Active learning for robot manipulation

Antonio Morales, Eris Chinellato
Robotic Intelligence Lab.
Universitat Jaume I
Castellon, Spain
email: {morales, eris}@uji.es

Andrew H. Fagg
Laboratory for Perceptual Robotics
University of Massachusetts
Amherst, Massachusetts, USA
email : fagg@cs.umass.edu

Angel P. del Pobil
Robotic Intelligence Lab.
Universitat Jaume I
Castellon, Spain
email: pobil@ieee.org

***Abstract*— Learning techniques in robotic grasping applications have usually been concerned with the way a hand approaches to an object, or with improving the motor control of manipulation actions. We present an active learning approach devised to face the problem of visually-guided grasp selection. We want to choose the best hand configuration for grasping a particular object using only visual information. Experimental data from real grasping actions is used, and the experience gathering process is driven by an on-line estimation of the reliability assessment capabilities of the system. The goal is to improve the selection skills of the grasping system, minimizing at the same time the cost and duration of the learning process.**

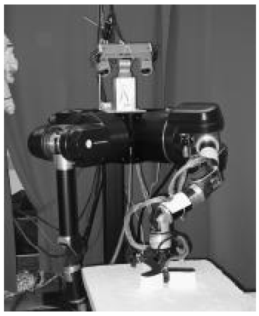
I. INTRODUCTION

Manipulation skills for service robotic applications require the development of techniques able to manage the natural complexity and uncertainty of such applications. In order to deal with these circumstances approaches based on the use of sensorial information and the application of learning techniques have been proposed by many different authors. Here, we face the problem of visually-guided grasp selection. Given an object, many different feasible grips can be performed on it, and it is thus critical to characterize the quality of candidate grips in order to execute the most reliable ones.

Applications of robotic manipulation have not been an exception in the use of learning techniques. In general, these approaches have been more concerned with the way a hand approaches to an object [14, 16] than in the best hand configuration for grasping a particular object. The object is usually abstracted by a simple volumetric model. Other works use learning techniques to discover, or simply tune, the motor control parameters that produce certain manipulation actions [2, 4, 7, 15].

More directly related to our work is the the approach followed by Coelho Jr, et al. [5], uses visual features to categorize and indicate to a haptic controller the best starting hand orientation. This controller starts, then, a tactile exploration of the object with a dexterous robot hand aimed at finding the best hand configuration in terms of applied wrench.

In this paper we focus on the use of active learning techniques that try to use experience of real grasping actions to tune the behavior and the reliability assessment capabilities of the grasping system. According to this learning paradigm, the agent is allowed to interact with its environment. It can execute actions which have an impact on the generation of training data. *Exploration* refers to the process of selecting actions in active learning. In the framework of our problem, the actions are the different candidate grips, at a given moment. Actions are selected by the agent in an ”intelligent” way in order to minimize the cost and duration of the learning process.



(a) UMass Torso

(b) Barrett Hand

Fig. 1. The UMass Humanoid Torso and the Barrett Hand at the Laboratory for Percetual Robotics of the University of Massachusetts, `http://www-robotics.cs.umass.edu`

We represent each grip as a point in a multidimensional space, and we use a procedure to predict a query point based on its similarity to its neighbors. This is a case of *instance-based* also known as *memory-based learning* [1], which is a numeric version of the more symbolic *case-based reasoning* [19]. These approaches do not construct an explicit representation of the target function when training samples are provided, but simply store them.

The main contribution of this paper is an exploration algorithm that makes use of a problem representation previously built to decide the next action, the grasp to be executed, in order to obtain a better knowledge of the environment with a lower cost, that is, with a minimum number of executions (see sec. V). The work presented in this paper is based on results and techniques developed in previous papers. Sections II to and IV develop the project framework of the current paper. Finally, in section VI we describe the validation of this algorithm by using real experimental data.

II. PROJECT BACKGROUND

The robotic grasping system we have been using throughout our project is the UMass Torso, developed at the Laboratory for Perceptual Robotics in the University of

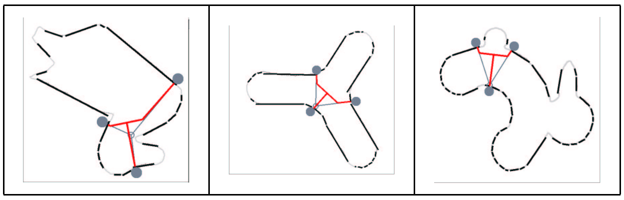


Fig. 2. Three samples grasps.

Massachusetts [18]. This humanoid robot consists of two Whole Arm Manipulators from Barrett Technologies, two Barrett hands with tactile sensors and a BiSight stereo head (see figure 1).

The main modules/steps of the functioning of this robotic grasping system are the following:

1 **Image processing:** The stereo vision system of the robot estimates the two-dimensional location of un unknown planar object placed on the table, which is the grasping target. A monocular image of the object for surface curvature analysis is provided and analyzed to extract the object contour and identify triplets of grasping regions on it;

2 **Grasp synthesis:** Many feasible candidate grasps are generated, by selecting the grasping points for each region triplet and the finger configurations that can be applied to the object in order to perform a grip; we use the kinematic and geometric constraints of the three-finger Barrett Hand. A *grasp* is defined by a set of three contact points on the object contour, and the corresponding force directions, applied along the directions in which the fingers close. Fig. 2 shows some examples of grasps. A particular kind of grasp is the *virtual two-finger grasp*, in which the two fingers opposing the thumb are positioned on the same grasping region and act in the same direction, facing the thumb;

3 **Grasp selection:** The candidate grasps are characterized in order to perform an 'intelligent' selection of the grasp to execute;

4 **Execution:** The hand is preshaped and positioned above the object. It moves down, closes the fingers so that the object is grasped, lifted and transported to a designated location. All this is performed with support of visual and tactile feedback.

The work presented in this paper is mainly focused on the third step. Details about the other sections of a system of this kind, concerned with the generation of candidate grasping configurations, are given in [12, 13].

III. GRASP CHARACTERIZATION SCHEME AND RELIABILITY MEASUREMENT

A. Grasp characterization

A characterization scheme which provides a way to describe grasps has been developed, so that a learning stage can be applied to the process. The scheme includes nine high-level features, all used for each of the candidate grasps, so that every grasp is represented by nine measurements becoming a point in a *9-dimensional* space.

The features have been designed in order to meet the following requirements:

- **Vision-based computation.** The features are computed from visually-extracted information.
- **Hand constraining.** Features take into account particular characteristics of the hand.
- **Location and orientation invariance.** Displacements and rotations of the object do not affect the values of the features.
- **Object independence.** Grasps with the same physical properties have the same characterization independently of the object for which they are computed.
- **Physical meaning.** Features are computed to measure physical properties relevant to grasping.
- **Stability and reliability.** Features consider stability and reliability hazards of a grasp.

The nine features derive from a set of grasp quality criteria, defined in [3]. Such criteria were normalized and a stability analysis was performed on them. They have been further improved in [11].

B. Experimental measurement of grasp reliability

A key issue in our experimental approach is the definition of a practical measurement of the reliability of a grasp. In order to do this a single object is placed on a table within the robot workspace. Using visual information the robot locates the object and computes a set of feasible grasp configurations. One of the configurations is selected, either manually by a human operator, or automatically by the robot, and executed.

If the robot has been able to lift the object safely, a set of stability tests are applied in sequence. These are aimed at measuring the stability of the current grasp. They consist of three consecutive shaking movements of the hand which are executed with an increasing acceleration. After each movement the tactile sensors are used to check whether the object has been dropped off.

This protocol provides us with a qualitative measure of the success of a grasp. Thus, an experiment may result in five different reliability classes: *E* indicates that the system was not able of lifting the object at all; *D, C, B* indicate that the object was dropped, respectively, during the first, second, or third series of shaking movements; finally *A* means the object did not fall and was returned successfully to its initial position on the table. Hence, we define $\Omega = \{A, B, C, D, E\}$ as the set of reliability classes.

IV. PREDICTION SCHEME

The learning methodology that we propose is composed of two main parts, an off-line and an on-line learning components.

First, an off-line prediction scheme computes the most likely reliability class of an untested grip, using previous experience as reference. This component assumes the existence of a set of previously executed grasps, i.e. experience, which descriptors values and reliability classes are known.

Second, the on-line component, which will be referred to as *e*xploration function, is responsible for building such a set of previous attempts by successive selection of the most appropriate grip candidates. This paper is mainly focused on this second component. First we outline the off-line learning component [3, 11].

Information about N executed triplets constitutes the previous experience dataset. Each grip $g_i, i = 1 \dots N$ is described by the nine visual features $q_1, \dots q_9$ introduced in subsection III-A. The 9-dimensional space G_S is formed by the ranges of the features. Moreover, we have also recorded the experimental grip performance and have assigned it to a class $\omega_i \in \Omega$ for each g_i.

A. Weighted KNN classification rule

A prediction function has the form $F(g) = \bar{\omega}$ where $g \in G_S$ and $\bar{\omega} \in \Omega$. Posed in these terms, the prediction problem is essentially a classification problem. There exists a wide bibliography on the building of such functions based on the Bayesian decision theory [6]. In our previous works we developed different classification approaches, among which one based on the *voting k-nearest neighbor (KNN) rule* [6, 9] showed the most convenient properties [11].

This approach is classified among the non-parametric techniques, and it does not assume any density distribution of the descriptors and the classes. To predict the class of a *query* point g_q, the KNN rule counts the K-nearest neighbors and chooses the class that most often appears in them.

In our implementation we have introduced some modifications to the basic schema. First we use the euclidean metric for measuring the distance between the points in the G_S. We weighted the contribution of each of the KNN points according to its distance to the query point. This gives more importance to the closer points. The kernel function used is $K(d) = \frac{1}{1+(d/T)}$, where T is an adjustable parameter, and d is the distance.

We define $KNN(g_q) = \{(g_i, \omega_i), i = 1 \dots k, g_i \in G_S, \omega_i in \Omega\}$ as the *k* closest points to g_q and d_i their corresponding distances from g_q. The probability corresponding to a class $\bar{\omega}$ are computed using this expression:

$$p(\bar{\omega}, g_q) = \sum_{\substack{g_i \in KNN(g_q) \\ \omega_i = \bar{\omega}}} \frac{K(d_i)}{\sum_{g_j \in KNN(g_q)} K(d_j)} \quad (1)$$

Function P is also an expression of the posterior probability [9]. Our predictor would be defined as

$$F(g_q) = argmax_{\omega \in \Omega}\{p(\omega, g_q)\} \quad (2)$$

That is, the predicted class ω is the one with the largest probability $p(\omega, g_q)$.

Error and risk functions

The classification performance is measured in terms of successful or wrong classifications. Our classes have an important particularity, their qualitative order (i.e.: class A means a higher stability for a grip than any other class). Having this in mind, we penalize in a different way failures which are qualitatively small (i.e.: predicting B when the outcome is C) compared to larger ones (i.e.: predicting A when the outcome is D). This penalization is realized with the error function $E(\bar{\omega}, \omega)$, being $\bar{\omega}, \omega \in \Omega$, where $\bar{\omega}$ is the predicted outcome and ω the real one. Such function is easily implemented with a table, and there exist many criteria to define it [10].

The error function is composed with the posterior probability to obtain the *risk function*:

$$R(\bar{\omega}, g_q) = \sum_{\omega \in \Omega} p(\bar{\omega}) E(\bar{\omega}, \omega) \text{ ,where } \bar{\omega} \in \Omega \quad (3)$$

The class $\omega \in \Omega$ selected for the prediction is the one that minimizes the risk:

$$F_R(g_q) = argmin_{\omega \in \Omega} R(\omega, g_q) \quad (4)$$

Equations 2 and 4 show two different functions for the class prediction of a given query grip. The first one is based on the underlying assumption that any wrong classification has the same cost, whilst the second assigns a different cost depending on the type of misclassification, making it possible to introduce in the prediction step the qualitative ordering of the classes.

Finally, it is worthy to note that in [11] we compared the KNN prediction approach to other well-known classification techniques, like those based on Artificial Neural Networks (ANN). In that paper we showed that the KNN approach presented better predictive performance that the other methods.

V. ON-LINE EXPLORATION

The goal of the on-line exploration procedure is to select the next grasp to execute among a set of candidates. This selection is done in order to improve the predictive capabilities of the stored experience, i.e., the set of already executed grasps.

The algorithm we propose assumes that at any point during the training of the grasping system a set of candidate grips $g_i \in G_S$ is proposed and the algorithm has to select the next grasp to be executed. To accomplish this task, it can take into account the results of previous experiments.

The approach we propose for the selection is inspired in the idea hinted by Thrun [17], “queries are favored that have the least predictable outcome”. That is, those candidates which category is less predictable are preferred. This idea is based on the insight that such candidates are located in areas where the implicit model represented by the experience dataset is less clear, so that their execution should provide critical information for the classification task.

We implement this idea by defining the term *prediction confidence*. For every grip candidate g_i, a class $\omega_i \in \Omega$ is computed using the prediction schemes defined in the previous section.

The formal meaning of confidence varies depending on the expression used for the prediction, more precisely, whether considerations about the error cost are included in the computation of the probability of each class. We distinguish two cases in the formal definition of confidence. In the first, equation 2 is used for computing the prediction of a class. In this case the confidence of a prediction is expressed by $p(\omega_i, g_i)$. In the second case, that considers the different cost errors, prediction confidence is derived from equation 4, and its value is $R(\omega_i, g_i)$. Summarizing, the confidence of a prediction is defined by this expression,

$$G_{conf}(g_q) = \begin{cases} \max\{p(\omega|g_q), \omega \in \Omega\} & \text{from eq. 2} \\ \min\{R(\omega, g_q), \omega \in \Omega\} & \text{from eq. 4} \end{cases} \quad (5)$$

Once defined the notion of confidence, it is easy to describe the exploration function, though it also depends on whether the prediction scheme is based on conditional probability or conditional risk. In the former case, the candidate with minimum confidence value is chosen, in the latter, the candidate with maximum risk value is preferred.

Given a set of m grasp candidates $G_q = \{g_1, \ldots, g_m\} \subset G_S$, the exploration function is defined as,

$$X(G_q) = \begin{cases} argmin_{g_i \in G_q} G_{conf}(g_i) & \text{minimum confidence} \\ argmax_{g_i \in G_q} G_{conf}(g_i) & \text{maximum risk} \end{cases} \quad (6)$$

VI. Validation and Results

A. Experimental sample dataset

The methodologies described above need to be validated using real data from real experiments. In order to acquire a sample database large enough to validate the proposed methods, a series of exhaustive experiments have been carried out.

Four real objects have been built for the experiment, two with simple shapes and two with more complex ones. In order to build the sample database the four objects are presented to the grasping system, and a sufficiently large number of grips are executed. The reliability of these grips is obtained applying the test described in section III-B.

A particular execution of a grip configuration can be influenced by many unpredictable factors. To avoid this problem, each grip is executed a sufficiently large number of times, changing the initial location and orientation of the object. In this way, statistically significant conclusions can be obtained. A collateral consequence is that the samples obtained are naturally grouped depending on repeated grips.

The number of feasible grips that are computed for each single object is usually large, varying from several dozens to more than one hundred. This could lead to a non practical number of executions, so for each object only a few configuration grips are selected to be executed. This selection consists of the most representative configurations of each object. Each configuration grip is executed 12 times, 4 times for three different orientations of the object.

Finally, in order to study the grasping performances in different circumstances, several characteristics of the environment are tested. These are the weight of the objects and the friction coefficient. Two qualitative categories for each of the two conditions are distinguished: heavy and light objects, and low and high friction. The different weights are obtained with two different sets of objects, that have the same appearance, but are made of different materials (wood in the heavy case and polystyrene in the light). The different frictions are obtained by placing and removing latex fingertips on the metal-made fingers.

More than nine hundred samples divided in four different datasets were obtained from this exhaustive experimentation. Table I shows the number of different grasps executed and the percentages of grips that resulted in each class of Ω.

TABLE I
SAMPLE DATASETS

	E	D	C	B	A	Total
Light	102	84	33	27	18	264
Low	38.6%	31.8%	12.5%	10.2%	6.8%	(22)
Light	51	97	56	38	118	360
High	14.2%	26.9%	15.6%	10.6%	32.8%	(34)
Heavy	95	92	29	2	2	220
High	43.1%	41.8%	13.2%	0.9%	0.9%	(23)
Heavy	50	26	0	0	0	76
Light	65.8%	34.2%	0.0%	0.0%	0.0%	(11)

Sample distributions among classes for the different data sets. The figures in brackets in the "Total" column indicates the number of different grip configurations tested.

B. Validation of the exploration procedure

The performance index of the exploration procedure is the predictive capability of the set of samples selected/executed. This can be estimated predicting the classes of the samples contained in a *validation test*, and comparing them with the real reliability classes obtained from the practical experiments.

The most straight and realistic way of measuring and studying the performance of the exploration would have been to design a set of running experiments on the real robot. However, the potentially large number of executions that this would require in order to tune and obtain statistically significant results, represents a main drawback.

For this reason, we have designed a validation framework that uses the available data so as the robot were performing its training process according to the active learning procedure. In this situation the robot is meant to execute the following sequence of *selection-execution* actions:

1) One or more objects appear in the workspace of the robot. The possible candidate grasps for all of them are computed.
2) The robot selects one of the generated grasps by using the exploration function.
3) The chosen grasp is executed and the reliability test is applied.
4) The new grasp with its experimental reliability class is added to the experience dataset.

The core of the validation framework is composed by the algorithm 1. This algorithm tries to emulate the running of a training session, which is a sequence of selection-executions, that starts with experience at all.

Algorithm 1 SINGLE-RUN EXECUTION

```
REQUIRE: S = {g_q, ω_q}, g_q ∈ G_S, ω_q ∈ Ω % Sample dataset
  % pool: Points that have not been selected yet
  % memory: Points that have been already
  explored

  S_V ← select_subset(S) % Extracts a subset from S
  S_pool ← S − S_V
  S_mem ← ∅
  WHILE S_pool ≠ ∅ DO
    % Selects a subset of candidates from S_pool
    G_Q ← subset_candidates(S_pool)
    % Selects a candidate from the set of
    candidates
    g_x ← select_candidate (G_Q)
    % Updates the memory
    S_mem ← S_mem ∪ {g_x}
    S_pool ← S_pool − {g_x}

    % Computes the prediction errors
    Y_V ← predict (S_V using S_mem)
    ē ← compute_error (Y_V, S_V)
  END WHILE
```

The output of this algorithm is the evolution of the error measure. Several datasets appear in this algorithm: S_V is the validation set; S_{pool} is the pool dataset, which contains the points that have not been explored yet; S_{mem}, is the memory dataset, which contains the points already explored.

The operation `select_candidate` implements the exploration procedure explained in the previous section. `select_subset` selects a subset from a larger dataset; in this case it is used to extract the validation set from the initial whole sample dataset. This can be done randomly, or in terms of grasps and objects. In the latter cases, all grasps labeled with a given object or grasp are selected for the validation test. Similarly, `subset_candidates` chooses randomly a set of candidates from a larger set. It is possible to select in all iterations the same number of candidates, or allow this number to change. Since there is no clear argument against or in favor for any of the options we implement the simplest, that is selecting a fixed number of candidates. An interesting option to this is to select all the points, $G_Q = S_{pool}$.

The operation `predict` represents a prediction that computes a class for a set of candidates using a set of previous executed grasps. In the case of this paper we implement the *weighted k-nearest-neighbor rule* as it is explained in section IV-A. The value of *K* has been experimentally set to 31, and T for the kernel function is set to 1.5.

The error metric represented by the operation `compute_error` is based on the concept of *misclassification error distance*. The distance between two consecutive classes is defined as 1, that between A and C as 2, etc. In this way define the error distance $e(g_q) = \{0, \ldots, 4\}$ for the prediction of a given query

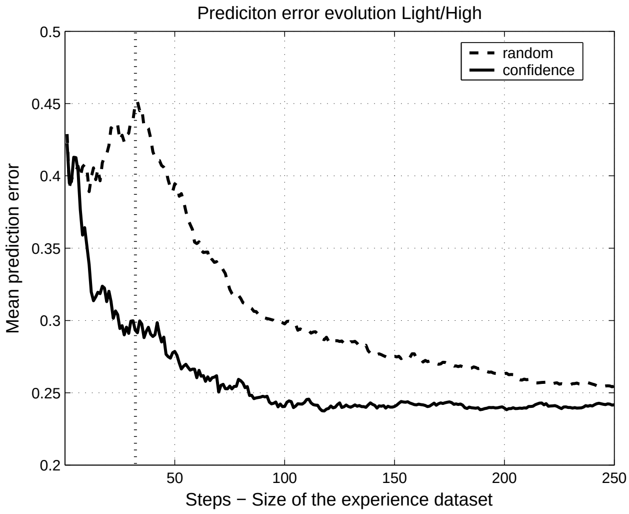


Fig. 3. Graph with the evolution of the prediction error when the Light/High sample dataset is used for training.

grip. Given a set of predictions $G = \{g_i, i = 1 \ldots n\}$, we compute the average error metric:

$$\bar{e}(G) = \sum e(g_i)/4 \tag{7}$$

As the presented algorithm is designed for a single run, the execution has to be repeated more than once in order to obtain statistically significant results. We have carried out a variation of an *m-fold cross validation*. The whole sample dataset is divided in m partitions and in each run one of them is assigned to S_V. Due to the stochastic nature of some of the operations, especially `subset_candidates`, it is also convenient to execute the whole cross-validation a sufficiently large number of times. The final result is the average of the error evolution $\bar{e}$ of the single runs.

Finally, in order to have a reference for comparison, we also implement a naive random exploration function, which chooses the next grasp to be executed randomly.

C. Results

Figures 3 presents (for a particular different datasets) the evolution of the prediction error for different sizes of the *experience dataset*. At each iteration of the algorithm described above 20 new samples are selected as query candidates. The error metric used is $\bar{e}$. The graphs in solid black line show the evolution of the error when the *minimum confidence* exploration procedure is used. The graphs in dashed lines show the evolution of the prediction error when the sample to execute is selected randomly among the set of candidates. This case represents the error evolution when no specific exploration rule is applied. The vertical dashed line indicates the point where the accumulated experience is larger than 31, which is the K used by the KNN prediction procedure.

We present the results obtained for the “Light/High”as representative of the other datasets, except for the “Heavy/Light” case, that was discarded right from the beginning because of the great disequilibrium it shows in the distribution of the reliability classes. The graph in figure 3 shows a very good performance, as the prediction error is able to fall to a stationary level in less than a hundred steps when using the proposed exploration strategy, clearly improving the random exploration.

Finally, is worthy to comment a work that is directly related to our problem. Kamon et al. [8] follow an approach similar to ours for estimating the quality of a two-finger grasp using stored experience. They keep a memory of successful grasps and a knowledge base of vision-based low-level quality features, that they use to estimate the quality of a set of candidates. They also define a test for the stability of a grasp based on the observation of the object gripped by the hand. The main difference with our work is that they keep an explicit separation between the description of a grasp and its quality features, thus keeping a different memory for each kind of data. On the contrary, our approach keeps a unified memory that characterizes uniquely a grasp and keeps information about its quality. Moreover, our work is focused on three-finger grasp. Finally, our results are supported by an exhaustive an larger amount of experimental data.

VII. Conclusion

In this paper we have presented an active learning approach to the problem of assessing robot grasp reliability. This algorithm tries to use the information that has been accumulated through successive grasping attempts. More precisely, the algorithm leads the succession of trials in order to increase the knowledge contained in the previous experience.

Data from real experiments and a validation methodology has been applied in order to test the appropriateness of the algorithm proposed. The results show both the usefulness of the exploration procedure. The approach leads in the right direction but important aspects remain yet to be improved.

Finally, this work is a completion of a larger project having as goal the development of a practical grasping system. This system, implemented on a humanoid robot, uses sensorial information and learning techniques to overcome the uncertainty that appear in manipulation tasks on complex working environments.

VIII. Acknowledgments

This work has been funded in part by the Ministerio de Ciencia y Tecnologa under project DPI2001-3801, by the Generalitat Valenciana under projects inf01-27, GV01-244, CTIDIA/2002/195, by the Fundaci Caixa-Castell under project P1-1B2001-28. The first and second author have been supported by the Ministerio de Ciencia y Tecnologia under a FPI Program graduate fellowship.

This work is also partly funded by grants CISE/CDA-9703217 and IRI-9704530, DARPA MARS DABT63-99-1-0004, and NASA/RICIS. The authors wish to thank Roderic Grupen, David Wheeler, Robert Platt and Danny Radhakrishnan, who have provided much of the foundation which allowed the robot experiments.

References

[1] D.W Aha. Lazy learning. *Artificial Intelligence Review*, 11:7–10, 1997.

[2] E. Cervera. *Perception-based Learning for Fine Motion Planning in Robot Manipulation*. PhD thesis, Universitat Jaume I, 1997.

[3] E. Chinellato, R.B Fisher, A. Morales, and A.P. del Pobil. Ranking planar grasp configurations for a three-finger hand. In *IEEE Intl Conf. on Robotics and Automation*, Taipei, Taiwan, September 2003.

[4] A. D. Christiansen, M.T Mason, and T.M. Mitchell. Learning reliable manipulation strategies without initial physical models. In *IEEE Intl Conf. on Robotics and Automation*, volume 2, pages 1224–1230, May 1990.

[5] J.A Coelhop Jr., J. Pieter, and R. Grupen. Developing haptic and visual perceptual categories for reaching and grasping with a humanoid robot. *Robotics adn Autonomus Systems*, 37:195–218, 2001.

[6] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification*. John Wiley & Sons, Inc., 2nd edition, 2001.

[7] M. Ferch and J. Zhang. Learning cooperative grasping with the graph representation of a state-action space. *Robotics and Autonomous Systems*, 38:183–195, 2002.

[8] I. Kamon, T. Flash, and S. Edelman. Learning to grasp using visual information. In *IEEE Intl Conf. on Robotics and Automation*, pages 2470–2476, Minnesota, April 1996.

[9] T. M. Mitchell. *Machine Learning*. McGraw Hill, 1997.

[10] A. Morales, Chinellato E, A.H. Fagg, and A.P. del Pobil. Experimental prediction of the preformance of grasps tasks from visual fetures. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pages 3423–3428, Las Vegas, Nevada, October 2003.

[11] A. Morales, Chinellato E, A.H. Fagg, and A.P. del Pobil. Using experience for assessing grasp reliability. In *International Conference on Humanoid Robots (Humanoids 2003)*, Karlsruhe, Germany, October 2003. on CD-ROM.

[12] A. Morales, P.J. Sanz, and A.P. del Pobil. Vision-based computation of three-finger grasps on unknown planar objects. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pages 1693–1698, Lausanne, Switzerland, October 2002.

[13] A. Morales, P.J. Sanz, A.P del Pobil, and A. H. Fagg. An experiment in constraining vision-based finger contact selection with gripper geometry. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pages 1711–1716, Lausanne, Switzerland, October 2002.

[14] M.A Moussa and M.S. Kamel. A connectionist model for learning robotic grasps using reinforcement learning. In *IEEE Intl. Conf on Neural Networks*, pages 1771–1776, Whasington, DC, May 1996.

[15] M. T. Rosenstein. *Learning to Exploit Dynamics for Robot Motor Coordination*. PhD thesis, University of Massachusetts Amherst, 2003.

[16] M. Salganicoff, L.H. Ungar, and R. Bajcsy. Active learning for vision-based robot grasping. *Machine Learning*, 23:251–278, 1996.

[17] S. Thrun. Exploration in active learning. In Michael A. Arbib, editor, *The Handbook of Brain Theory and Neural Networks*, pages 381–384. MIT Press, 1995.

[18] UMass humaniod torso. `http://www-robotics.cs.umass.edu`.

[19] D. L. Waltz. Memory-based reasoning. In Michael A. Arbib, editor, *The Handbook of Brain Thery and Neural Networks*, pages 661–662. MIT Press, 1995.