tactile sensors for contact points could facilitate and improve the results, e.g. using tactile sensor would be possible to know when the fingers are in contact to the object to initialize the tracker device points acquisition, avoiding to start the application with the hand on the object or to fill some voxel on the workspace during the reach-to-grasp. In this work we have introduced the idea of work with multimodal perception. Our future directions is also to use the visual perception to acquire more information of the object and then through sensors fusion we can reach better results for the probabilistic representation of the object to be used for features extraction for its characterization.

6 Conclusion and Future Work

In this work is shown the probabilistic representation of a 3D object model achieved by grasp exploration through probabilistic volumetric map due to the simplicity to recover, by this method, a shape given the sensors measurements. The center of mass of the object is computed to define the object referential for its representation in the spherical coordinate system. As future work, it will be used the extended model for sensor fusion using stereo vision and grasp exploration for better representation of the object. For that, it is necessary sensors calibration, and it is necessary a joint distribution decomposition of the relevant variables to show the dependency assumptions according to Bayes' rule for the data fusion. The experiments will be performed with non-fixed object (in-hand exploration using two hands). For that, it is necessary to have an initial position of the object to be used as reference, and then it will be computed a transformation between the object referential when it is moving to the initial position referential. This will be done every time that the object moves to be possible to map all points in the origin to have the object model representation. After acquiring the representation of the object will be possible to recognize and characterize it.

Acknowledgments. This work is partially supported by the European project HANDLE (23-16-40). Diego Faria is supported by the Portuguese Foundation for Science and Technology (FCT).

References

1. Klatzky, R.L., Lederman, S.: Chapter 4 Intelligent exploration by the human hand. In: Dextrous robot hands. Springer-Verlag New York, Inc., New York, NY, USA, pp. 66--81 (1990)
2. Allen, P.K., Roberts, K.S.: Haptic object recognition using a multi-fingered dextrous hand. In: IEEE International Conference on Robotics and Automation, pp. 342--347 (1989)
3. Bierbaum, A.; Welke, K., Burger, D., Asfour, T., Dillmann, R..: Haptic Exploration for 3D Shape Reconstruction using Five-Finger Hands. In: IEEE/RAS International Conference on Humanoid Robots. (2007)
4. Bierbaum, A. Gubarev, I., Dillmann, R.: Robust shape recovery for sparse contact location and normal data from haptic exploration. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 3200--3205 (2008)
5. Faria, D. R., Prado, J. A. S, Drews Junior, P., Dias, J.: Object shape retrieval through grasp exploration. In: ECMR-European Conference on Mobile Robots. (2009)
6. Moravec, H.: Sensor fusion in certainty grids for mobile robots. AI Mag. 9, pp. 61--74 (1988)
7. Elfes, A.: Using occupancy grids for mobile robot perception and navigation. Computer 22, pp. 46--57 (1989)
8. S. Thrun: Robotic mapping: a survey. In: Exploring artificial intelligence in the new millennium. Morgan Kaufmann Publishers Inc., San Francisco, CA,USA, 1-35 (2003)
9. Ferreira, João F., Pinho, C., Dias, J.: Active exploration using bayesian models for multimodal perception. In: ICIAR '08: Proceedings of the 5th International conference on Image Analysis and Recognition, Berlin, Heidelberg, Springer-Verlag, pp. 369--378 (2008)
10. Polhemus Liberty Electromagnetic Motion Tracking System: http://www.polhemus.com