Active learning for robot manipulation

Antonio Morales [1] and **Eris Chinellato** [1] and **Andrew H. Fagg** [2] and **Angel P. del Pobil** [1]

Abstract.
This paper describes a novel application of active learning techniques in the field of robotic grasping. A vision-based grasping system has been implemented on a humanoid robot. It is able to compute a set of feasible grasps and to execute any of them and measure their actual reliability.

An algorithm aimed at predicting the performance of an untested grasp using the results observed on previous similar attempts is presented. The previous experience is stored using a set of vision-based grasp descriptors. Moreover, a second algorithm that actively selects the next grasp to be executed in order to improve the predictive quality of the accumulated experience is introduced.

An exhaustive database of experimental data is collected and used to test and validate both algorithms.

1 INTRODUCTION

Manipulation is one of the most useful skills in any robot system and constitutes a key component for many robotic applications on all kind of areas such as industrial, medical, service, and space robotics. In this paper we focus in a subfield of manipulation named *fixturing* that consists in the task of restraining or immobilizing objects with the fingers. We manipulate objects by caging them with the fingers and, then, using the robot arm to move and orient them.

Extensive research on this field during the last two decades has established a strong theoretical framework. However, most of this research has been based on perfect models or ideal operational conditions. These assumptions often become unrealistic in real world applications.

We face the problem of grasp selection. Given a 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.

The approach introduced in this paper uses experience of real grasping actions to tune the behaviour and the reliability assessment capabilities of the grasping system. More specifically, we follow an active learning approach. According to this paradigm, the agent is allowed to interact with its environment. More specifically, 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.

[1] Intelligent Robotics Lab., Univertitat Jaume I, Castellon, Spain. email: {morales, eris, pobil}@dsic.upv.es
[2] Computer Science Department, University of Massachusetts, Amherst, Massachusetts. email : fagg@cs.umass.edu

Each grip is characterized by a set of vision-based high-level features [2] that measure different aspects related with the stability of the grip. This permits to represent each grip as a point in a multidimensional space.

We present (sec. 4) a procedure to predict a query point based on its similarity to its neighbours. This is a case of *instance-based learning*, also known as *memory-based learning* [1]. These approaches do not construct an explicit representation of the model to be learnt when training samples are provided, but simply store them, and build the model when a query is presented.

We also present in sec. 5 an exploration algorithm that makes use of the problem representation previously defined 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 minimun number of executions.

Finally, we carry out an experimental validation of this methods using real data from repeated grasping actions of the robot (sec. 6). We develop (sec. 3.1) on the robot a practical test for measuring the reliability of a grip, and through it we collect an extensive set of samples from real grasping executions, and use them to tune, test and validate our methods. Moreover, we also develop.

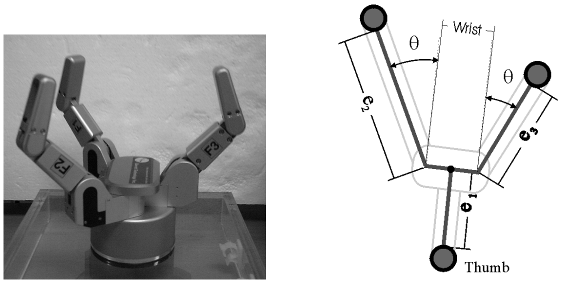
2 PREVIOUS RESEARCH: BACKGROUND

This paper presents partial results of a larger project that aims to provide an experimental approach to the grasping problem. Within this project we have implemented a robotic grasping system on the UMass humanoid torso, at the Laboratory for Perceptual Robotics in the University of Massachusetts [8]. This humanoid robot consists of two Whole Arm Manipulators from Barrett Technologies, two three-finger Barrett hands with tactile sensors at the fingertips and a BiSight stereo head.

The stereo vision system estimates the two-dimensional location of the target object on the table, and provides a monocular image for surface curvature analysis (see [6] for more details). Once a grip is selected (consisting of contact locations and a hand posture), 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.

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

1 **Image processing:** analyzes an image of an unknown planar object, extract its contour and identify triplets of grasping regions;
2 **Grip synthesis:** determines a number of feasible grasps selecting the grasping points for each region triplet; after that, generates finger configurations that could actually be applied to the object in order to perform a grip action;
3 **Grasp selection:** perform an 'intelligent' selection of the grip to execute;



(a) Photo (b) Kinematics

Figure 1. Barrett Hand, `http://www.barretttechnology.com`

4 **Execution:** execute the grip with support of visual and tactile feedback.

The work presented in this paper is mainly focused on the third step. Details about the other modules of a system of this kind, concerned with the generation of candidate grasping configurations, are given in [5, 6], though in the next subsections we introduce the necessary background concepts.

2.1 Grasp synthesis

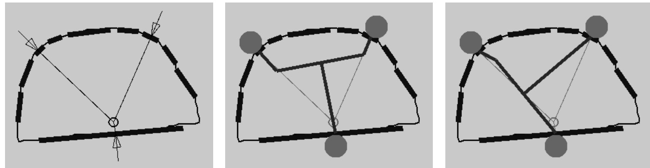
We define a *grasp* as the set of three contact points on an object contour, and the corresponding force directions, perpendicular to the contour, which meet in the grasp force focus. We call *hand configuration* each possible grip obtained applying the kinematics constraints of a robot hand to a grasp as defined above. To avoid misunderstandings, in all this text when referring to grasps and configurations together, the term grip is used.

We assume a real-time system acting in an unstructured environment, which detects unknown objects and, through analysis of visual data, selects and executes a stable grip of such objects.

Fast computation is necessary in order to achieve a real-time interaction with the external world. The ability to cope with uncertainties, in terms of knowledge of friction coefficients or visual and positioning errors, is a must in an uncontrolled environment.

With a perfectly homogeneous three-finger hand, for which the fingers are all the same, the three possible ways of combining fingers with contact points in a grasp are not distinguishable. This is not the case for the Barrett Hand, for which the kinematics of the thumb is different from that of the other two fingers. A photo of the hand is reproduced in Fig. 1(a). Its kinematics are depicted in Fig. 1(b). The hand has four degrees of freedom: the three finger extensions e_1, e_2, e_3 and the spread angle θ.

For each grasp there are three possible positions of the thumb. After deciding where to place the thumb, there are still potentially infinite ways of making the hand touch the object at three contact points. However, when the action line of the thumb is fixed as well, only one solution is possible. A one-dimensional search along all possible thumb force directions gives the best Barrett Hand configuration for a grasp after the thumb position has been defined . Thus, every grasp ideally generates three different configurations, one for each thumb position. When no solutions are found for a thumb position within a grasp, due to the constraints deriving from the hand geometry and kinematics, no corresponding configurations are produced.



(a) Grasp (b) Config. 1 (c) Config. 2

Figure 2. Generating configurations from a grasp

Typically, dozens of configurations can be generated for an object, mostly depending on the number of regions found. In Fig. 2(b) and 2(c) two configurations generated from the grasp of Fig. 2(a) are depicted.

3 GRASP CHARACTERIZATION SCHEME AND RELIABILITY MEASUREMENT

A characterization scheme to provide a way to describe grasps so that they can be used by the learning procedures has been developed. We have opted for a scheme that measures a set of properties of each grasp. In this way a grasp will be represented by *n* measurements becoming a point in an *n-dimensional* space. This scheme consists of nine of these high-level features that have been designed in order to meet the next 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.

To summarize, every grip is described by a nine-elements tuple, and therefore, can be abstracted as a point in a nine-dimensions space. This space would contain all the possible grip descriptors. For further details and a better explanation of all the descriptors the reader is referred to [2].

3.1 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.

4 PREDICTION SCHEME

The learning methodology that we propose is composed of two main components. The first is a prediction scheme that 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 grips having the values of the descriptors and their reliability class known.

The second component, that will be referred as exploration function, is responsible of building such set of previous attempts by successive selection of the most appropriate grip candidates. In this subsection we focus on the first component.

In theoretical terms a data set of previous experience is composed of N executed triplets. Each grip $g_i, i = 1 \ldots N$ is described by the nine visual features $q_1, \ldots q_9$ introduced in subsection 3. The 9-dimensional space G_S is formed by the ranges of the values of the features. Moreover, we have also recorded the performance of the grip and have assigned it to a class $\omega_i \in \Omega$ for each g_i.

4.1 Voting KNN classification rule

A prediction function has the form $F(g) = \bar{\omega}$ where $g \in G_S$ and $\bar{\omega} \in \Omega$. There exists a wide bibliography on the building of such functions based on the Bayesian decision theory [3]. In this paper we have chosen the approach of the nonparametric techniques, in particular the *voting k-nearest neighbor (KNN) rule* [4, 3], for modeling this function. The nonparametric techniques do not assume any density distribution of the features 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, the most voted.

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*$(g_q) = \{(g_i, \omega_i), i = 1 \ldots k, g_i \in G_S, \omega_i in \Omega\}$ 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:

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

Function p is also an expression of the posterior probability [4]. To conclude, our predictor would be formally defined by the expression $F(g_q) = argmax_{\omega \in \Omega}\{p(\omega, g_q)\}$. That is, the class predicted ω is the one with the largest probability $p(\omega, g_q)$.

5 ACTIVE LEARNING FOR EXPERIENCE ACQUISITION

The goal of the exploration procedure is to select the next grasp to execute among a set of candidates. This selection must be 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 takes into account the results of previous experiments.

The approach we propose for the selection is inspired in the idea hinted by Thrun [7], "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 intuition that such candidates are located in areas where the implicit model represented by the experience dataset is less clear.

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 scheme defined in the previous section. The confidence of that prediction is simply $p(\omega_i, g_i)$. In formal terms the prediction confidence for a grip g_q is defined as $F_{conf}(g_q) = \max\{p(\omega|g_q)\}, \omega \in \Omega$.

Once defined the notion of confidence, it is easy to describe the exploration function. It chooses the candidate with a minimum confidence value. Given a set of m grasp candidates $G_q = \{g_1, \ldots, g_m\} \subset G_S$, the exploration function is defined as,

$$F_{exp}(G_q) = \underset{g_i \in G_q}{\text{argmin}}\, F_{pred}(g_i) \tag{2}$$

Hereinafter, we will refer to this method as the *minimum confidence exploration*, or simply the risk exploration function.

6 VALIDATION AND RESULTS

6.1 Experimental sample dataset

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 this experiment. Two with simple shapes and two with more complex shapes. 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 3.1.

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, by varying the location and orientation in the presentation of the object. In this way, statistically significant conclusions can be reached. 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. The repetition above mentioned 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.

Table 1. SAMPLE DATA SETS

E	D	C	B	A	Total
51	97	56	38	118	360
14.2%	26.9%	15.6%	10.6%	32.8%	(34)

Sample distributions among classes for the sample data set. The figure in brackets indicates the number of different grip configurations really tested.

More than three hundred samples were obtained from this exhaustive experimentation. Table 1 shows the number of different grips executed and the percentages of grips that resulted in each class of Ω. This sample databases are used as training data is the validation of the two learning algorithms.

6.2 Validation of the prediction function

Two basic questions need to be answered about the prediction capabilities of the rules described in section 4.1: first, is it able to generalize across different objects?, and second, did we have enough data to properly construct a function ? To answer these questions we have developed a cross-validation method named *leave-one-grasp-out validation* similar to the well known *leave-one-out validation* and *n-fold cross-validation* [3]. This consists of the following steps: 1) given the whole data set, remove all the points of a particular grasp configuration and use this subset as validation set; 2) use the remaining samples for predicting the outcomes of the validation set and compute the mean error; 3) repeat steps 1) and 2) for all configurations. The validation error will be the mean error of the iterations of step 2). The reason for removing all the points of a configuration from the data set is that all the points of a particular configuration are very close in the G_S and the KNN rule would be affected by them instead of using points of unrelated configurations, farther in G_S.

The error metric 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 grip. Given a set of predictions $G = \{g_i, i = 1 \ldots n\}$, we define the average error metric $\bar{e}(G) = \sum e(g_i)/4$.

Table 2 shows the results obtained after repeating the validation procedure explained above and averaging the resulting percentages for all the cases of classification error distance. The final column shows the error values computed using error metric $\bar{e}$. The value of parameter k of KNN has been experimentally set to 31. We compare these results against the theoretical figures that would be obtained by a prediction method that would have chosen *randomly* the predicted class.

Table 2. FULL DATASET EXPERIMENT

	0	**1**	**2**	**3**	**4**	$\bar{e}$
random	23.5%	26.2%	20.3%	20.7%	9.3%	**0.415**
knn	51.1%	21.7%	13.3%	11.1%	2.8%	**0.223**

Percentages of misclassifications depending on the error distance. Distance 0 indicates the successful classification rates.

We are also interested in the sensitivity of the error with respect to the size of the data set. We can analyze it by modifying the second step of the *leave-one-grasp-out validation* procedure. Instead of using the whole remaining data set, we chose randomly a set of given size. This introduces a random factor, and to reduce the effect of this randomness we repeat this step a sufficiently large number of times. Figure 3 shows the results of this experiment. Again, the results are compared against the results that would be obtained by a *random* prediction function.

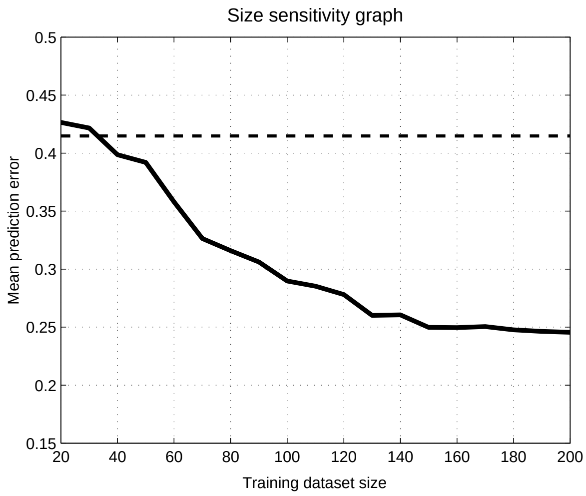


Figure 3. Evolution of the error when the size of the available data set varies. The Solid black line represents the errors obtained by the KNN prediction method, while the dashed line is the threshold of the random error.

Two main conclusions can be drawn from these results. First, the proposed prediction strategy clearly improves the performance of a *naive* random selection. In addition, from a practical point of view, when performing a strongly stochastic action like grasping an unmodeled real object with a robotic hand, an error between two neighbor classes can be considered acceptable, especially in the case of a false negative. Indeed, it means that the reliability of the grasp is only slightly better than the predicted one. Taking this into account, the sum of errors of distance 0 (successful classifications) and 1 (acceptable error) is about 72%. which is a quite good performance for a complex problem like this ones.

Second, the error is reduced as the size of the available data set increases. These two conclusions are satisfactory enough to suggest that this part of the methodology is adequate and justifies the use of the exploration procedure in a further step.

6.3 Validation of the exploration procedure

The performance of the exploration/selection procedure is measured by the predictive capability of the set of samples executed, which reliability class is known. This can be easily measured by using this dataset to predict the class of the samples contained in a secondary *validation test*. We have designed a validation framework that follows this principle. We also take inspiration from the running of the robot in the training environment or in a learning experiment. In this situation the robot will execute a sequence of *selection-execution* actions. Each of these actions will follow the next steps:

1. One or more objects appear in the workspace of the robot. The grasps for them are computed. These are the grasp candidates
2. The robot selects one of them by using the exploration function.

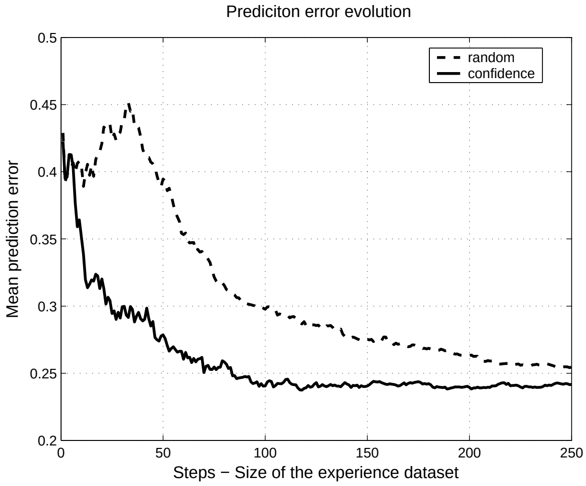


Figure 4. Graph with the evolution of the prediction error

3. The grasp is executed and the reliability test is applied.
4. The new grasp and the performance outcome are added to the experience dataset.

For the execution of the validation algorithm, we take the whole sample dataset available and extract a subset, *validation dataset* from it. The remaining is used as a *pool dataset*. In a sequence of selection steps, a small subset of *candidate samples* are extracted randomly from this pool. The exploration function, in our case, the *minimum confidence* rule, is applied to select one of these candidates. The selected candidate is added to the *experience dataset* and the discarded candidates are returned back to the pool. The performance measurement is done by using the samples in the *experience dataset* for predicting the samples in the *validation set*. The sequence is repeated until the pool dataset is emptied or it contains few samples.

This procedure is repeated a sufficient number of times varying the contents of the pool and validation datasets and the performance measurements for each size of the *experience dataset* are averaged.

Figure 4 presents the evolution of the prediction error for different sizes of the *experience dataset*, that is equivalent to the number of steps of the algorithm described in the above paragraphs. In this case the number of samples that are selected as candidates at each step is 20. The error metric used is $\bar{e}$. The graph in solid black line shows the evolution of the error when the *minimum confidence exploration* procedure (see sec. 5) is used. The graph in dashed lines shows the evolution of the prediction error when the sample to execute is selected randomly among the set of candidates. This case represents the evolution when no specific exploration rule is applied.

The results shown in this figure clearly indicate that the exploration function proposed in this paper reaches in less than a hundred trials the same level of prediction performance than a random unbiased selection procedure would reach in more than two hundred. This is a good argument in favor of using this exploration procedures since it would save more than a hundred trials, that is, more than a hundred executions of the whole reliability test, with the saving in time (about two minutes per execution) and mechanical wearing.

7 CONCLUSION

In this paper we have presented the development of a learning framework for assessing robot grasp reliability. This framework is based on two learning algorithms and a representation of the data which is built on a grasp characterization scheme composed of nine high level vision-based descriptors.

In this paper we focus on the two learning algorithms. The first one is aimed at predicting the reliability of an untested grip from its comparison to previous recorded attempts. This algorithm makes use of the voting k-nearest neighbour rule to perform such prediction.

The second algorithm, based on the idea of active learning, is an exploration rule that has to select among a set of candidate grips the next one to execute, having the goal of improving the predictive performance of the accumulated experience.

An experimental measurement of the reliability of a grasp have been developed and used to gather an exhaustive database of sample grips. Several validation frameworks that make use of this database, have been designed to test and validate the usefulness and properties of the proposed algorithms.

The results have proved that the algorithms proposed in this paper are able to carry out the expected tasks with a reasonable level of performance, despite the unpredictable nature of the task space.

Moreover, the experimental and practical approach followed in this paper indicates a path that service robotic applications willing to be used in every-day human environments could walk. The inclusion of active learning schemes in robot systems is an appropriate way to improve their adaptability to unmodeled or partially unknown environments and, thus, building real intelligent robot systems.

ACKNOWLEDGMENTS

This work has been funded in part by the Ministerio de Ciencia y Tecnología under project DPI2001-3801, by the Generalitat Valenciana under projects GRUPOS 03/153, GV01-244, CTIDIA/2002/195, by the Fundació Caixa-Castelló under project P1-1B2001-28. The second author has been supported by the Ministerio de Ciencia y Tecnología 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. Chinellato, A. Morales, R.B. Fisher, and A.P. del Pobil, 'Visual features for characterizing robot grasp quality', *IEEE Transactions on Systems, Man and Cybernetics (Part C)*, (2004). In Press.

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

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

[5] 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*, pp. 1693–1698, Lausanne, Switzerland, (October 2002).

[6] 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*, pp. 1711–1716, Lausanne, Switzerland, (October 2002).

[7] S. Thrun, 'Exploration in active learning', in *The Handbook of Brain Theory and Neural Networks*, ed., Michael A. Arbib, 381–384, MIT Press, (1995).

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