



University of Pennsylvania
ScholarlyCommons

Technical Reports (CIS)

Department of Computer & Information Science

September 1991

Robotic Sensorimotor Learning in Continuous Domains

Marcos Salganicoff
University of Pennsylvania

Ruzena Bajcsy
University of Pennsylvania

Follow this and additional works at: https://repository.upenn.edu/cis_reports

Recommended Citation

Marcos Salganicoff and Ruzena Bajcsy, "Robotic Sensorimotor Learning in Continuous Domains", .
September 1991.

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-91-85.

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/cis_reports/426
For more information, please contact repository@pobox.upenn.edu.

Robotic Sensorimotor Learning in Continuous Domains

Abstract

We propose that some aspects of task based learning in robotics can be approached using nativist and constructivist views on human sensorimotor development as a metaphor. We use findings in developmental psychology, neurophysiology, and machine perception to guide a robotic learning system's level of representation both for actions and for percepts. Visually driven grasping is chosen as the experimental task since it has general applicability and it has been extensively researched from several perspectives. An implementation of a robotic system with a dexterous three fingered hand, compliant instrumented wrist, arm and vision is used to test these ideas. Several sensorimotor primitives (vision segmentation and manipulatory reflexes) are implemented in this system and may be thought of as the "innate" perceptual and motor abilities of the system.

Applying empirical learning techniques to real situations brings up some important issues such as observation sparsity in high dimensional spaces, arbitrary underlying functional forms of the reinforcement distribution and robustness to noise in exemplars. The well established technique of non-parametric projection pursuit regression (PPR) is used to accomplish reinforcement learning by searching for *generalization directions* determining projections of high dimensional data sets which capture task invariants. Additionally, the learning process generally implies failures along the way. Therefore, the mechanics of the untrained robotic system must be able to tolerate grave mistakes during learning and not damage itself. We address this by the use of an instrumented compliant robot wrist which controls impact forces.

Comments

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-91-85.

**Robotic Sensorimotor Learning In Continuous
Domains**

**MS-CIS-91-85
GRASP LAB 282**

**Marcos Salganicoff
Ruzena Bajcsy**

**Department of Computer and Information Science
School of Engineering and Applied Science
University of Pennsylvania
Philadelphia, PA 19104-6389**

October 1991

Robotic Sensorimotor Learning in Continuous Domains

Marcos Salganicoff and Ruzena Bajcsy*

GRASP Laboratory

Department of Computer and Information Science

University of Pennsylvania

email: sal/bajcsy@grip.cis.upenn.edu

September 18, 1991

Abstract

We propose that some aspects of task based learning in robotics can be approached using nativist and constructivist views on human sensorimotor development as a metaphor. We use findings in developmental psychology, neurophysiology, and machine perception to guide a robotic learning system's level of representation both for actions and for percepts. Visually driven grasping is chosen as the experimental task since it has general applicability and it has been extensively researched from several perspectives. An implementation of a robotic system with a dexterous three fingered hand, compliant instrumented wrist, arm and vision is used to test these ideas. Several sensorimotor primitives (vision segmentation and manipulatory reflexes) are implemented in this system and may be thought of as the "innate" perceptual and motor abilities of the system.

Applying empirical learning techniques to real situations brings up some important issues such as observation sparsity in high dimensional spaces, arbitrary underlying functional forms of the reinforcement distribution and robustness to noise in exemplars. The well established technique of non-parametric projection pursuit regression (PPR) is used to accomplish reinforcement learning by searching for *generalization directions* determining projections of high dimensional data sets which capture task invariants. Additionally, the learning process generally implies failures along the way. Therefore, the mechanics of the untrained robotic system must be able tolerate grave mistakes during learning and not damage itself. We address this by the use of an instrumented compliant robot wrist which controls impact forces.

*Acknowledgements: This research was supported in part by Navy Grant N0014-88-K-0630, AFOSR Grants 88-0244, AFOSR 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770; and Du Pont Company

1 Introduction

We propose that some aspects of task based learning in robotics can be approached using nativist and constructivist views on human sensorimotor development as a metaphor.

We use findings in developmental psychology and neurophysiology, as well as machine perception, to guide the overall design of robotic system which attempts to learn sensorimotor binding rules for simple actions. Visually driven grasping is chosen as the experimental task since: it is a generally applicable primitive action; allows for incremental gradation of complexity of the task; and has been extensively investigated by a number of research communities.

Before one can talk about learning, we have to put forward what the assumptions are about the system, i.e. what is “Innate.” We postulate the following as innate: A set of data reduction mechanisms for processing sensory data; Cost/benefit (utility) functions as part of the task model which control the internal resources of the system; motivation in terms of importance of the success/failure in accomplishing the task; and memory mechanisms that include indexing and matching with already stored experiential data and computation of the frequency and saliency of the stored information.

The learning is empirical in nature, and is done by having the robot observe itself in repeated interactions with the task environment. The resulting parameter binding rules then link the observed perceptual variables to appropriate operator action parameters during future executions of the task.

The approach taken here gains direct inspiration from Piaget, Bower, Von Hofsten, Diamond and Held [19, 3, 12, 5, 11] in the developmental psychology communities. Many researchers have looked at such ideas for robot learning, most notable are Drescher [6], who developed statistical approaches to learning in boolean valued simulated worlds, Mel [18], who takes a more neuromorphic approach using Sigma-Pi units, and Kuperstein [14] who looked at simulated hand/arm topographic maps and learning to coordinate them. Mason [17] has looked at action learning for manipulation planning via self-observation.

Several sensorimotor primitives (vision segmentation and manipulatory reflexes) are defined and implemented using this system and may be thought of as its “Innate” perceptual and motor abilities. In a visual scene, objects are represented parametrically by their position, orientation and gross shape parameters in a superquadric model. The execution of the motor activity is modelled by various parameterized actions such as approach to location, preshape hand, acquisition and lift. Collision retraction [3] and palmar traction grasping reflexes [21] are also used.

1.1 Progressive Refinement of Action and Perceptual Representations

We put forward and test the following working hypothesis: inductive learning must happen incrementally with respect to the number of parameters to be characterized, otherwise the learning becomes intractable due to the combinatorics of the task, given all possible parameter values.

As the system maturation proceeds (in our case as the tasks are getting more complex) the sensors deliver increasingly differentiated information (more perceptual parameters) about the world and the actuation. The action parameters of the system must correspondingly adjust. At the same time actions must progressively differentiate into more refined actions with more controlling parameters [3, 12, 20]. Each learning level guides exploration in subsequent learning levels, permitting the escape from the combinatorics of statistical learning with no prior information.

There are several tasks presented to the system, each with progressively increasing complexity. In executing these tasks, the system is learning to use more parameters as the task and perceptual complexity increase. This complexity with respect to parameters mimics the maturation process in biological systems. At first, the system has very insensitive perceptual capabilities and correspondingly, the task cannot be very demanding. Hence, its actions are very primitive, but almost always successful. An example of this zeroth order task could be: Make a tactile contact anywhere in the reachable workspace with any object or support surface. At this level, the system learns about the characteristics of its reachable workspace.

Let's consider a more complex task. There is a desired object in the workplace and the arm/hand must contact and move it, although it need not grasp and lift it. Mastery of this level is equivalent to a biological system that has learned to discriminate the object in the foreground from the background. The hand/arm system has learned the constraint that the hand position and object position must match roughly in order to change the state of the object.

The next level of difficulty in the progression of tasks is one in which the system must grasp the object, although not necessarily lift it. This is similar to the previous task, except that the hand/object matching constraint is much tighter since the hand must enclose the object. The information from the previous task is used to guide the exploration in this level so that each grasping trials has a higher probability of success. Thus, the system does not waste time attempting grasps far away from the location of the object which are information poor with respect to the current task.

The next task is to grasp and lift the seen object. Again, the success constraint is progressively tighter and we bootstrap our exploration using the previous tasks.

We model this empirical learning process as a multivariate statistical regression. Projection Pursuit Regression [7] (PPR) developed specifically for use in high dimensional spaces ($d \geq 3$) is used to approximate the distribution of reinforcement (success) in this parameter space. This technique also allows salient variables for the successful outcome to be identified if the space contains information poor parameters. Such techniques must be used in order to work with the small sample sizes required in learning, since there is a cost associated with completing each trial. Finally a method is proposed to index this domain information, allowing the reinforcement distribution to be efficiently accessed for decision-making during future planning and real-time execution of the actions.

What we seek is to characterize the distribution of reinforcement in the attribute space. We view this distribution as a *prediction surface*. Having such a predictive mechanism yields several benefits. It is a tool for guiding task execution and subsequent learning because it provides a means of compactly characterizing the peaks in reinforcement in the space of relevant sensorimotor attributes. We can partition this space into volumes which have a high predicted reinforcement. The iso-reinforcement surfaces of the volumes become decision hypersurfaces whose projections onto the action parameter axes can provide feasible, as well as preferable, intervals for parameterizing a given perception/action pair. These intervals then have a high likelihood of success. In essence, this allows us to identify relevant constraints for goal success in a given state.

1.2 Data Reduction Mechanisms and Parameterized Environments

Since our paradigm is for parameterized worlds, it is important to define what is meant by this term. A parameterized world is one whose configuration can be reduced to some set of real-valued description vectors for the configuration of objects and relationships between them. This world also includes a set of stereotypical primitive parameterized actions whose execution behavior is a function of some finite set of parameters that describe them. This fits quite nicely with current advances in computer vision as well macro-operators in planning and schema based descriptions of actions in the motor control literature.

As an example in the perceptual domain, consider the superquadric part representation as developed by [10]. Superquadrics are a generalized form of parametric surfaces which can represent a wide variety of shapes. A 3D superquadric shape in the scene is completely defined by the parameter set $(x, y, z, \theta, \phi, \gamma, a_1, a_2, a_3, \epsilon_1, \epsilon_2)$ which defines its position, shape, and orientation.

On the other hand, we might like to sense the position and orientation of the robot wrist, which would be represented as $Q_{wrist} = (Q_1, Q_2, Q_2, Q_3, Q_4, Q_5) =$

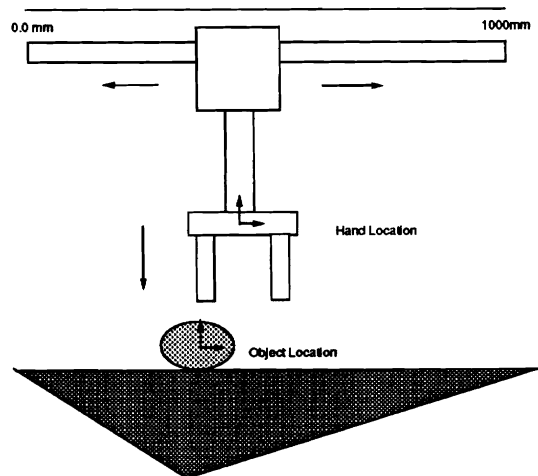


Figure 1: A schematic of a hypothetical simplified task setup for learning hand position selection based on object position for the purposes of this exposition. O_x refers to the position of the object along the x-axis. H_x refers to the position of the hand along the x-axis.

(x, y, z, r, p, y) using the roll, pitch, yaw transformation description.

1.3 Statistical Learning as a Form of Induction

It is useful to think of inductive learning as process of searching for regularities and structure in data sets. It is a data reduction mechanism applied to stored experiential information. The discovery of such regularities corresponds to the induction of a generalized rules about the data set. Normally, the data set consists of a preclassified set of instance descriptions and class assignments that are typed in by a human expert. An autonomous system does not have this luxury. It must be capable of data reduction from a real-valued domain to the appropriate level of granularity which permits the system to function effectively, yet not be over-represented.

Most of the structure to be found in perceptual data consists of correlations between perceptual inputs and action parameters. Once this relationship is found, its degree of reliability must also be categorized if the knowledge extracted from the learning is to be operationalized. The degree of reliability is also estimatable from the variability of the reinforcement measurements in the attribute space about the conditional expected value.

As an example, consider the simple task illustrated in fig. 1, which is the simplest pick and place. We define some simple sequencing order and parameterized actions to accomplish this task. We take some number of measurements of the reinforcement for different parameterizations of actions in the attribute space (see fig. 2(a)) and

attempt to form a least-squares response surface as in fig. 2(b)) which is then used as an estimation function for predicted reinforcement given new combinations of the sensorimotor attribute valuations. The form of this function is a non-parametric least-squared fit of the data or possibly, some other non-parametric means of characterizing modes and widths of the distribution. In either case, smoothing such a distribution allows a generalization to novel instantiations over a given range by the properties of interpolation afforded by the regression fitting process.

The relationships discovered between independent perceptual parameters and controlled action parameters can be expressed in terms of a functional motor mapping approximation $Q_i = \hat{M}(P_{reduced})$ where Q_i is a dependent actuator value and $P_{reduced}$ is the reduced perceptual state vector for that function, which is some subset of the attribute variables. The maximum reinforcement regions of the attribute space form constraints which can be used to generate action parameter binding relations which describe feasible as well as locally maximal estimated reinforcement values.

In order to represent the regions of high reinforcement in an efficient manner, a 2^n -tree representation of hyper-rectangular volumes in the n -dimensional parameter space is used (fig. 3 (a)). This allows arbitrarily shaped regions to be represented as unions of hyper-rectangular volumes of varying size which are accessible using time efficient tree structure to store them. These regions are then merged as in fig. 3(b). Once we have an 2^n -tree representation of the desirable regions, the question is how do we utilize and index this information in a useful and efficient manner? Since we have the information stored in a tree representation we can perform an associative search based on the attributes that are currently being observed. The result of such a process is illustrated schematically in fig. 4 where a given observation indexes through to associated volumes in the parameter space and finds the orthographic projection of that volume onto the motor attribute axis. We also reinterpolate along this volumes so that the expected reinforcement and associated variances inside of the volumes are also easily available so as to permit the computation of the local maxima within that volume and the variances about the expected values.

2 Non-Parametric Regression

2.1 Function Learning

Using regression allows one to build up a predictive mechanism for future success as a function of what the robot is observing and the action parameterization. This amounts to learning the expectation of reinforcement value conditioned on the valu-

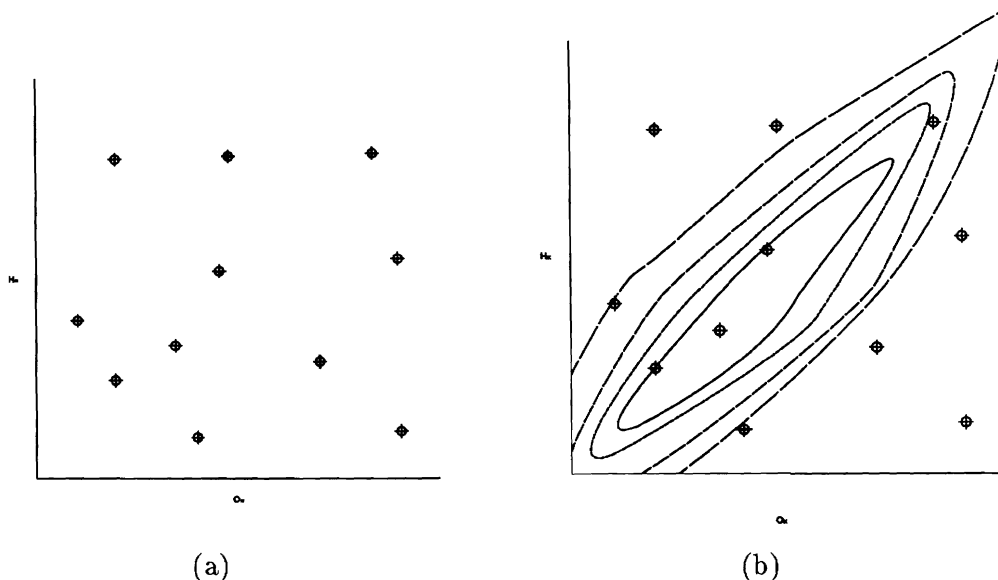


Figure 2: (a) The raw scatter plot of reinforcement values obtained by executing an action with different perceptual states of affairs. (O_x is the horizontal and H_x is the vertical). Each point in the plot has a reinforcement mass associated with it which is the result of the degree of success of the execution an action given some observed perceptual parameters. (b) The non-parametric regression fit to measured data which attempts to smooth and predict the reinforcement mass density over the entire domain of input parameters.

ations of the perceptual attributes from a series of noisy and sparsely spaced observations. This problem can be solved using multivariate statistical regression techniques. If a smooth function well-approximates the underlying distribution, then we can extrapolate and interpolate this expectation function to novel sensorimotor instances. In other words, we have a system which is able to generalize with respect to the action parameters.

The idea of learning a function by at a set of input/output pairs is not a new one in robotics. A common approach has been to use look-up tables with interpolation between measured points. Indeed, non-parametric statistical regression on a set of measurements may be considered to subsume these techniques. An early example of table lookup is Albus' CMAC polynomial hashing interpolation [1]. More recently, Atkeson et.al.[2] have explored task level robotic learning using polynomial interpolation as well as non-parametric locally weighted regression with some success. Mel [18] has used a connectionist approach approximate functions of several variables. These approaches are interesting, but in general, suffer from high sample size requirements as the dimensionality of the input space increases.

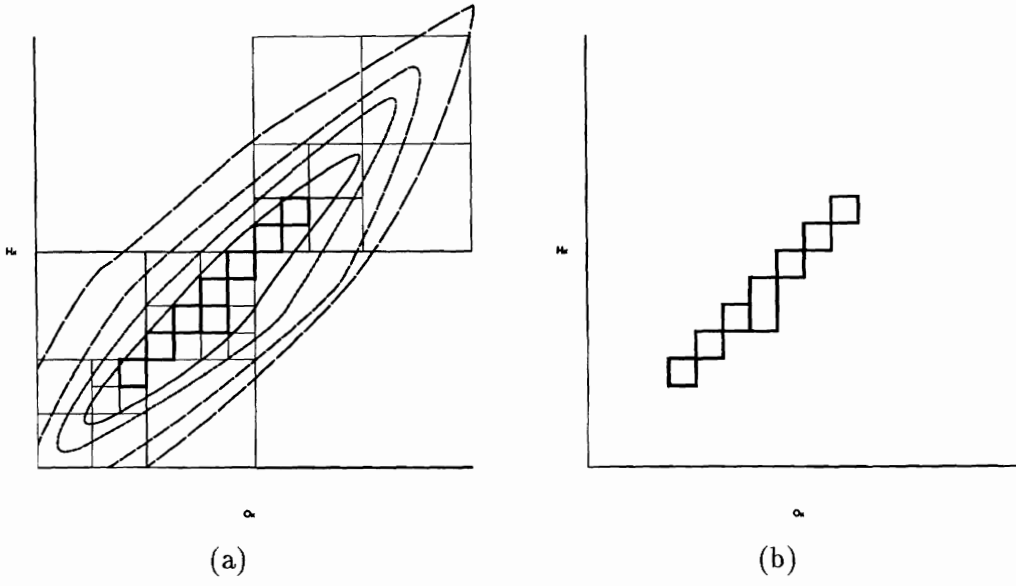


Figure 3: (a) The quadtree (2^2 -tree) reduction of the regression surface
 (b) The resulting constraint regions which represent the reduced empirically derived domain knowledge for the given subtask.

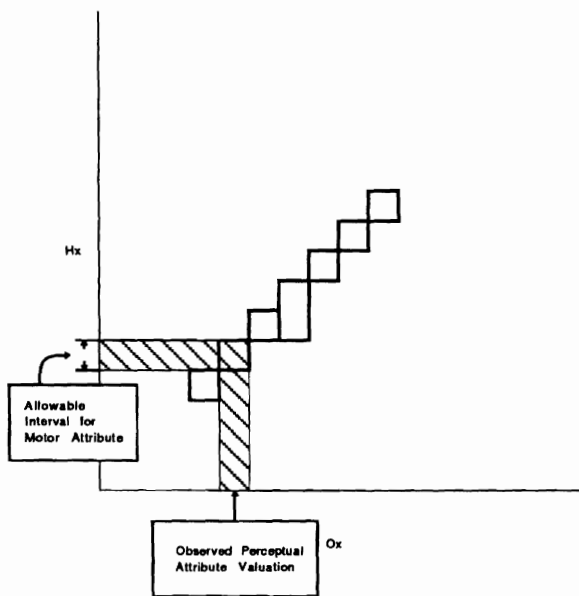


Figure 4: Execution-time indexing and retrieval of domain knowledge after learning. The perceptual parameter, along the O_x axis, in this case the location of the object, is mapped via an associative lookup to an allowable interval in the H_x axis which determines where the robotic hand is placed

Many of the interpolation schemes are not designed to be robust with respect to noise in the training samples and can be unduly influenced by this. Non-parametric regression locally weighted techniques [4] as advocated by Atkeson [2] remedy the noise immunity problem to some extent.

2.2 The “Curse of Dimensionality”

All of the above approaches suffer from the “curse of dimensionality.” The “curse” can be defined as the need for exponentially larger sample sizes as the dimensionality of the input space increases. A common illustration is as follows [13]. Consider a locally weighted regression or interpolation scheme which relies on 10% of the total samples for making an estimate of a given query point. Assume we are interested in the function over the domain of a unit 9-dimensional hypercube. If we assume uniform distribution of exemplars over this cube, then we must have 10% of the volume of this 9-d cube, i.e. $(f_i)^9 = .1$, where f_i is the fraction of the unit distance along each axis. Then $f_i = (.1)^{1/9} \approx .77$ which is a huge portion of domain. If we attempt to narrow f_i , then the fraction of volume necessary for the local fit rapidly decreases, and in order for it to contain sufficient number of points for a reasonable estimate requires a huge number of samples. This is the reason that most table lookup approaches have been applied primarily to lower dimensional functions. This problem has been addressed by the statistical community in a number of interesting ways, the approach we will select here is that of projection pursuit regression (PPR) as devised by Friedman et.al. [9, 7].

2.3 Projection Pursuit Non-Parametric Regression (PPR) Methods

We summarize the Smooth Multiple Additive Regression Technique (SMART) of Friedman [7]. Assume we have some underlying function $f(x_1, \dots, x_p)$ which we wish to approximate from some set of noisy observations $\{(x_{11}, \dots, x_{p1}, y_1), \dots, (x_{1n}, \dots, x_{pn}, y_n)\}$, (in our case y is either a success or a failure, although it could be a continuous reinforcement value) where there are n observations. Each observation may also be scaled by some weighting factor w_i . Assume the observations come from the following process:

$$Y = f(X_1, \dots, X_p) + \nu \quad (1)$$

where ν is a random variable with $E[\nu] = 0$ and $E[f] = 0$. In regression we endeavor to estimate the conditional expectation

$$\hat{Y}(x_1, \dots, x_p) = E[Y | X_1 = x_1, \dots, X_p = x_p] \quad (2)$$

The SMART method searches for an expansion of the form

$$\hat{Y} = \hat{f}(X_1, \dots, X_p) = \hat{f}(\vec{X}) = \bar{Y} + \sum_{i=1}^M \beta_i g_i(\alpha_i^T \vec{X}) \quad (3)$$

where $g_i(z)$ is a smooth “ridge” function of scalar z . Here α is the unit direction vector which projects the various covariates and β is a scalar weighting coefficient. The approach is therefore, to simultaneously find some “good” projection directions of the data and smooth functions $g_i(z)$ which are the smoothed versions of the set of values $\{(z_1, y_1), \dots, (z_n, y_n)\}$, where $z_i = \alpha^T \cdot [x_{1i}, \dots, x_{pi}]$. By a good choice of direction vectors, weighting coefficients and smooth functions, we mean those that minimize the unexplained variance of the case responses along those projections and mapped through the smooth functions.

Since the $g_i(\cdot)$'s are the smoothed versions of all of the cases projected onto one dimension, achieving a large enough sample size is much less of a problem than methods which form estimates over the raw high dimensional neighborhoods.

The search for the parameter set minimizing those values is done using standard Gauss-Newton minimization techniques and by grouping the parameters, holding some fixed, and minimizing the others in turn, so that the residual error is always decreasing.

The SMART technique can also be used to solve classification problems [7], that is to come up with an assignment rules conditioned on (X_1, \dots, X_p) that minimizes the classification risk for a categorical response variable. That is, a variable which takes on only one of a set of discrete and unordered values. This is useful when only binary reinforcement (a thresholded success or failure) of the task goal is available. The risk of misclassification is defined in general as

$$R = E[\min_{1 \leq j \leq q} \sum_{i=1}^q l_{ij} p(i | X_1, \dots, X_p)] \quad (4)$$

l_{ij} is the loss for predicting $Y = c_j$ when in actuality its value is c_i , $p(i | X_1, \dots, X_p)$ is the conditional probability that $Y = c_i$ given some valuation for the predictor variables. The l_{ij} allows the incorporation of the cost/benefit notion in the classification. $p(i | X_1, \dots, X_p)$ is the conditional probability. The conditional probabilities are then estimated and j^* which minimizes the R is chosen as the class for a given future observation.

2.4 Using Projection Pursuit for Attention Learning

Attention learning involves identifying salient variables for a goal, or learning what to attend to during given points of the execution of the task, this type of learning has been investigated by [6, 16]. A salient variable for a goal is one that has influence

on the outcome of the task as measured by the reinforcement function. In general the number of salient attributes for a given goal is much less than the total number of available perceptual attributes. Therefore, if we can have a system which learns to focus on only the attributes which are currently relevant, then we may more efficiently allocate sensing resources.

We use the relative importance of variable measure [7] to select relevant variables. This is defined as the product of the variance of a predictor variable times the magnitude of expected sensitivity of the component in the ridge functions to it. This is expressed as

$$I_j = \sigma_j E[|\partial \hat{Y} / \partial X_j|] \quad (5)$$

Therefore, PPR may present some advantages with respect to interpolation schemes such as CMAC etc., as well as the robustness of non-parametric regression techniques without the problem of poor sample economy in higher dimensions, and may identify salient input attributes. Now, let us present an experimental example of the use of the technique.

3 Experimental Protocol and Setup

The experiment consisted of learning to position the gripper in the plane given a visual observation of the centroid of the object in the plane from a top visual view (see fig. 5).

3.1 Experimental Protocol

A workspace was defined in which the object to be grasped may be placed at random. The object consists of a 1/2 lb. aluminum soda can (12 cm tall, 6 cm diameter) covered with white paper to simplify vision processing. The workspace consists of square 80cm by 40cm area. A pair of numbers in the workspace interval is generated by a random number generator. The experimenter (human) manually positions the target object at that position. The robot arm is retracted from the workspace and the CCD camera vision system acquires a top-view digitized image of the scene. The vision software then thresholds the scene based on sample mean, grows 8-connected regions and culls the regions by a minimum area criterion to eliminate artifact and noise induced smaller regions. The surviving region (there should be one region since there is only one target object in the field of view) is then used to compute a superquadric fit and the centroid of this superquadric is stored as the position of the object. Since the task is a two dimensional one, a monocular camera view is used to determine the location of the object in the plane.

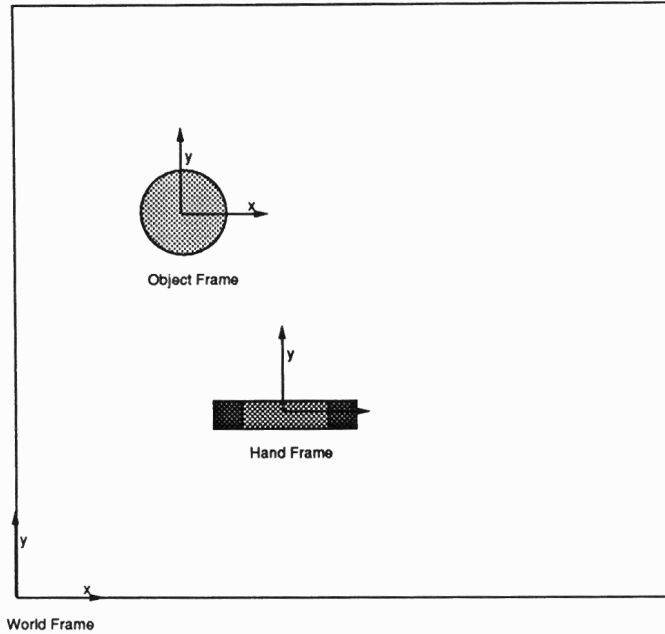


Figure 5: Schematic of the planar object retrieval task from the ceiling mounted camera view. Here, the object frame (O_x, O_y) , the hand frame (H_x, H_y) , and the contact/no contact are stored for each trial grasp.

The grasping trial set consists of the following actions. The arm is retracted upwards and laterally out of the workspace to prevent visual occlusion. Another visual sample is taken and logged since the object may have moved due to interaction with the hand. The system then computes a bounding box of 40cm around the location where object contact took place and chooses random location in that box from a uniform distribution. The robot then moves to that location and begins a downward motion. This motion is terminated by one of two conditions. Either a wrist displacement is sensed at which point the grasp reflex is initiated (and that trial labelled a success) or a positional stop at 8cm above the table is reached. A wrist tactile event is logged as a tactile interaction with the object. If the positional stop is reached, then the arm missed the target and the given grasping trial is labelled a total failure and the arm again retracts and this step begins at another random point in the bounding box. Otherwise the grasp reflex occurs in the next step. If more than some maximum number of grasping trial failures occur consecutively, another image of the scene is taken and the grasping trial set begins again.

The grasp reflex (see fig. 6) consists of the instantaneous closure of the finger until either a object interaction detected on each of the fingers or a desired position is reached with no contact. If the positional stop occurs then the grasp attempt is termed a failure since the finger reached its maximum position without encountering

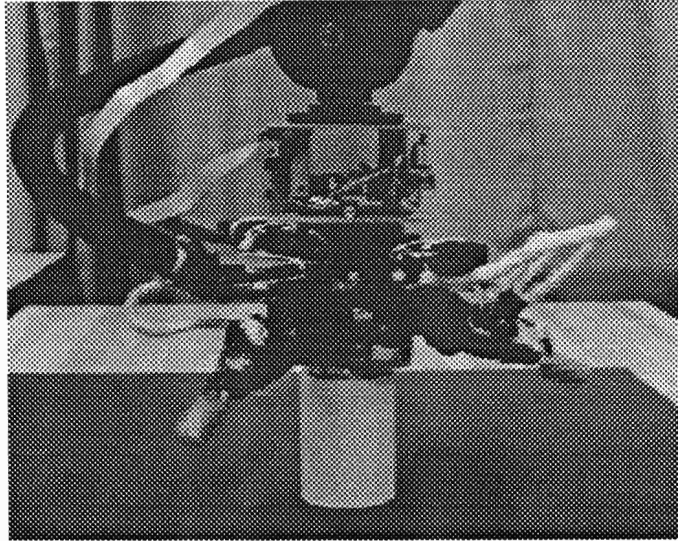


Figure 6: The initiation of the grasp reflex. A relative cartesian deflection or equivalent angle axis rotation of magnitude above the given threshold value is detected by the wrist and causes the arm motion to cease. The hand then immediately closes, analogous to the triggering of a spinal-level reflex.

the object. The outcome of this event logged and a grasp trial set step begins again. If 2 of the 3 fingers are still contacting the object, the finger states are logged, and the arm is retracted upwards for 10cm and the contact information saved. If the object was not enclosed at the end of the trial, another image is acquired and processed and another grasp trial set begins.

3.2 Hardware Setup

The experimental System consisted of a PUMA 560, instrumented compliant wrist and Penn Hand controlled and coordinated using a common MicroVAXII with shared memory (see fig.7). The Penn Hand [22] is controlled using a serial link to a high high-level controller which interprets commands and servos the hand configuration to desired forces or positions. The CCD camera output is digitized on the MicroVAXII and processed for a superquadric fit using on a SUN4/260 via an ethernet connection. The only real-time sensitive component of the system is the connection between the wrist and arm which occurs via the shared memory connection within the Microvax.

3.3 The Instrumented Compliant Wrist

The compliant wrist serves two important functions. It controls contact forces on the hand and also serves to detect contact with the object during data gathering. Since

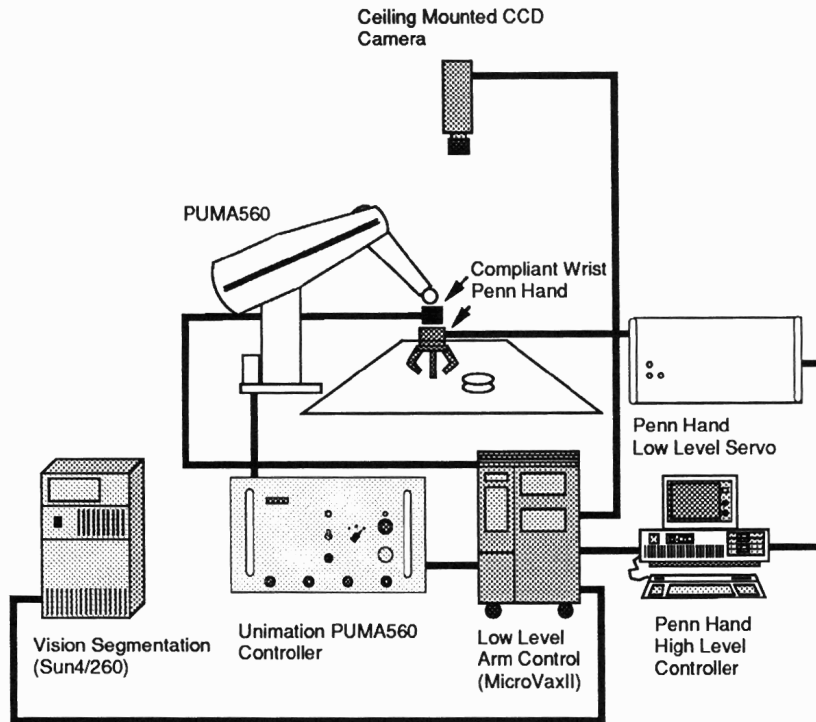


Figure 7: The Experimental System. The PUMA560, Hand and Wrist are controlled and coordinated using a MicroVAXII with shared memory. The MicroVAXII sends commands via a serial link to a high-level controller which interprets commands and servos the hand configuration to a desired force or position. The CCD camera output is digitized on the MicroVAXII and processed on the SUN4/260 via an ethernet connection. The only real-time sensitive component of the system is the connection between the wrist and arm which occurs via the shared memory connection.

the Penn Hand is a somewhat delicate mechanism and the PUMA is capable of large forces, we must take care to control the forces exerted on it (especially its fingers) during impacts in the data gathering phase. This is accomplished by mounting an instrumented passively compliant wrist [15] behind the hand. The wrist has intrinsic low stiffness, which dissipates impact energy on contact, thus protecting the fingers from excessive forces. The wrist also serves to detect collisions with the object in uninstrumented areas of the hand. Since all exposed areas of the hand cannot be sensorized, a contact in a non-sensorized area will still lead to a wrist displacement.

As can be seen in fig. 6, the wiring requirements for the hand's actuators and sensors are significant. This leads to arm configuration dependent forces being exerted on the wrist due to the cables. These undesirable artifact forces would be sensed by the wrist and might lead to grasping reflexes being inappropriately triggered. In order to compensate for this, two low-pass filters are used on the wrist output. The first has a very low band pass which tracks the wrist positional baseline at roughly the frequency of gross arm motions. The output of the filter is subtracted from the current cartesian readings and the resultant signal is again low pass filtered, although this time with a higher pass filter. This second filter blocks arm vibrations but permits contact events to be passed. The relative displacement thresholds and filter cutoffs were empirically determined and proved quite reliable throughout the experiments. Additionally, excessive arm accelerations could lead to artifact forces and moments being generated at the wrist. Therefore, all arm accelerations were carefully limited and reflexes were only enabled during the terminal phase of the hand approach.

4 Results

Figure 8 is a histogram for 303 grasping trials which were actually performed in the workspace. This figure illustrates a rough outline of the shape of the hand, since a collision causes a wrist displacement, no matter where it occurs on the hand. One can view this figure as the resulting image of the hand as yielded by the object being used as a probe to trace out the presence or lack of the hand. Notice that the width of the "fingers" of the histogram is approximately 6cm (which is the diameter of the can). Therefore the histogram also encodes information about the target object as well as the hand itself.

This gathered data from experimental trials was used to generate an augmented data set which consisted of simulating the process of positioning the object at 20 uniformly distributed random points in a rectangular workspace of $\pm 1\text{m}$ around the base of the robot. At each of the different locations, 100 points from the experimental

data were rotated by a random θ in $[0, 2\pi]$ and translated to the current simulated object location. This process yielded the 2000 simulated trial points shown in figure 9. Each instance is recorded as O_x, O_y, H_x, H_y, G , where (O_x, O_y) is the perceived object location, (H_x, H_y) the hand position and G is either a 0 for no contact or a 1 for contact. This corresponds to randomly oriented the hand and moving it to a random point in a 40 cm by 40 cm interval around the object, and moving the hand downwards to see if it contacts the object. The larger points in fig. 9 indicate successes and the smaller points indicate failures.

The projection pursuit algorithm classification was attempted on this data (SMART Routines Version 10/10/84 [8]) and yielded the results depicted in fig. 10. After training, the classification function was able to correctly predict, given the perceived location of the object in the plane, whether placing the hand in a given location would yield a tactile percept in a region $\pm 1m$ of the base of the robot. It was also able to generalize to regions of the workspace where empirical information was not taken as is illustrated by fig. 10 which shows the correct classification given that the object is in location $(-.6m, 0m)$ (see fig. 10 (a)) although this position was not in the learning set.

5 Discussion

The result shown above illustrates the usefulness of the approach, and also brings up several important issues. First, the PPR method exhibits an inductive bias which searches for invariances of arbitrary distributions under affine transformations. However, since the fitting of the projected data uses a non-parametric smoother, it does not exhibit bias for preferring certain distributions over others (i.e. multimodal versus unimodal). This is advantageous in the case that the task being learned can be satisfied by several action valuations. Consider the task of mating a smooth part into a cylinder with both ends uncapped. The distribution of success relative to the cylinder coordinate frame would have two peaks $\pm\pi$ which would be characterized by the fit with enough samples. The corresponding tree representation of fig. 3 (b) would then capture the feasible bi-modal distribution of valuations for a given cylinder position and could subsequently provide this domain information to a higher level spatial planner which could incorporate it into its plan building. This is in contrast to a connectionist type learning system which would not encode possible alternatives explicitly if they existed.

It is clear that there are several important tradeoffs in the application of this technique. The first tradeoff is between the width of the distribution of successes relative to the total size of the workspace in which the task is take place. If the

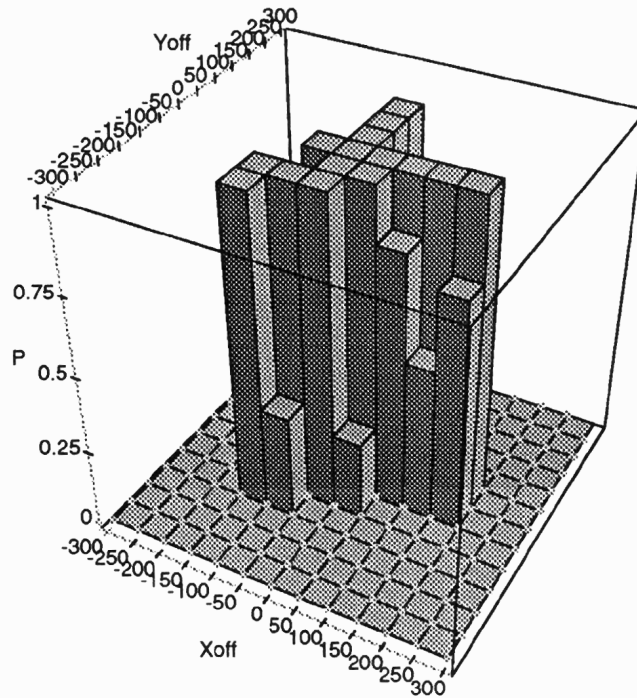


Figure 8: Histogram of centered data from 303 trial grasps in the workspace. It depicts the raw data centered at the perceived centroid of the target object and is the proportion of success to failure for summed over 50mm regions of the workspace. Notice that since the hand orientation was fixed for all trials, the histogram outlines the shape of the hand in the spherical grasp configuration

width is too large relative to the workspace or the sampled locations are too close together then the finding a projection direction vector which organizes the data and minimizes its variance is ill-conditioned since the ensemble variance varies little as a function of the direction chosen. This was evidenced by the fact the the results for the fit on positions in the original workspace (40cm by 80cm) were poor given that the diameter of the distribution relative to the object center is approximately 40cm due to the physical hand width span. By augmenting this data set using empirical data as a base and increasing the domain size to $\pm 1\text{m}$ of the robot the successful result shown here was obtained. At the other extreme, if the width of the distribution is too small relative to the workspace, then the sample economy of the learning process is very small, i.e. many trials have to be attempted for a success to be logged and therefore the learning process is very slow.

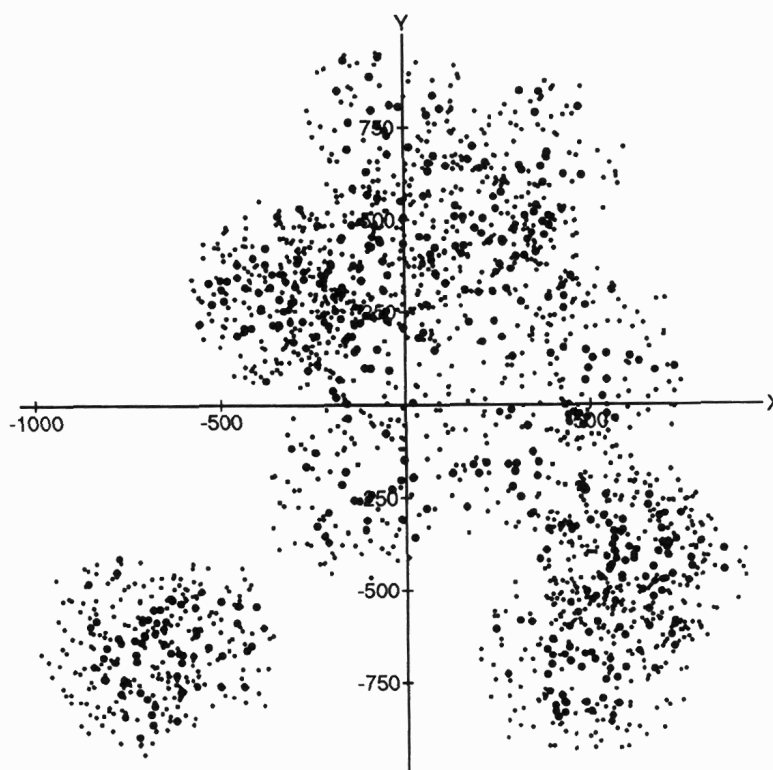


Figure 9: The raw data (generated simulation using the empirically obtained data distribution).

6 Conclusion and Future Extensions

This use of projection type regression techniques has shown promise in reducing the sample sizes necessary for generalization in continuous domains. Immediate extensions include using the prior information of this level to guide exploration in subsequent levels and attempting to learn the full grasping task, as well as learning to select from among different stereotypical grasps based on object shape. Other issues to be investigated include adaptivity in terms of forgetting rules such as weighting each observation by an appropriate discount factor based on its recency. Also, other interactive schedules for varying the locations of data gathering based on ambiguities in the current fit would serve make the method more on-line in nature and should be pursued.

Acknowledgements

We would particularly like to thank Dr. Max Mintz for contributing to many of the ideas in this paper; Dr. Jerome H. Friedman for graciously providing the the

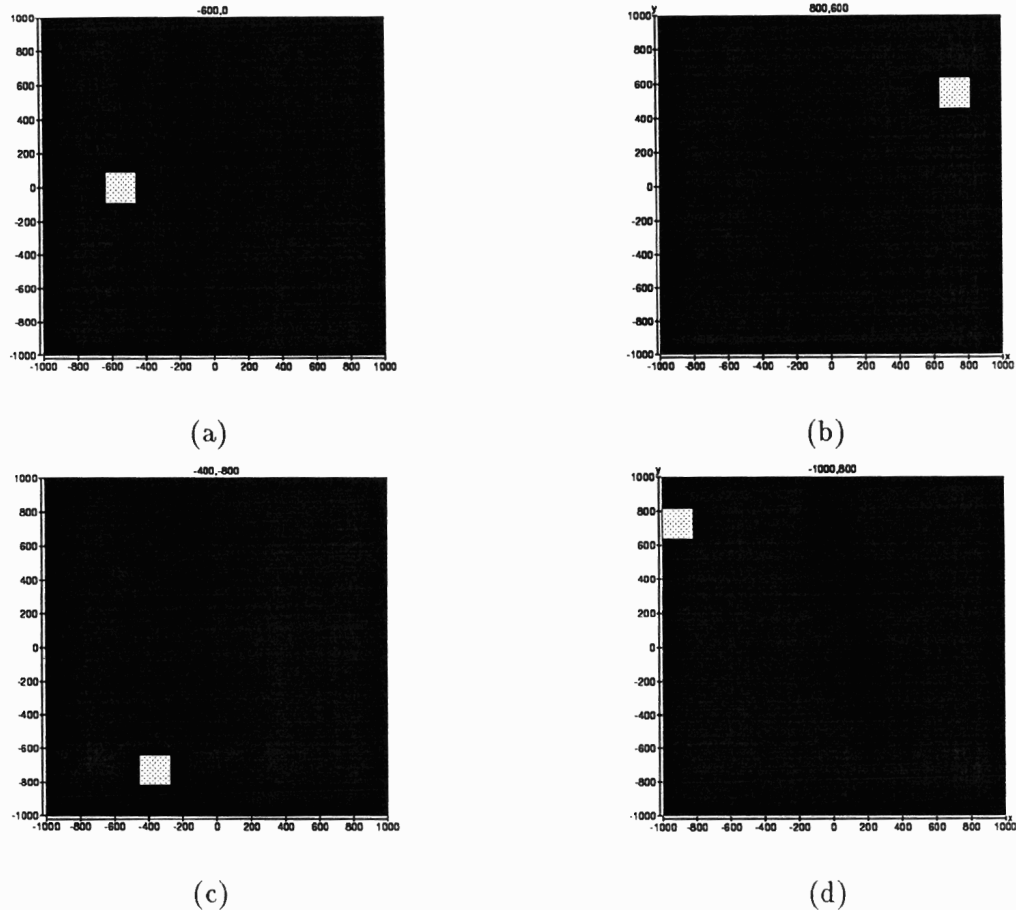


Figure 10: Resulting classification for objects positioned at (a) (-6.m,0m), (b) (.8m,.8m), (c) (.4m,-.8m) (d) (-1m,-.8m) relative to the robot base. Each density plot represents the resulting classification where white represents an expected tactile stimulus and black is a miss. From this, it can be seen that a correct decision rule for placing the hand has been induced.

source codes for his PPR algorithms; Tom Lindsay for making the wrist available; and to Luca Bogoni for many insightful discussions. Phil Fuma and John Bradley contributed much to the success of the implementation.

References

- [1] J.S. Albus. A new approach to manipulator control: the cerebellar model articulation controller (cmac). *Transactions of the ASME: Journal of Dynamic Systems, Measurement, and Control*, 97:220–227, September 1972.
- [2] C.G. Atkeson. Using locally weighted regression for robot learning. In *Proceedings, 1991 IEEE International Conference on Robotics and Automation*, pages 958–963, IEEE Press, 1991.

- [3] T.G.R. Bower. *Development in Infancy*. W.H. Freeman and Company, New York, 1982.
- [4] W.S. Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74:829–835, 1979.
- [5] A. Diamond. Developmental time course in human infants and infant monkeys, and the neural bases of inhibitory control in reaching. *Annals of the New York Academy of Sciences*, 608, 1990.
- [6] G.L. Drescher. *Genetic Ai – Translating Piaget into Lisp*. A.I. Memo 890, MIT AI Laboratory, Cambridge, Massachusetts, February 1986.
- [7] J.H. Friedman. *Classification and Multiple Regression Through Projection Pursuit*. Technical Report LCS 12, Stanford University, Laboratory for Computational Statistics, Stanford, CA, 1985.
- [8] J.H. Friedman. *SMART User’s Guide*. Technical Report LCS 1, Stanford University, Laboratory for Computational Statistics, Stanford, CA, 1984.
- [9] J.H. Friedman and W. Stuetzle. Projection pursuit regression. *Journal of the American Statistical Association*, 76:817–823, 1981.
- [10] A. Gupta. *Part description and segmentation using contour, surface and volumetric primitives*. Technical Report MS-CIS-89-33, University of Pennsylvania, Dept. of Computer and Information Science, Philadelphia, Pa., 1989.
- [11] R. Held and J. Bauer. Visually guided reaching in infant monkeys after restricted rearing. *Science*, 155:718–720, 1970.
- [12] C. Von Hofsten. *Motor Development in children: aspects of coordination and control*, chapter The emergence of manual skills. Martinus-Nihoff, Dordrecht, 1986.
- [13] P.J. Huber. Projection pursuit. *The Annals of Statistics*, 13:435–475, 1986.
- [14] M. Kuperstein. Neural model of adaptive hand-eye coordination for single postures. *Science*, 239:1308–1311, 1988.
- [15] Thomas Lindsay and R.P. Paul. *Design of a Tool-Surrounding Compliant Instrumented Wrist*. Technical Report MS-CIS-91-30, GRASP LAB 258, University of Pennsylvania, Dept. of Computer and Information Science, Philadelphia, Pa., 1991.

- [16] P. Maes and R.A. Brooks. Learning to coordinate behaviours. In *AAAI-90*, pages 796–802, 1990.
- [17] M.T. Mason, A.D. Christiansen, and T.M. Mitchell. Experiments in robot learning. In *Proceedings of the Sixth International Workshop on Machine Learning*, pages 141–145, Morgan-Kaufman, 1989.
- [18] B. Mel. *Connectionist robot motion planning: a neurally inspired approach to visually guided reaching*. Academic Press, San Diego, CA, 1991.
- [19] J. Piaget. *The origin of intelligence in the child*. Norton, NY, 1952.
- [20] E.A. Roy and J. Starkes. *Motor development in children*, chapter The contribution of Neuroscience to the understanding of the development of reaching, pages 443–456. Martinus-Nijhoff, Dordrecht, 1986.
- [21] W. Twitchell. *Mechanisms of motor development*, chapter Reflex mechanisms and the development of prehension. Academic Press, New York, 1970.
- [22] N.T. Ulrich, R. Paul, and R.K. Bajcsy. A medium complexity end-effector. In *Proceedings, 1988 IEEE International Conference on Robotics and Automation*, pages 434–437, 1987.