Human reach-to-grasp generalization strategies: a Bayesian approach

Diego R. Faria, Ricardo Martins and Jorge Dias
Institute of Systems and Robotics
Department of Electrical Engineering
University of Coimbra - Polo II
3030-290 Coimbra, Portugal
Email: {diego,rmartins,jorge}@isr.uc.pt

***Abstract*— In this work we present a general structure of an initial Bayesian framework to describe the mechanisms underlying the human strategies that define the appropriate characteristics of the reach-to-grasp movements to specific contexts, objects and how these strategies can be extended and replicated to other contexts and objects. The Bayesian framework uses information extracted from data about the pose of the hand, fingers and head acquired by a magnetic tracker device, finger flexure data acquired by a data glove, as well as, data about eye gaze and saccade movements of the subject.**

I. Introduction

Some of the most performed actions by humans in their daily activities involve the manipulation of objects from one place to another and its in-hand manipulation to adjust the pose of the object with the final goal of the action which will be performed with it.

Typically, the global hand's trajectory during a manipulation task can be segmented in different stages: reach, lift, transport and release[6]. We focus our attention in the reach stage (reach-to-grasp movement). The conclusions of several previous studies suggest that the human decision making process of how (type of grasp [5], hand trajectory characteristics and maximum grip aperture [4][10]) the reach-to-grasp movement is performed, is influenced by different factors, such as the presence or absence of visual feedback [9] or the eye-hand coordination in object manipulation with and without obstacles [7].

We intend to develop a Bayesian framework which will describe what are the specific and elementary factors which are involved in the planning and control of a reach-to-grasp movement. This can contribute to the description of the human strategies that are used to extract the physical properties of the object which will be grasped and how humans relate those highlighted physical properties with the object's possible applications (affordance). Our main concern is to provide a framework that can define a general description of the human strategies (generalization), instead of a model that is specific and restricted to the learnt object/context. An initial structure of the framework is presented in section II. The experiments carried out to validate this initial framework will be used to refine and detail it in the future.

II. Bayesian Approach

Based on the studies about the grasp model [8] and studies of neuroscience of grasping [4], we intend to develop a Bayesian framework to generalize the human strategies to perform a reach-to-grasp tasks and to transfer and integrate this knowledge to robotic platforms.

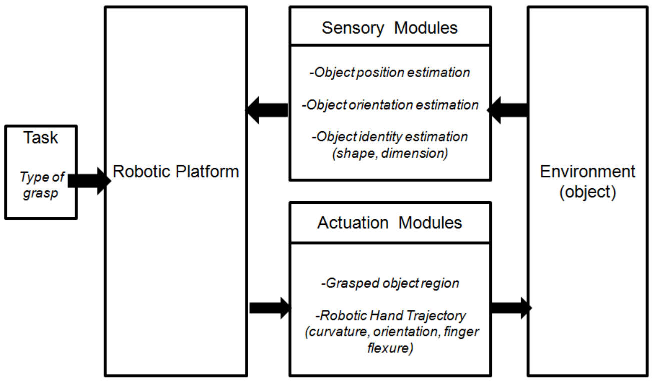

Figure 1 - Overview of the expected inputs and outputs of the Bayesian framework integrated in a robotic platform.

Figure 1 shows a schematic representation of the global overview of the expected inputs and outputs of the framework. The robotic platform must have the ability to estimate the identity (O_i) (dimensions, shape), position (O_p) and relative orientation (O_o) of the object to be manipulated. The robotic platform must be instructed about the task (T) which should be performed with the object. In this initial approach the considered possible tasks are restricted to the indication of what type of grasp must be performed, such as top-grasp or side-grasp. In future developments, it is expected to increase the set of possible tasks and automatically establish the relation between the object's identity and its affordances. The type of task can be indicated through a pre-defined instruction or it can be extracted from a human demonstration. In the latter situation, the reach-to-grasp movement can be classified as top-grasp or side-grasp by the Bayesian classifier developed in a previous work [5]. This set of information is combined in the Bayesian framework, which estimates the region of the object which will be grasped (R) and several characteristics of the trajectory and behaviour of the robotic hand during the approach to the object: hand trajectory curvature (H_c), hand orientation (H_o) and the level of fingers flexure (F). These expected outputs of the Bayesian framework can be written using the Bayesian formalism as $P(R|o_p, o_o, o_i, t)$, $P(H_c|o_p, o_o, o_i, t)$, $P(H_o|o_p, o_o, o_i, t)$, $P(F|o_p, o_o, o_i, t)$, respectively. These informations are transmitted to the robot motor control modules which are responsible by its execution.

A. Learning

The learning phase corresponds to the acquisition of knowledge from human demonstrations of the required tasks performed with different objects (variable identity, size and shape) in different contexts (positions, orientation).

Six magnetic sensors of the *Polhemus Liberty* device [2] are used to provide the pose (6 DoF) of the fingers and head. Five of them are attached to a data glove [1] on the fingertips region and the other one is attached to the head-mounted eye-tracker device. The head-mounted eye-tracker device, such as [3], is used to extract information about the regions of the object which are observed during the reach-to-grasp movement. A magnetic sensor is also attached to the grasped object, in order to acquires its initial position and orientation.

This acquired knowledge is represented by Learned Histograms. For each implemented task, object orientation and object identity, four types of learning tables are defined: Trajectory Curvature Learned Table (CLT), Hand Orientation Learned Table (HLT), Finger Flexure Learned Table (FLT) and Grasped Region Learning Table (RLT). The definition of the CLT, HLT are described in more detail in a previous work [5]. The CLT represents the probability of a specific curvature in each of the eight segmented parts of a normalized hand trajectory. Similarly, the HLT and FLT represent the probability of a specific hand orientation and level of fingers flexure, respectively, in each of the eight trajectory segments. The RLT represents the probability of the object being grasped in each of three defined regions: the upper, middle or bottom part of the object. The probability is determined based in the principle studied in [7], which relates the observed regions of an object during a manipulation task with the region of the object where the grasping is performed. The mapping of the observed regions of the object is made by an eye-tracker device such as [3].

B. Generalization Strategy

As previously referred, for all types of learned features we have learned tables (mean histograms calculate from all observations): $P(R|o_p, o_o, o_i, t)$, $P(H_c|o_p, o_o, o_i, t)$, $P(H_o|o_p, o_o, o_i, t)$, $P(F|o_p, o_o, o_i, t)$.

For a estimated object identity, orientation, position and pre-defined task, it is possible to determine the appropriated learned table for each of the wanted features: hand trajectory curvature, hand orientation, finger flexure and grasped region. The estimated distance between the robotic platform and the object is segmented in eight parts, as the data collected during the construction of the learned tables. For each of these segments, it is determined the most probable value of each of the features referred before, based in the learned tables previously identified.

For instance, to generalize the trajectory curvatures, we get the higher probability of the features in each hand displacement, in a curvature learned table (CLT) for the top-grasp , so that we have the curvatures: U-U-U-UR-DR-DR-D-D, and for each curvature we have associated the (r, θ, ϕ) information, that is, the angles of each curvature and the information of forward or backward direction. The same happens for the hand orientation, eg. for top-grasp we have the features with higher probability in each hand displacement: T-T-S-T-T-T-T-T (T: top orientation, S: side orientation, more details see [5]) and the θ angles of the hand plane. The same for the finger flexure (joint angles).

The learned information: (r, θ, ϕ) curvatures; θ angle of the hand plane (hand orientation); the fingers joint angles acquired from $P(F|o_p, o_o, o_i, t)$ and the region of the object to grasp given $P(R|o_p, o_o, o_i, t)$, is mapped to the robot referential to the robot reproducing the same task according with its DoF.

III. CONCLUSION

In this preliminary work, we present an initial approach to a framework to represent the human strategies to perform reach-to-grasp tasks. The analyses that were made will support and guide our future works. We focus our attention in the analysis of the visual behavior (conjugation of head and eyes movements) of the subjects, in the tracking of hand, fingers and head pose, as well as, the flexure level of the fingers along the reach-to-grasp movements.

ACKNOWLEDGMENT

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

REFERENCES

[1] *http://www.5dt.com/products/pdataglove5u.html*. Accessed in 20 April 2009.

[2] *http://www.polhemus.com/*. Accessed in 20 April 2009.

[3] *http://www.smivision.com/en/eye-gaze-tracking-systems/products/iview-x-hed.html*. Accessed in 20 April 2009.

[4] U. Castiello. The neuroscience of grasping. *Nature Reviews Neuroscience*, 6(9):726–736, 2005.

[5] D. R. Faria and J. Dias. 3d hand trajectory segmentation by curvatures and hand orientation for classification through a probabilistic approach. In *(to appear) 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS'09.*

[6] J. R. Flanagan, M. C. Bowman, and R. S. Johansson. Control strategies in object manipulation tasks. *Current Opinion in Neurobiology*, 16(6):650–659, 2006.

[7] R. S. Johansson, G. Westling, A. Bckstrm, and J. R. Flanagan. Eye-hand coordination in object manipulation. *The Journal of Neuroscience*, 21(17):6917–6932, 2001.

[8] E. Oztop, N. S. Bradley, and M. A. Arbib. Infant grasp learning: a computational model. *Experimental Brain Research*, 158:480–503, 2004.

[9] M. K. Rand, M. Lemay, L. M. Squire, Y. P. Shimansky, and G. E. Stelmach. Role of vision in aperture closure control during reach-to-grasp movements. *Experimental Brain Research*, 181(3):447–460, 2007.

[10] M. K. Rand, Y. P. Shimansky, A. B. M. I. Hossain, and G. E. Stelmach. Quantitative model of transport-aperture coordination during reach-to-grasp movements. *Experimental Brain Research*, 188(2):263–274, 2008.