Learning Visually Guided Grasping: A Test Case in Sensorimotor Learning

Ishay Kamon, Tamar Flash, and Shimon Edelman

Abstract—We present a general scheme for learning sensorimotor tasks which allows rapid on-line learning and generalization of the learned knowledge to unfamiliar objects. The scheme consists of two modules, the first generating candidate actions and the second estimating their quality. Both modules work in an alternating fashion until an action which is expected to provide satisfactory performance is generated, at which point the system executes the action. This design decomposes the learning problem and thus simplifies it and allows direct generalization among objects for the quality estimation. Since the proposed scheme requires some initial knowledge about the task, we developed a method for off-line selection of heuristic strategies and quality predicting features, based on statistical analysis. The usefulness of the scheme was demonstrated in the context of learning visually guided grasping. We consider a system that coordinates a parallel-jaw gripper and a fixed camera. The system learns successful grasp configurations using a special coding, which allows it to apply stored examples to unfamiliar target objects. The system learns to estimate grasp quality by learning a function from relevant visual features to the quality. An experimental setup using an AdeptOne manipulator was developed to test the scheme. The system demonstrated an ability to grasp a relatively wide variety of objects, and its performance had significantly improved with practice following a small number of trials.

I. INTRODUCTION

A GENERAL goal of robotics is to develop systems that improve their performance based on experience, since complete information about the robot and the environment is not available in realistic settings. In recent years, there has been a growing interest in incorporating learning into robotic systems [4]. In the context of object manipulation, for example, physical characteristics of the robot and the environment can be learned from interaction (e.g., friction between a gripper and target objects), and the proper way to handle unknown objects can be learned by experimenting. Learning can use either symbolic or numeric representations. Symbolic encoding includes a variety of means, from decision trees [22] to classification rules [21]. Numerical encodings are usually in the form of artificial neural networks (e.g., radial basis function (RBF) networks [3]), nearest neighbor (NN) rules [13], or coefficients that describe discrimination hyperplanes (e.g., perceptron [17]). Kaelbling [9] has pointed out that numerical oriented methods tend to be more robust with respect to noise in examples, while symbolic representations are easier to interpret.

Grasping is a basic and common task in robotics, and there is much research on various aspects of the grasping process. Previous work on methods for choosing grasp configurations can be divided into two main approaches. First, the “analytic” approach takes a model of the target object and finds optimal grasping points on it, relative to some criteria for optimality [2], [12], [15], [25]. In the context of sensorimotor learning, we look for a function $f_2: S \rightarrow A$ that maximizes $G_O$, where $S$ is the sensory information, $A$ is the action, and $G_O$ is the observed quality of the action. The learner receives the state $S$ as input, chooses an action $A$ and then gets the grade $G_O$ from a teacher. This formulation is usually categorized as reinforcement learning. Second, the alternative “comparative” approach generates a set of candidate grasps, evaluates the quality of each grasp, and chooses the best candidate [1], [6], [7], [26]. In the context of sensorimotor learning, we look for a function $f_2: S \times A \rightarrow G_E$, which gets both the sensory information $S$ and the action $A$ as input, and returns the estimated grade $G_E$. This formulation is categorized as supervised learning since the teacher feedback, which is the observed quality $G_O$, is the desired output of $f_2$.

In recent years, there has been a growing interest in integration of sensing capabilities into robotic systems [8], [11]. Several papers consider the problem of visually guided grasping of unknown objects, using a parallel-jaw gripper. A system for grasping three-dimensional (3-D) objects, which drives a camera to explore the target object and chooses optimal grasping points on antipodal planar patches, is reported in [23]. Other papers [1], [24] emphasize the use of task-depending representations, in contrast to choosing optimal grasping points based on the reconstructed shape of the target object. A discussion about several aspects of visually guided grasping appears in [20].

Only few works deal with the problem of learning to grasp. Dunn and Segen [5] presented a system that first tried to recognize its target object using a stored library. If the object was recognized, a stored grasp was applied to it. Otherwise, the system tried to grasp the unknown object by trial and error. Tan [22] used a set of features to distinguish among objects. A decision tree was built in the training stage, making it possible to distinguish among the different objects by taking

Manuscript received May 19, 1994; revised September 22, 1995, January 7, 1997, and July 5, 1997. This paper was presented in part at the IEEE International Conference on Robotics and Automation, Apr. 1997. This work was supported in part by the National Foundation for Research and Development, Israeli Ministry of Science. The work of S. Edelman was supported by BARD under Contract US 1992-91.

I. Kamon is with the Department of Computer Science, Technion, Haifa 32000, Israel (e-mail: ishay@cs.technion.ac.il).
T. Flash is with the Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 76100, Israel.
S. Edelman is with the School of Cognitive and Computing Sciences, University of Sussex, Brighton BN1 9QH, U.K.

Publisher Item Identifier S 1083-4427(98)02742-8.
sensory measurements. In the working stage, a sequence of actions was planned for recognizing the target object and it was then grasped, using a predefined grasp configuration. No generalization to unfamiliar objects was performed by those systems. Salganicoff and Bajcsy [19] presented a general framework for learning sensorimotor tasks. They suggested to approximate the function \( f_2 : S \times A \mapsto G_E \) from examples. Given such an approximation scheme and a set of sensory parameters obtained for a new situation, they suggested a method for finding the action parameters that are expected to provide good performance. In other words, they tried to solve the “analytic” problem (given sensory information find an action) using the formulation of the “comparative” approach (the function \( f_2 \) is learned from examples). A recent paper by Moussa and Kamel [14] describes a connectionist scheme that learns generic grasping functions. These functions are not object-specific, and therefore they allow generalization among objects. This approach is somewhat similar to the concept of learning in the quality parameters space, that we present in this paper.

II. THE SCHEME FOR LEARNING SENSORIMOTOR TASKS

We present a general scheme for learning sensorimotor tasks which allows rapid on-line learning and generalization of the learned knowledge to unfamiliar objects. The proposed scheme consists of two modules, the first generating candidate actions and the second estimating their quality. Given a new task, the working system first uses its sensors to perceive the situation. The system then generates a sequence of candidate actions, based on examples from earlier successful actions and heuristic knowledge. Each candidate action is evaluated by a quality estimation function which was learned from earlier examples. As soon as an action gets a high grade from the quality estimator, the system performs this action, evaluates its actual quality and uses this feedback information to update its knowledge-base.

The design of the scheme decomposes the learning problem into two subproblems: learning to generate candidate grasp configurations and learning to predict the quality of a given grasp. Generating candidate grasps requires domain-specific knowledge. We solve this problem by storing successful examples from earlier grasps, using a special coding which allows the system to apply stored examples to unfamiliar target objects. Predicting grasp quality is formulated as learning a function from examples. We choose a subset of features, which are correlated with grasp quality and can be extracted from the input image, and term them the quality parameters \( Q \). The system learns the function from the quality parameters \( Q \) to the estimated grade \( G_E, f_2 : Q \mapsto G_E \).

The space of quality parameter has several advantages for learning. When the target object shape is encoded explicitly in the parameter space, as was suggested in [19], generalization among objects is difficult. In contrast, each point in the quality parameters space represents a class of grasp configurations that can be used on different objects, thus achieving generalization among objects (Fig. 1). When shape is encoded, the number of shape parameters grows with the complexity of the target objects, and that increases the number of examples necessary for learning. In contrast, the number of quality parameters is fixed and does not depend on the complexity of the target objects. When shape is encoded, parameters that encode shape far from the grasping points may be irrelevant for
quality prediction while relevant information is implicit, hence the function which is the target of learning is expected to be complicated. In contrast, the mapping from the quality parameters to the grade is expected to be relatively simple and smooth.

We may rely on a few quality parameters, which constitute only a partial model of the grasping process, since the system can visually observe the outcome of the grasping trials. The observed quality is used as the teacher feedback in a supervised learning process, which is an integral part of the proposed scheme. Since the system learns the correlation between visual features and the overall observed quality, it can adjust to non-visual features, e.g., the force and the contact characteristics of the gripper.

The proposed scheme requires three types of initial knowledge: a heuristic strategy for choosing grasping points on unknown objects, a set of informative quality parameters, and a candidate action representation which allows the system to apply stored examples to unfamiliar target objects. We present a method for off-line selection of the heuristic strategy and the quality parameters using statistical analysis. The overall flowchart of the scheme, including both off-line selection and the on-line working system, is presented in Fig. 2. The rest of the article is organized as follows. Off-line selection of the initial knowledge is discussed in Section III. In Section IV, we present the on-line working system. Experimental results are presented in Section V, followed by discussion in Section VI.

III. OFF-LINE SELECTION

To facilitate the system’s on-line learning, we supply it with initial knowledge, which includes a heuristic strategy for choosing grasping points on unknown objects and a set of informative quality parameters. This knowledge is selected off-line by applying statistical analysis to the results of simulated grasping trials. We used simulated trials since a large amount of data was necessary for the statistical analysis. Twenty synthetic images, obtained by cross-sectioning random generalized-cone objects, were used as target objects. For every input image, the main axis of the synthetic object was known, and the center of mass was calculated assuming horizontal main axis and constant 3-D density. Every grasp configuration was assigned a grade by a mechanical model of grasping. The model took into consideration two aspects: the stability of the fingers and the resistance to rotation of the grasp (for more details see Appendix A). Next we describe the selection of the heuristic strategy and the statistical methods for choosing the quality parameters. We then present a novel representation for grasp configuration, which is based on the selected heuristic strategy and quality parameters.

A. Choosing Heuristic Strategies

A heuristic strategy is used for choosing grasping points when stored knowledge from earlier trials is not applicable to a novel situation. In this case, candidate grasp configurations are generated by the heuristic strategy until the predicted quality of a candidate grasp is satisfactory, at which point the system performs the grasp. We compared the average performance of several strategies in the following way. Each heuristic strategy generated 2000 grasp configurations, 100 for each synthetic object, and the mechanical model assigned a grade for each configuration. Next, we briefly describe the tested strategies and their performance (for more details refer to [10]).

The initial strategy $s_1$ randomly chose two grasping points on the boundary. In strategy $s_2$, the second grasping point $p_2$ was chosen opposite to the first grasping point $p_1$, by considering the orientation at $p_2$. In strategies $s_3, s_4, s_5$ we tested several distributions for random variables which determined the distance from the center of mass and the location of the second point relative to the first. The strategy $s_5$ yielded the best mean grade 70 (compared to 8 for $s_1$), and the probability to get a grade higher than 80 was 0.57 (compared to 0.02 for $s_1$). The advantage of strategy $s_5$ was also observed in experiments with the real system. A detailed description of strategy $s_5$ now follows.

The strategy $s_5$ locates the grasping points by executing a search process on the input image. The search process traverses a distance $f_1$ from the center of mass along the main axis to the point $p_1$ (Fig. 3). The distance $f_1$ is randomly chosen, using a zero-mean Gaussian distribution. From $p_1$ the search process then moves along a perpendicular direction to the main axis, until the boundary is reached at point $p_L$, which is the first grasping point. The boundary orientation at $p_1$ is calculated, and the search process moves in the direction $f_2$ relative to the internal normal until the boundary is reached at point $p_2$, which is the second grasping point. The angle $f_2$ is randomly chosen, using a Gaussian distribution. The mean of the Gaussian is slightly shifted from the internal normal at $p_1$ toward the center of mass of the object.

B. Choosing Quality Parameters

The quality parameters are used to predict the quality of a candidate grasp configuration. We used statistical methods to choose a small subset of visual features that can reliably predict the grasp quality. The selection was performed using 2000 grasp configurations, generated by the strategy $s_5$. For

![Flowchart of the scheme for learning sensorimotor tasks.](image-url)
each grasp, a grade was assigned by the mechanical model and 20 visual features were extracted from the input image. The features were distances and angles related to the grasping points and the center of mass.

We estimated the statistical relations between the features and the grades using multivariate regression and the non-parametric method of conditional average prediction. The regression revealed that the proportion of the total data variance accounted for by the best model was relatively low ($R^2 = 0.82$), meaning that the grades were not approximated well by linear combinations of the first, second, and third powers of the visual features. However, small subsets of features gave high proportional variance relative to that of the full model. In particular, the three features $a_1, a_2, d$, shown in Fig. 1, yielded $R^2 = 0.748$, which amounts to 91% of the variance accounted for by the full model. The conditional average prediction method yielded similar results. The configurations were divided into bins and the expected grade for each bin was computed as the average grade of the configurations in it. The prediction quality in the method of conditional average is defined as $\text{Prediction} = 1 - \text{SSE}/\text{SST}$, where SSE is the variance of the grades in the bins and SST is the overall variance. The results show that a small subset of features can reliably predict the quality of the grasp. The best subset of three features $a_1, a_2, d$ yielded prediction quality of 0.97.

The visual features $a_1, a_2, d$ have direct influence on the grasp quality according to the mechanical model of grasping, which is described in Appendix A. The features $a_1, a_2$ estimate the angles between the grasping line, which is the line connecting the grasping points, and the normals at the grasping points. When these angles are below a certain threshold, the fingers do not slide. The feature $d$ is the distance from the center of mass to the grasping line, normalized by the length of the main axis. This feature estimates the moment arm. As $d$ decreases, the torque that rotates the object around the grasping line decreases, and the grasp becomes more stable. This interpretation, along with the statistical results presented above, supports the assumption that grasp quality can be predicted rather reliably based on a few visual features.

C. Choosing Action Representation

Since representation of candidate actions requires domain-specific knowledge, we use the information selected in the off-line process to design an efficient coding of a grasp configuration. This coding, which consists of quality parameters and location parameters, is used by the working system for storing successful configurations. The quality parameters are the visual features $a_1, a_2, d$. Note that the quality parameters are not specific to a particular object, and the same parameters may correspond to different objects (Fig. 1). The location parameters are the parameters of the search process which was defined for the heuristic strategy $s_5$. The parameters $f_1, f_2$ are sufficient to locate grasping points on an object, given its boundary, its center of mass, its main axis and the preferred direction along the main axis (Fig. 3). The location coding is not specific to a particular object, therefore stored configurations can be applied to unfamiliar target objects. Note that the curvature of the boundary near the grasping points $p_1, p_2$ must be low, otherwise the boundary orientations are not defined and the configuration is considered invalid.

IV. THE WORKING SYSTEM

The working system performs grasping trials while incorporating on-line learning into the two basic modules of the proposed scheme, generating grasping trials and predicting their quality. The primary sensory input to the system consists of a single top-view image of the target object, which is used to choose the grasping points. A second side-mounted camera is used for evaluating the grasp quality, as we explain in Section IV-B. The systems performs grasping trials using a parallel-jaw gripper with simple open/close control, mounted on a 4-DOF AdeptOne manipulator. When a new target object is presented, the system performs the following algorithm:

1. Get sensory information.
   Take a top-view image. Calculate boundary, center of mass and main axis of the target object.

2. For every stored grasp configuration $C$ do:
   $L \leftarrow$ Consider location parameters from $C$.
   Locate grasping points on image using $L$.
   $Q \leftarrow$ Extract quality parameters from image.
   $G_E \leftarrow \text{EstimateGrade}(Q)$.
   If $G_E \geq \text{SuccessfulGrasp}$ then goto step 5.

3. For $i = 1$ to MaxTrials do:
   $L \leftarrow$ Apply heuristic strategy.
   Locate grasping points on image using $L$.
   $Q \leftarrow$ Extract quality parameters from image.
   $G_E \leftarrow \text{EstimateGrade}(Q)$.
   If $G_E \geq \text{SuccessfulGrasp}$ then goto step 5.

4. Report that the object is not graspable. Stop.

5. Perform the grasp.
   $G_O \leftarrow \text{EvaluateGraspQuality}()$.
   Update Knowledge Base($L, Q, G_O$).

Generating candidate configurations is performed in steps 2 and 3 of the algorithm. In step 2, the location parameters of stored examples are applied to the new target object. Note
that the order of the stored examples, which is determined by $\text{UpdateKnowledgeBase}()$, affects the results. Since determining the preferred direction of the main axis from the input image is unreliable, the two possible directions are considered for every stored configuration. If a configuration $C$ is not applicable for a certain target object then $C$ is skipped. After all the stored examples were exhausted, the heuristic strategy is used to generate candidate configurations in step 3. Estimating grasp quality is performed by the function $\text{EstimateGrade}(\cdot)$, which receives the quality parameters and returns the estimated grade. After performing a grasping trial, the actual quality is evaluated by the function $\text{EvaluateGraspQuality}(\cdot)$. The knowledge-base is then updated for future use by the function $\text{UpdateKnowledgeBase}()$, which receives all the parameters of the current grasping trial. Next we present in detail these three functions. For more information about the working system, refer to [10].

A. Estimating Action Quality

We want to predict, based on previous experience, whether or not a candidate grasp configuration will be successful. Grasp quality is predicted by learning from examples the function $f_3: Q \mapsto G_F$ from the quality parameters $Q$ to the estimated grade $G_F$. The examples are the quality parameters of successful grasp configurations with their associated observed grades. As we need only a yes/no classification, basic learning techniques are appropriate for the task. We use numerical methods in which each set of quality parameters is considered as a point in a parameter space. Considering the quality parameters from a candidate grasp as a new point, its quality is predicted based on the quality of similar stored examples. Similarity between examples is defined using the Euclidean distance in the quality parameters space. We tested the following two methods for estimating the function $f_3$. In the NN method, the grade of a new point is set to the grade of its nearest neighbor in the parameter space. For our purpose, a candidate grasp is labeled successful when its Euclidean distance from any stored example is below a threshold. In the RBF method, the grade of a new point is calculated by interpolating the grades of several neighboring examples, using each example as a center of a basis function.

We use an additional global criterion for filtering candidate grasps. We define ranges of acceptable values of quality parameters and use them as an initial filter. Candidate grasps that have quality parameter values out of these ranges are ruled out. As the system starts, all possible values are acceptable. As the number of stored examples increases, the acceptable ranges shrink (Fig. 4), and at the limit they are restricted to the range of values in the stored examples. To calculate the acceptable ranges we add margins to the range of values in the examples. The margin size is proportional to the difference between the maximal (minimal) example value and in the examples. The margin size is proportional to the acceptable ranges we add margins to the range of values in the stored examples. To calculate ranges shrink (Fig. 4), and at the limit they are restricted to the range of values in the stored examples. To calculate the acceptable ranges we add margins to the range of values in the examples. The margin size is proportional to the difference between the maximal (minimal) example value and the maximal (minimal) possible value, using the following formula: Margin = $(\text{Max}_{\text{possible}} - \text{Max}_{\text{example}}) \times \exp^{-\delta N}$, where $N$ is the number of learned trials and the constant $\delta$ influences the learning speed. (In our implementation $\delta$ was set to 0.1.)

Ideally, once a successful configuration $C_O$ is stored for an object $O$, the system will be able to grasp the object $O$ when presented later, due to the following reasons. When the stored configuration $C_O$ is applied to $O$, the location parameters of $C_O$ will locate the proper grasping points. The quality parameters that will be extracted will match those stored in $C_O$ and therefore predict a good grasp quality, and the grasp will be performed. The configuration $C_O$ will not be applied to the object $O$ only if another configuration, for which the predicted quality is satisfactory, is found before $C_O$ during the loop in stage 2 of the algorithm.

As the number of stored examples increases, the probability that a successful combination of quality parameters will match a stored example increases, making it possible to generalize to novel configurations. On the other hand, the ranges filter becomes more restrictive as the number of examples increases, thus making it more likely to reject unsuccessful combinations of quality parameters. Therefore the system’s performance is expected to improve over time.

B. Evaluating Action Quality

The overall quality of a grasping trial is evaluated using a fixed side-mounted camera. After grasping the target object, the manipulator places the object in front of this camera, and the grasp quality is automatically calculated based on the minimal height of the object (as was suggested in [5]). If the object is held firmly, its height is above a certain value. If the object slides, its minimal height is lower, and so is the quality (Fig. 5). Note that more complex analysis of the object’s pose is not applicable here since the initial pose of the object is not known.

C. Updating the Knowledge-Base

After performing a grasping trial, the system updates its knowledge-base according to the current configuration $C = (L, Q)$ and the observed grade $G_O$. The current quality parameters are matched against all the stored examples, using the Euclidean distance $d_E$ in the quality parameters space.
We say that the current configuration $C$ matches an example $E$ when the distance $d_E$ is below a threshold. When a match occurs, we update the information associated with $E$. For each example $E$ we store the number of matched trials and the average grade over all the matched trials. A grasp is stored as a new example if $\text{quality} \neq \text{stored examples}$ and the grasp is not too similar to stored examples. In this case, similarity is measured as the Euclidean distance between configurations, considering both quality and location parameters. A stored example is removed from the knowledge-base when its average grade is lower than 90, or when it matches an unsuccessful grasp whose observed grade is lower than 70 (these values were set experimentally).

It is important to note that grasping trials with the same quality parameters may differ in their actual quality, because the quality parameters do not capture all the information that affects the grasp quality. The main reasons for the quality differences are described in Appendix B. Some grasp configurations are general and can be applied to many objects, whereas other grasps are specific to the objects for which they were used. For example, $a_1 = a_2 = d = 0$ and $f_1 = f_2 = 0$ are examples of general sets of quality and location parameters, respectively. The quality parameters refer to grasping an object on two parallel faces, where the grasping line passes through the center of mass. The location parameters determine that the grasping line is perpendicular to the main axis (Figs. 1 and 3). To differentiate between general and specific grasp configurations, we added a mechanism that controls the size of the matching neighborhood for the NN method. The current configuration $C$ matches a stored example $E$ only if its distance from $E$ is smaller than a threshold associated with $E', d_E$. When an unsuccessful grasp matches an example $E$, the matching neighborhood of $E$ is reduced by decreasing $d_E$. In this way specific configurations have more local effect in the parameter space than general configurations.

In order to use general examples before specific examples while generating candidate grasps, during step 2 of the algorithm, we sort the examples based on the following observation. General examples are expected to have a higher average grade, and to match more grasping trials. Therefore we sort the stored examples according to their average grade and their number of matched trials.

V. EXPERIMENTAL RESULTS

The scheme was used for learning visually guided grasping, in both real and simulated scenarios. First, we describe the experiments with the real robot, and then present the results of the simulated experiments.

A. Experiments with the Real Robot

A grasping trial with the robot consisted on the following steps: the user presented an object in the field of view of the top-view camera and initiated the trial (Fig. 6). The system...
chose two grasping points, using the ranges filter and the NN method for predicting grasp quality, and performed the grasp in the following way. The gripper was first located above the grasp position, moving high enough to avoid collision with the target object. It then moved down to a fixed grasping height and closed the fingers. The object was taken in front of the side-mounted camera, to evaluate the grasp quality. Finally, the object was returned to its initial position. Each grasping trial lasted about 1 min.

We performed two experiments with the real robot. In the first experiment, five metal generalized cones were used as target objects [Fig. 7(a)]. Each iteration of the experiment consisted of one grasping trial for every object, in the randomly chosen order $o3, o3, o4, o1, o2$. The average observed grade was measured for each iteration. The experiment consisted of three sessions, each starting with the heuristic strategy for choosing grasping points and no knowledge about quality prediction. The first two sessions included 20 iterations each. The third session included 25 iterations since the system required a larger number of trials to stabilize on a high level of performance. To compare the results to nonlearning performance, ten iterations were performed using the heuristic strategy alone. In this scenario every candidate grasp was performed. The average observed grade of these 50 grasping trials was considered a baseline for comparison. In the second experiment, we used 15 target objects which varied in weight, size, rigidity, and color (Fig. 7). The weights of the target objects ranged from 830 g to 30 g. At each iteration the system performed one trial for every object, and measured the average grade. One session which included 15 iterations was performed. To compare the results to nonlearning performance, five iterations were performed using the heuristic strategy alone.

The results show a significant improvement in the performance of the learning system relative to the results without learning (Fig. 8). In most cases the system stabilized very rapidly on a high level of performance. In session 3 of the first experiment, the system went through two phases of improvement/degradation before stabilizing. This behavior was due to the generation of “marginal” configurations by the heuristic strategy, which are defined as follows. Suppose that the quality parameters space is divided into two types of labeled regions, which contain either successful or unsuccessful examples. A marginal configuration is a successful example which is close to the boundary of a successful region. Consequently, configurations which are similar to a marginal example have a high probability to belong to unsuccessful regions and thus to cause failures. Session 3 demonstrated the system’s ability to delete unsuccessful examples from the knowledge-base. As the system gathered more experience, general grasp configurations became more dominant and specific configurations were used only for the objects that cannot be grasped using the general configurations. The system successfully tolerated inaccuracy of the measurements, in particular in the location of the center of mass and in the direction of the main axis. Weight differences between objects were also handled successfully, as grasp configurations that were learned for light objects but failed when applied to heavier objects were discarded from the knowledge-base (for more details see Appendix B).

Filtering candidates based on ranges of quality parameters proved to be very effective. The acceptable ranges that were learned in experiment 2 are presented in Table I, and they can be interpreted in the following way. Learning the ranges
for the angles $a_1, a_2$ corresponds to learning the value of the friction coefficient $\mu$ for the current gripper and the current set of target objects. Learning the range for the distance $d$ corresponds to learning the maximal moment arm $R$ that the gripper can resist.

### B. Experiments with the Simulated Robot

During the experiments we noticed that grasp configurations which were learned for certain objects were applied to other objects. After completing experiment 2 the system successfully grasped unfamiliar objects in several informal experiments. These results confirmed the ability to generalize to unfamiliar objects. No formal experiments were performed with the real robot due to technical problems, which made it impossible to use the experimental setup.

To test the generalization ability in a systematic way, we compared the predicted quality, using the knowledge learned in experiment 2, with the calculated quality, provided by the mechanical model. Three prediction methods were tested: the ranges filter and two methods for function estimation from examples (NN and RBF). Combinations of these methods were also tested. The knowledge from experiment 2 consisted of 45 grasp configurations, with grades ranging from 93 to 99. Two-thousand grasp configurations, chosen by the heuristic strategy $s_3$ on 20 synthetic objects, were used for the comparison. The quality parameters $a_1, a_2, d$ and the mechanical model grade $G_M$ were calculated for each configuration. A configuration was considered successful when $G_M \geq 80$.

The ranges filter, described in Section IV-A, ruled out grasp configurations which had quality parameter values out of the learned ranges. The performance of the ranges filter is summarized in Table II. A “false-alarm” mistake means that a successful grasp was ruled out by the test, while a “miss” mistake means that an unsuccessful grasp was not ruled out by the test. The ranges filter ruled out 60% from the configurations, and proved to be conservative. The overall “false-alarm” rate was 21%, and the “miss” rate was 4%. Most of the low grade configurations were ruled out by this test, making it a successful tool to reject unsuccessful grasp configurations.

Two function estimation methods, NN and RBF, assigned an estimated grade $G_E$ to each configuration. A configuration was expected to succeed when $G_E \geq 80$. The performance of the estimation methods is summarized in Tables III and IV. The NN method achieved a lower “miss” rate since configurations that were too far from all stored examples were considered as failures. The RBF method had a lower “false-alarm” rate for good configurations, $90 < G_M \leq 100$, since it uses interpolation between the stored examples.
To estimate the contribution of learning, we compared the success rate of the simulated grasping trials with and without learning. Without any previous knowledge the system would perform every candidate grasp. Using the ranges filter, only configurations that passed the filter would be performed. Adding the grasp quality prediction, using either the NN method or RBF, only the configurations that passed the ranges filter and their expected grade satisfied $G_F \geq 80$ would be performed. We also tested a combination of the ranges filter, NN and RBF methods. Only configurations which passed the ranges filter and matched at least one example in the NN method were checked by the RBF method. The comparison results are presented in Table V.

Surprisingly, the simple ranges filter gave a success rate similar to the more sophisticated learning methods, although estimating the grasp quality using NN and RBF methods ruled out more grasp configurations. These results suggest that the underlying structure of the quality parameters space is rather simple. The fact that the ranges filter is conservative implies that the training set of experiment 2, from which the system acquired its knowledge, is not large enough to enable generalization for all the random objects in the simulated experiment. Training on a larger set of examples, the ranges filter would become less restrictive and the more specialized prediction methods might have a more significant effect on the results. It is important to note that the mechanical model was not tuned for the specific characteristics of the real system, therefore an accurate fit between the model grade and the predicted grade was not expected. However, the strong correlation between the learned grasp quality and the model grade shows that the system successfully learned the important features which predict grasp quality.

Table V
Performance of Simulated Grasping Trials, Using Different Learning Methods. NN Stands for the Nearest Neighbor Method

<table>
<thead>
<tr>
<th>operation mode</th>
<th># trials</th>
<th>success, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>no learning</td>
<td>2000</td>
<td>56.9</td>
</tr>
<tr>
<td>ranges filter</td>
<td>788</td>
<td>89.8</td>
</tr>
<tr>
<td>ranges + NN</td>
<td>712</td>
<td>90.9</td>
</tr>
<tr>
<td>ranges + RBF</td>
<td>712</td>
<td>90.6</td>
</tr>
<tr>
<td>ranges + NN + RBF</td>
<td>674</td>
<td>90.7</td>
</tr>
</tbody>
</table>

In our view, the successful learning demonstrated by the system supports the use of task-dependent representations for learning sensorimotor tasks. When learning to predict grasp quality, for example, coding only the relevant features for grasp quality makes it possible to generalize among objects. The experimental results show that when the question of “what to learn” is addressed properly, simple techniques are sufficient for the actual learning.

Appendix A
Evaluating Grasp Quality: A Mechanical Model

We assigned grades to simulated grasping trials using a mechanical model of the grasping operation. Given an image that contains a cross section of a generalized cone and two grasping points on the boundary, we calculated the grade based on two considerations: no sliding and resistance to rotation.

Sliding of the fingers is not allowed. This constrain limits the angles between the line along which the fingers apply force, which we term the grasping line, and the normals at the grasping points (angles $\alpha_1, \alpha_2$ in Fig. 9). If these angles are above a threshold, $\mu$, which is related to the coefficient of static friction, the grade is set to zero. We set the value of $\mu$ to 30° in our calculations. If $(\alpha_1 > \mu)$ or $(\alpha_2 > \mu)$ then Grade = 0.

The resistance to rotation is the maximal torque the gripper can apply to the object. It depends on the shape and size of the contact areas, the pressure on them, and the friction and viscoelasticity of the object and the fingers. We consider soft-finger contacts, and assume that the contact areas are small, flat, and have a constant area. Consequently, the contact quality. Since the quality parameters are not object-specific, this formulation allows direct generalization among objects. As the proposed scheme requires some initial knowledge about the task, we developed a method for off-line selection of heuristic strategies and observable quality parameters, based on statistical analysis.

The scheme was used for learning visually guided grasping, in both real and simulated scenarios. We showed that grasp quality can be predicted rather reliably using a few visual features, extracted from a single image of the target object. We developed a special coding for grasp configuration which allows the system to generate candidate grasps by applying stored examples to unfamiliar target objects. Using this representation, and using the NN method for predicting grasp quality, the experimental system successfully learned to grasp a large variety of objects, with very different characteristics (geometry, weight, rigidity, color). The system showed an appreciable improvement of performance after a small number of trials, and maintained a high level of performance over sessions of several dozens of trials. We showed in simulations that the learned knowledge could be used for other objects. These results have demonstrated better performance compared to previous studies which have dealt with the same problem: Dunn and Segen [5] and Tan [22] used three target objects each, and Salganicoff [18] considered only cylinders and boxes. None of the previous studies presented an improvement of grasping performance over a continuous session of work.

In our view, the successful learning demonstrated by the system supports the use of task-dependent representations for learning sensorimotor tasks. When learning to predict grasp quality, for example, coding only the relevant features for grasp quality makes it possible to generalize among objects. The experimental results show that when the question of “what to learn” is addressed properly, simple techniques are sufficient for the actual learning.
characteristics are the same for all the grasps, and each contact can be modeled considering only the normal at one point of contact. We assume that the grasping force applied by each finger is constant, and its direction is along the grasping line. Thus the pressure of a finger on a contact area depends only on the normal component of the grasping force. The resistance at each contact area is given by

$$\tau_i = K \cdot \mu_s \cdot \text{Force} \cdot \cos a_i$$

where $K$ represents the geometrical characteristics of the contact area, $\mu_s$ is the coefficient of static friction, and Force is the magnitude of force applied by each finger. Note that the product $K \cdot \mu_s \cdot \text{Force}$ is constant over all grasps. The maximal torque the gripper can apply to the object is the sum of the maximal torques at the contact areas, Resistance $= \tau_1 + \tau_2$.

The torque that rotates the object around the grasping line depends on the mass distribution on both sides of the grasping line. We assume that the images are cross sections of generalized cones, with horizontal main axis and constant density. The volume of the generalized cone is calculated assuming a circular cross section in depth, perpendicular to the grasping line. The torque around the grasping line is given by the formula: Torque $= \int \gamma \cdot \rho \cdot dV \times g = R \times M \times g$. We integrate over the volume $V$, where $\rho$ is the density, and $\gamma$ is the moment arm of a mass element $\rho \cdot dV$. The gravitational acceleration, $g$, operates in the vertical direction. The equivalent moment arm, $R$, operates on the mass of the object, $M$ (see Fig. 9). To calculate the grade we consider the net torque, which is the difference between the resistance and the rotation torque. We use the following formulas:

$$\text{Diff} = \text{Best} - \min(\text{Best}, \text{Resistance} - \text{Torque})$$

$$\text{Grade} = 100 \times \exp \left(-\frac{\text{Diff}}{\gamma}\right)$$

where Best is the minimal net torque that gives the grade 100. The grade generated by this formula is 100 for Resistance $\geq$ Best, and goes down as Torque gets bigger, or Resistance gets smaller. The quality of the grasp decreases exponentially with the net torque (there is no special importance to the exponential relation, and $\gamma$ is a constant whose value was set experimentally).

**APPENDIX B**

**ASSUMPTIONS AND SOURCES OF UNCERTAINTY**

Given only a single image of the target object, certain assumptions must be made about its geometry and other nonvisual properties. We assume that the object height is within an acceptable range, and that low curvature of the silhouette corresponds to a low curvature (graspable) surface. The center of mass of the object is calculated as the center of the silhouette area. We do not make specific assumptions about the mass distribution of the object, but assume that the deviation between the real and the calculated centers of mass is small. We also assume that the center of mass falls within the silhouette, for the proper application of the heuristic strategy and the location coding.

In realistic settings, the assumption that grasp quality can be predicted based on a few visual parameters does not always hold. Consequently, similar points in the quality parameter space may correspond to grasping trials with different actual quality. Next we describe the main reasons for quality differences that we have encountered in our experiments. Weight differences between objects are a major cause of quality differences. Applying configurations which were learned on light objects to heavier objects may result in failure, since grasps of light objects are not sensitive to the distance from the center of mass. Since the knowledge-base is updated based on the observed quality of grasping trials, the system tends to discard such grasps. On the other hand, configurations which were learned on heavy objects are applicable to lighter objects. Mistakes in the calculated (silhouette based) center of mass are another major cause of quality differences. Some of the target objects in our experiments, especially the objects o2 and o3 (Fig. 7), had a significant difference between the calculated and the real centers of mass. The experiments showed that the system tolerated these mistakes by using specific configurations for problematic objects. Some grasp configurations rely on specific geometrical details of the grasped object. For example, the fingers may slip from an initial unstable position into a stable grasp (Fig. 10). Applying such a grasp to a different object may cause a failure. Last, some grasp configurations are marginal, and may cause different grades for very similar situations, as we explained in Section V-A. The quality difference may result from small differences in the image measurements or in the gripper location relative to the object, due to contact of one finger before the other, etc.

Variations in the silhouette shape, another source of uncertainty, can harm the matching between the current configuration and stored examples. Changes in the silhouette cause changes in the search process which locates the grasping points, and consequently cause changes in the quality parameters. As a result, applying a stored configuration, which was learned for object $O$, to the same object $O$, can result in a new configuration which does not match the original one. The shape
variations are caused either by noise and discretization of the image processing or by variations in the pose of the target object relative to the camera. (The target objects were placed by hand, without exact positioning and in random orientations.) In our experiments, these problems caused multiple storage of similar grasp configurations. However, in most cases the noisy values were similar enough to the stored ones to allow matching and proper prediction of the grasp quality.

REFERENCES


Ishay Kamon received the B.A. degree in computer science from the Technion, Technological Institute of Israel, Haifa, in 1989, and the M.Sc. degree in applied mathematics and computer science from the Weizmann Institute of Science, Rehovot, Israel, in 1993. He is currently pursuing the Ph.D. degree in computer science at the Technion. The subject of his thesis is incorporating local shortest path considerations in Bug-type motion planning.

His research interests include sensor based motion planning, robot navigation, and task-oriented vision.

Tamar Flash received the B.S. and M.S. degrees in physics from Tel-Aviv University, Tel-Aviv, Israel, in 1972 and 1976, respectively. She received the Ph.D. degree in medical physics in 1983 from the Massachusetts Institute of Technology (MIT), Cambridge.

From 1983 to 1985, she conducted postdoctoral research in the Department of Brain and Cognitive Science, Artificial Intelligence Laboratory, MIT. In 1985, she joined the Department of Applied Mathematics and Computer Science, Weizmann Institute of Science, Rehovot, Israel, where she is currently an Associate Professor. Her research interests include motor control, computational neuroscience, robotics, and artificial intelligence.

Shimon Edelman received the B.S. degree in electrical engineering from the Technion, Technological Institute of Israel, Haifa, in 1978, and the Ph.D. degree in computer science from the Weizmann Institute, Rehovot, Israel, in 1988.

He is now a Senior Researcher with the Weizmann Institute. He took leave from the Weizmann Institute to work as a Visiting Scientist at Brown University, Providence, RI, in 1991, at the Max Planck Institute for Biological Cybernetics, Tübingen, Germany, in 1995, and at the Massachusetts Institute of Technology, Cambridge, from 1988 to 1990 and 1996 to 1997. His main interests are cognitive science and neural computation. He has published nearly 40 journal papers, 20 refereed conference articles, and a dozen book chapters, on topics ranging from visual psychophysics to natural language processing. His current research concentrates on developing mathematical solutions to the problems concerning similarity in visual object categorization and in other cognitive functions, experimental assessment of these solutions as models of human performance, and bringing the theoretical understanding to bear on data from neurobiological studies of cognition in the brain.