First IEEE-RAS International Conference on Humanoid Robots, 7–8 September 2000, Cambridge, MA, USA.

Developing Haptic and Visual Perceptual Categories for Reaching and Grasping with a Humanoid Robot

Jefferson Coelho, Justus Piater, and Roderic Grupen

Laboratory for Perceptual Robotics,
Department of Computer Science
University of Massachusetts
Amherst, MA 01003
{coelho, piater, grupen}@cs.umass.edu
http://www-robotics.cs.umass.edu/

Abstract. Properties of the human embodiment – sensorimotor apparatus and neurological structure – participate directly in the growth and development of cognitive processes. It is our position that relationships between morphology and perception over time lead to increasingly comprehensive models couched in the dynamics of interactions with the world. We are studying humanoid robots in order to apply insight derived from neuroscience, neurology, and developmental psychology to the design of advanced robot architectures. Increasingly compelling evidence from these communities indicate an intimate relationship between form, control structure, and patterns of cognitive growth. Somehow these interactions manage to defeat the awesome complexity associated with controlling a human's sensorimotor apparatus. To find out how this happens, we believe it is necessary to approximate the human sensorimotor configuration and to engage sensory and motor subsystems in developmental sequences. Many such sequences have been documented in studies of infant development, so we intend to bootstrap cognitive structures in robots by emulating some of these growth processes that bear an essential resemblance to the human morphology.

In this paper, we motivate a framework for cognitive integration of haptic and visual sensorimotor modalities. Specifically, we will model the development of prospective grasping behavior in human infants. Prospective behavior describes a process in which behavioral utility is discovered and associated with a prior perceptual condition and action (a *cause*) that reliably distinguishes between desirable and undesirable *effects*. Our example shows how exploratory grasping behavior using a dextrous robot hand can lead to candidate grasps of varying quality and illustrates a learning mechanism that reliably selects high quality grasps on the basis of haptic feedback. Subsequently, the haptic policy can be partially subsumed by learning visual features that bootstrap grasp formation. The result is a trajectory toward associative visual-haptic categories that bounds the incremental complexity of each stage of development.

1 Introduction

Human infants display a tremendous assortment of time-varying structure in their physiological and neurological responses to the world. We speculate that this growth process provides important insight into how infants manage the complexity of learning while acquiring increasingly sophisticated mental representations. Right now, developmental psychologists and roboticists are proposing similar theories of sensorimotor development – namely, that latent aptitudes expressed by virtue of the kinematic, dynamic, and "neurological" properties of a developing agent are exploited to simplify and structure learning in the context of an on-going interaction with the world. The temporal sequence of developmental processes appears to lead to tractable incremental learning tasks. We present a framework developed to provide the basic mechanisms in support of cognitive growth for a humanoid robot. The proposed system architecture has been used to study methods for learning control [31], acquired representations [12], and on visual behavior [55, 24].

The *interactionist* representation grounds human knowledge in *activity*. From this perspective, "motor timing in skilled actions is discovered . . . through perceptual exploration of the body's intrinsic (or autonomous) dynamics within a changing task and physical space [70]." It is the potential for rich and varied interaction with the world that we contend necessitates cognitive organization and development in humans – this is a critical issue that has been largely overlooked by the AI community.

The human hand has often been cited as an important factor in the development of the apparently superior ability of the human brain to form critical categories in sensorimotor experience [75]. Many are of the opinion that this faculty for building predictive models underlies much of what we recognize as human-level cognitive ability. While experts disagree on cause and effect, it is clear that the mechanical dexterity and redundancy afforded in the hand requires a neural architecture capable of modeling a huge variety of interactions with the world. Our decision to study reaching, grasping, and manipulation is motivated by our desire to understand one of the important missing aspects of intelligent systems research as well as by our desire to construct general purpose end effectors and integrated perceptual abilities for robots. We postulate that the processes underlying multifaceted world models effect problem solving in general as well as the formulation of skillful manipulation strategies.

We pose the development of robot programs as an incremental search for strategies that exploit the intrinsic dynamics of the robot/world interaction. "Intrinsic dynamics" is interpreted fairly broadly as any kinematic, dynamic, perceptual or motor synergy that produces characteristic and invariant temporal sequences of observable state variables. Humanoid robots are simply too complex to make use of traditional approaches from robotics and computer vision. The range of interaction possible and the kinds of perceptual distinction required challenge commonly used methodologies for control and programming. Consequently, the architecture proposed in this paper adopts an incremental and automatic approach to programming modeled after the sensorimotor development of human children in the first two years of life. In this period, genetically-mediated maturational mechanisms focus the infant on simple problems first and subsequently enrich these policies by including additional motor and perceptual systems [27]. Infants are constantly learning about the capabilities of their motor systems and adapting motor strategies in accord with their current level of sensory and motor control [4]. Early sensorimotor programs are not burdened with the full complexity of the infant neuroanatomy. Instead, maturational mechanisms in the brain, co-contraction of distal degrees of freedom, and evolving neurological structure organize and direct evolving motor programming. Attentional mechanisms participate in this growth process and, therefore, it is critical that flexible means of directing attention in humanoid robots are developed that can be varied as a function of time. The framework reported in this paper is intended to be a first step in that direction.

2 Relationship to the Literature

Three principle threads in the research community are immediately relevant to our on-going project. The first is the body of analytical results in the robot grasping community. The second is the growing interest among behavioral scientists and roboticists regarding the use of models of dynamics and the desire to exploit the intrinsic dynamics of controlled processes. Finally, we review methods for learning visual recognition strategies as they have been applied in computational systems.

2.1 Grasp Mechanics

A great deal of progress has been made in the mathematical analysis of phenomena associated with grasping and manipulation tasks [50]. We have relatively standard models of contact types consisting of point contacts with and without friction, and soft-fingers that can produce torque around the contact normal [42, 17]. To move control contact during manipulation, work has been done on how to exploit slippage [20, 8, 14, 74, 37], and rolling contact geometries [48, 29, 10, 50]. Some of the most widely read literature on grasping concerns the conditions under which a grasp can restrain an object. Motivated by fixturing problems in machining operations, a form closure grasp typically considers the placement of frictionless point contacts so as to fully restrain an object [41]. Force closure properties speak to the ability of a grasp to reject disturbance forces and usually considers frictional forces [21, 53, 19]. We have adopted insights from these results in the work reported here, however, the cited approaches rely ultimately on complete geometric models which is not appropriate for the problem specification we consider. Consequently, we propose a closed-loop grasp primitive that tend locally toward null spaces in the grip Jacobian [62] which is a necessary condition for force closure with frictional forces.

Theoretical analysis of the stability of an object within a grasp is typically focused on the size and steepness of a potential well determined by the grasp that tends to restore the object to a equilibrium position [73, 30]. This is useful insight, especially when comparing otherwise equivalent alternatives, but it is noteworthy to mention that humans use many grasps in everyday life that are technically unstable by this analysis. Moreover, it is quite difficult in practice to provide useful specifications of stability, especially in the context of other competing objectives.

Despite the significant theoretical impact of this literature, we have not yet developed an adequate model of the sensory and motor process of grasping and manipulation. This process moves fluidly through multiple contact regimes and can trade stability margins early during manipulation for constructive interactions later in the operation, e.g. as usually imposed by pick-and-place constraints. Moreover, nearly all the work on multifingered grasping considers a complete geometrical model of the object and most depend on geometrical reasoning to compute a grasp – this despite the fact that grasping is inherently a force domain task. Finally, we feel that the real challenge and opportunity afforded by multifingered hands is the automatic modeling of complex and non-stationary modes of interaction between a robot and the world. Modeling end-to-end manipulation sequences leads immediately to issues of representation and learning – issues that have been largely ignored to date.

2.2 Dynamical Systems – Movement Units for Robot Control

The human central nervous system (CNS) is organized according to *movement patterns* [2]. The basic form of packaged movement pattern is the reflex, which can reside at many levels of the central and peripheral nervous system. These processes contribute to the organization of behavior at the lowest levels. This form of native control structure speaks to the vegetative needs of the organism and provides an "instruction set" for composing more complex behavioral programs. All map stimulus to response – the so-called simple segmental reflex does so in a predominately open-loop fashion and the more advanced brain stem and cerebellar-mediated reflexes in a more closed-loop fashion. It is generally understood that reflexes exercise the musculature and provide useful motor responses to simple reoccurring situations, but it may also be true that they serve to increase the exposure of learning and developmental processes to conditions underlying important developmental milestones.

Infants are born with neural and skeletomuscular systems that produce patterned and timed movements. Constructive interactions within the anatomical substrate can be stimulated by a variety of environmental contexts. During the first several months in an infant's life, reflexive responses begin to organize into coherent motor strategies, sensory modalities are coordinated and attentional mechanisms begin to emerge. Native reflexive responses like the primary walking reflex and the palmar grasp reflex [2] provide primitive, closed-loop sensorimotor behavior that accomplish sensory-driven work in the world. Bruner [9] refers to these types of behaviors as "preadaptation" primitives for learning skillful motor policies. Subsequently, policies for coordinating multiple sensory and motor modalities appear as primary circular reactions [54] which are practiced until the infant finds it possible to prolong certain interactions with the world.

In 1989, Koditschek *et al.* argued that robot designers should focus on "finding controllers" with inherent dynamical properties that produce valuable artifacts in the world rather than computing the artifacts directly [60]. Assertions about the stability of the coupled system usually form the state space for such systems as in the attractor landscape proposed by Huber *et al.* [31] or the limit cycles proposed by Schaal *et al.* [63]. Koditschek uses prior stability assertions based on Lyapunov functions to predict when discrete attractors can capture the state of the system. This information supports switching policies to achieve "juggling" tasks [38, 61, 39]. This paper adopts these methodologies as well.

Currently there is a great deal of interest in the research community regarding adaptive control architectures for non-stationary, nonlinear processes [51, 28, 61, 58, 7]. Monolithic optimal controllers don't exist in general for highly non-linear processes, so these approaches postulate a family of local models that can be used to approximate the optimal, global control surface [47]. By switching controllers, or by reformulating local models, a linear control substrate can be applied more generally to nonlinear and/or non-stationary tasks [58, 1]. As a result, the robot control program is generally more robust. Some of these approaches incorporate learning methods for control law synthesis [47, 1, 35, 43, 45]. If local control models are stable (in the sense of Lyapunov), then they can actively restrict the state to a neighborhood surrounding the attractor, thus approximately preserving some property in the system until conditions permit a transition to another attractor [61, 35]. We will employ this approach to express grasping behavior on multiple object types and will learn grasping policies that switch between closed-loop grasp controllers.

2.3 Learning Discriminatory Visual Features

How do humans learn to recognize perceptual cues in the world? Two principal hypotheses can be identified [57]. According to the Schema Hypothesis, sensory input is matched to internal representations of objects that are built and refined through experience. On the other hand, the Differentiation Hypothesis holds that contrastive relations are learned that make relevant distinctions. Psychological evidence argues strongly in favor of a differentiation learning framework [57, 67, 76, 66]. As we interact, we learn to pay attention to perceptual cues that are behaviorally important. For instance, we learn to recognize and distinguish individual objects and form categories on the basis of relevance. This community argues against fixed features and for an active process of identifying new features that serve to form a relevant distinction in the task.

Most work in machine vision concentrates on Schema methods without a developmental component. Hence the performance characteristics of most existing machine vision systems are largely determined *a priori* by the design of features and matching algorithms. Nevertheless, impressive systems exist that use sophisticated statistical, texture- and shape-based features and recognition algorithms and perform very well on closed tasks where all training data are available at the outset [46, 49, 52, 64]. Despite the noteworthy progress advanced in the form of the Schema hypothesis and the visual recognition tasks, there has been a dearth of results drawing on the Differentiation hypothesis. This approach appears, nonetheless, critical for embedded and developing perceptual systems in open domains. These techniques promise to find a basis for learning discriminative abilities on the basis of behavioral utility [55] and hold a great deal of promise as descriptive computational accounts of infant development.

3 The UMASS Humanoid Platforms

The results reported in this paper require certain essential relationships to the human morphology. Perhaps foremost among these is a multifingered robot hand. As we cited previously, these devices require accompanying technologies for modeling the variety of interactions that they afford in open environments. We have previously reported results employing the Utah/MIT robot hand, but the work we describe here employs the Stanford/JPL (or Salisbury) hand. The Utah/MIT hand would not easily accommodate fingertip tactile sensation that was critical for this work. Brock sensors were fitted to the Stanford/JPL hand to provide observations of contact position and normal. This feedback provides the basis for our closed-loop grasp controllers. The hand requires a degree of mobility in space so as to permit the flexible application of contact resources. The robot hand is placed on the end of a 5 DOF GE P50 robot arm. Although not the subject of this paper, this hand/arm configuration permits kinematic properties of the hand to drive arm movement – a capability critical to our approach. The resulting 14 DOF effector provides the essential mobility and reconfigurability required to model human grasping processes.

In addition to kinematic and tactile dimensions of the human configuration, we require visual input to determine spatial targets for reaching tasks and to associate visual features with grasp control parameters. The data reported in this paper was derived from a monocular vision system with sufficient additional knowledge to recover range to objects. The search for features that parameterize grasping behavior is accomplished in monocular image frames acquired prior to the reach-and-grasp process. There is no specific requirement for the geometry of the hand and eye.

The ultimate target platform for this work is the newly constructed UMass humanoid torso, *Magilla*, illustrated in Figure 1. It consists of two Whole Arm Manipulators (WAMs – Barrett Technologies^{TM}), two multi-fingered

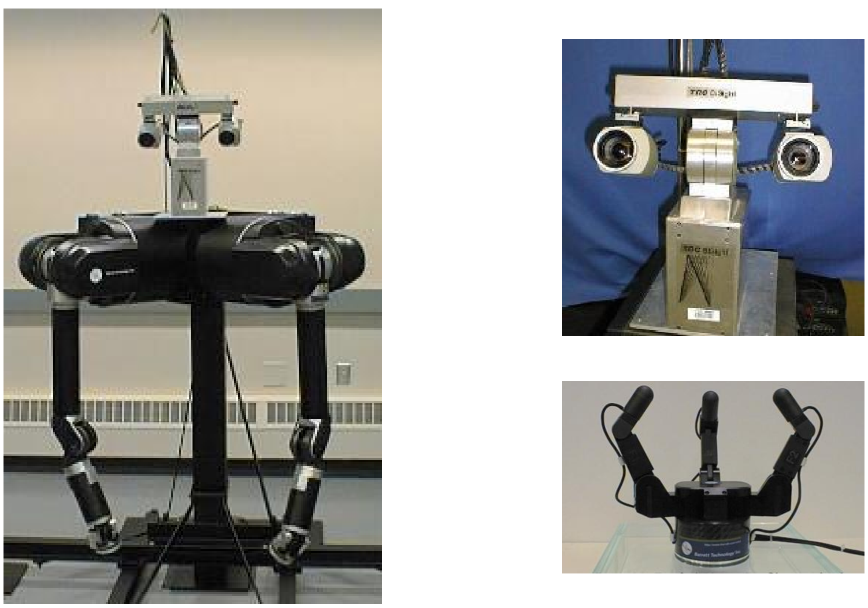


Fig. 1. Magilla – the UMASS Humanoid Torso.

Barrett hands, and a TRC BiSight stereo head. Later, we intend to add a multi-aural auditory system as well. Each arm is a seven degree-of-freedom (DOF) backdrivable manipulator with roughly anthropomorphic scale and kinematics. Magilla's hands are two BH8-255 Barrett Hands, each with three fingers and a total of 4 DOF. Two of the fingers track laterally around the perimeter of the palm through 180 degrees synchronously. This supports hook as well as opposition grasp types. Tactile (ATI Nano17 force/torque) sensors are implemented in the fingertips, allowing recovery of contact positions and normals.

Visual information for *Magilla* is provided by the articulated stereo head consisting of two video cameras mounted on a TRC BiSight head providing four mechanical degrees of freedom: pan, tilt, and independent left and right vergence. Six more optical degrees of freedom are controllable: iris, zoom, and focus, independently for each eye. Motion is controlled via a PMAC/Delta TAU interface. Images from each camera are input to a Datacube pipelined array (image) processor (IP).

4 Computational Framework for Humanoid Development

Figure 2 is a sketch of a computational framework that addresses the development of manual skills with robot hands. One dimension of development is viewed as a scheduling problem in which robotic resources are engaged to

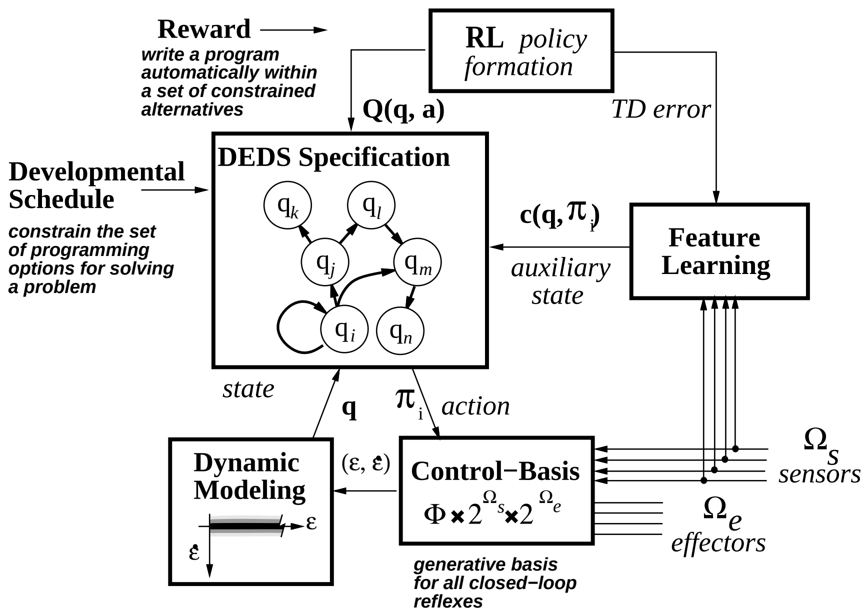


Fig. 2. Native Structure, Learning, and Development in an Integrated Architecture.

satisfy a task. Primitive actions are closed-loop control processes constructed by combining an artificial potential, $\phi \in \Phi$, with a subset of the available sensors, Ω_s, and effectors, Ω_e. As these controllers interact with objects and tasks, a set of prototypical dynamic models are constructed that identify haptic categories during grasp formation. We will introduce the control processes used for grasping in Sections 4.1.

In the Dynamic Modeling component of Figure 2, we show the initial, native model for all states and actions. This model describes convergence of the grasp controllers since it contains the set of states where the time derivative of the control error, $\dot{\epsilon}$, is near zero. This implies that initially control decisions may only happen when one or more of the working grasp controllers approaches equilibrium. Moreover, the very first policies will move between discrete equilibria in the working controllers. As the robot accumulates experience with an object, additional models are constructed that describe prototype transient responses of the grasp controllers. The manner in which the grasp controllers eliminate error over time depends on object geometry and the local grasp attractor, so these haptic categories can be used to predict the eventual grasp quality relative to the reward function. If the predicted quality is unacceptable, the haptic category can be used to make control decisions that cause the system to navigate through a landscape of attractors toward those that satisfy end-to-end task specifications. The pattern of membership in these models over a working set of controllers and the controllers themselves form the *states* and *actions* for a Markov Decision Problem (MDP). Nodes in the graph depicted in Figure 2 are states and transitions involve concurrent grasp control processes selected from the set of available actions.

To balance expressive power against computational tractability, the Discrete Event Dynamic Systems (DEDS) specification constrains the range of interactions permitted with the environment to those that:

- are consistent with a resource model specifying which combinations of resources are relevant;
- satisfy real-time computing constraints;
- guarantee safety specifications while learning;
- are consistent with kinematic and dynamic limitations; and
- express a developmental schedule to learn complex activities incrementally.

The DEDS specification is designed to eliminate irrelevant or unsafe control combinations during on-line, incremental learning tasks. Together with the task (reward function) it focuses exploration on the horizon of available control knowledge in order to build internal representations of important (sub)tasks. This, in and of itself, can lead a robot through a sequence of developmental milestones by shaping the sequence of policies acquired [33] and making "options" (temporally extended sequences of control) available as abstract actions [69].

Finally, once the utility of haptic categories and policies for moving between attractors are learned and compiled into value functions, we may use mature haptic value functions as the basis for visual discrimination tasks. Visual features are sampled from a possibly infinite set of alternatives to discriminate between haptic categories and index a set of relative hand postures distinguished by their utility in the end-to-end manipulation task. In a simple form of visual guidance, the Feature Learning component of Figure 2 is used to construct robust visual features that recommend particular spatial goals for reaching. These goals place the hand in positions relative to the object that are upstream of optimal equilibria in the grasp controllers. The overall architecture is designed to extend a native representation in two ways:

1. to accumulate models of grasp dynamics to enhance state and to derive temporally extended actions, and
2. to identify important visual distinctions on the basis of discernible differences in haptic utility.

Visual categories provide auxiliary (non-native) state information with which to make control decisions. The manipulation policy informed by this additional state may leapfrog toward valuable attractors by reaching directly to them rather than groping through intermediate and suboptimal haptic categories.

4.1 A Closed-Loop Motor Control Basis

All successful organisms exploit some form of native structure (neurological, muscular, skeletal), many employ mechanisms for neural adaptation, and all successful species settle into a stable dynamic relationship with their environment. Successful biological systems exploit the intrinsic dynamics of bodies and tasks. Moreover, humans learn, by means of a developmental trajectory, to exploit favorable dynamic relationships to the world by using acquired control knowledge. The design of a native control repertoire for a synthetic system should be both flexible and expressive enough without introducing undue complexity. The motor unit in this paper employs a closed-loop control basis, this design is consistent with perspectives in infant motor development and adult motor control [72, 71, 4, 5] and robotics [25, 18].

The control basis $\Pi = \{\pi_1, \pi_2, \ldots, \pi_n\}$ represents the agent's native control structure. In our formulation, each $\pi_i \in \Pi$ is a closed-loop controller based on simple, local models of how the agent affects its environment by applying inputs to its actuators. To structure control and to reduce the complexity of behavior composition, the control basis approach uses a small set of feedback control laws to construct complex behavior on-line. End-to-end tasks are solved by combining and sequencing elements of Π as proposed in [35]. The control basis effectively organizes a continuous, high dimensional state space into an enumerable set of attractors. In its simplest form, control is expressed in terms of activation and convergence events in the participating controllers. Navigation controllers [15, 16], contact controllers [12], and kinematic conditioning controllers [26] have been used in experiments in complex, multiple hand-and-arm systems [11, 68], in adaptive, aperiodic walking gaits [34, 40], in foraging tasks [1], and in visual servoing tasks [65].

Figure 3 depicts the closed-loop control scheme proposed. The plant consists of an k-contact grasp configuration on an unknown object geometry. The feedback controller relies on a frictionless point contact model represented by a unit force oriented along the inward directed surface normal. Sensor evidence in the form of contact positions and normals is used to construct the instantaneous grip Jacobian, G, with which to transform contact forces into object frame wrenches. This local characterization captures the ability of the object surface to carry contact forces and to generate object frame wrenches.

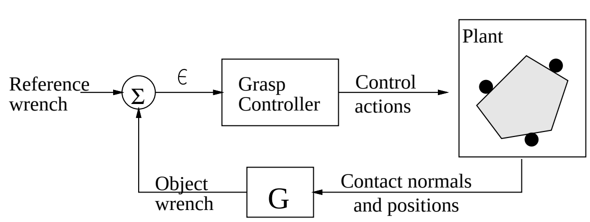


Fig. 3. Grasp synthesis as a control problem.

A necessary condition for force closure requires that the grip Jacobian, G [62], contains a null space consisting of wrenches derived from strictly positive-sense contact forces. The controller constructs the grasp configuration by continuously adjusting the contact coordinates to achieve a reference net wrench. If there exists a contact wrench that can be written as a positive linear combination of other contact wrenches, $\omega_i = \sum_{j \neq i} \omega_j$, then the object may be *squeezed* within a null space of the grip Jacobian without applying a net wrench to the object. It follows that the condition $\sum \omega_i = 0$ satisfies a necessary condition for force closure with frictional forces. While most researchers describe grasp synthesis as an optimization problem, we proposed that it is best characterized as a robust control problem. In this framework, the robot uses tactile feedback to compute incremental contact displacements.

Contact displacements are determined by the grasp controller π_c described in [12] whose potential field gradients are based on local models of interaction between the contacts and the object surface, and are aimed at reducing the squared wrench residual ϵ measured at the object's center of mass. The squared wrench residual ϵ is defined as the sum of squared force and torque residuals. Given the wrench residual vector for n contacts

$$\rho = \sum_{i=1}^{n} \left[f_x^i \;\; f_y^i \;\; f_z^i \;\; \tau_x^i \;\; \tau_y^i \;\; \tau_z^i \right]^T ,$$

then the squared wrench residual is defined by

$$\epsilon = \rho^T \rho.$$

The controller π_c displaces the subset c of contacts until a local minimum for ϵ is reached. Minima in ϵ correspond to the existence of a null space of rank 1 or higher in the grasp matrix, G. The subset of contacts c specify which fingers and surfaces are enlisted in the grasp task. A hand with 3 fingers labeled $\{T, 1, 2\}$ permits 4 distinct fingertip contact subsets, assuming that two or more fingers are required to grasp the object:

$$\mathcal{C} = \{(T,1), \;\; (T,2), \;\; (1,2), \;\; (T,1,2)\}.$$

Each instance of $c \in \mathcal{C}$ defines a new control law. The controller π_c is an element in the family of grasp controllers $\Pi = \{\pi_c | c \in \mathcal{C}\}$.

The control actions of the controller π_c are dependent solely on instantaneous, local tactile feedback. Convergent configurations for π_c correspond to local minima of ϵ. Each choice of control law $\pi_c \in \Pi$ leads to distinct convergent grasp configuration for a given object orientation. Therefore, there exists an optimal choice of grasp resources, c, for each orientation of the object that yields a convergent configuration with the minimum ϵ. The idea can be extended to controller sequencing: given a certain initial configuration, there exists an optimal sequence of controllers that lead the system state to the solution with the smallest possible ϵ. More importantly, controller sequencing expands the capabilities of the primitive controllers and allows one to build a system that can adapt to many operational contexts.

4.2 Constructing Models of Control Dynamics

When controller π_i is active, the observed system dynamics are partially obscured by noisy sensors and actuators, so probabilistic models are appropriate. The use of parametric models to represent a sequence of observations is common practice in the dynamical systems literature (e.g. see Fraser [23]), especially where insight about the underlying phenomena is available.

Parametric models presume the existence of a (parametric) generator mechanism for the data observed; the structure of the parametric model must be chosen according to the phenomenon one wants to model. We define an observation, $\mathbf{o} = [\epsilon \;\; \dot{\epsilon}]^T$ where ϵ and $\dot{\epsilon}$ are the squared residual and its time rate of change, respectively. We assume the the observation will evolve along a piecewise continuous contour in the "residual" phase portrait. to an equilibrium configuration, $[\epsilon_0 \;\; 0]^T$. We model these trajectories using linear segments – this assumption is justified for our controllers in [13]. A particular model, M_{π_i}, with parameters ϵ_0 and K, predicts an observation $\tilde{\mathbf{o}} = [\epsilon \;\; -K(\epsilon - \epsilon_0)]^T$ given the observed residual squared error ϵ. We further assume that ϵ may have superimposed noise $\sim N(0, \sigma^2)$ so that probabilistic membership of observation $\mathbf{o}$ in M_{π_i} can be estimated by:

$$\eta = (\mathbf{o} - \tilde{\mathbf{o}})$$
$$M_{\pi_i}(\theta, \mathbf{o}) = \frac{1}{L} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-\eta^T \eta}{2\sigma^2}\right), \tag{1}$$

where the parameter vector $\theta = [K \;\; \epsilon_0 \;\; \sigma^2]^T$ and L is a normalization constant.

The complete representation of system dynamics under policy π_i requires a set of m observation models, expressed as $\mathcal{M}(\pi_i) = \bigcup_{k=1}^{m} M_{\pi_i}(\theta_k, \mathbf{o})$. The set $\mathcal{M}(\pi_i)$ expresses empirical knowledge acquired by the agent during the execution of policy π_i. Each model is valid only within a bounded domain D_k, $M_{\pi_i}(\theta_k, \mathbf{o}) = 0$ if $\mathbf{o}$ is not in D_k.

The derivation of $\mathcal{M}(\pi_i)$ involves sampling system dynamics for a predetermined number of epochs τ, while recording the data $\mathcal{O} = \{\mathbf{o}_1, \mathbf{o}_2, \ldots, \mathbf{o}_n\}$ observed as the control error evolves toward equilibrium. The instantiation of new models involves two steps, (1) the derivation of τ observation models corresponding to the observation sequences recorded in each of the τ epochs, and (2) the elimination of redundant observation models when $|\theta_j - \theta_k| < \delta$. Many optimization procedures can be used to derive the parameter vector θ_k under the assumptions stated earlier. Model construction is concluded when $\mathcal{M}(\pi_i)$ has been constructed for every $\pi_i \in \Pi$. Any single object will present several unique models and models sets are not disjoint for different objects. Once $\mathcal{M}(\pi_i)$ is available, Bayesian estimation is used to identify the subset $\mathbf{q} \subset \mathcal{M}(\pi_i)$ of models compatible with a sequence of run-time observations. The *state* of the system is defined as the concatenation of the control law and the membership pattern, $(\pi_i, \mathbf{q})$.

The evolution of the grasp process captured by the sequence $\mathcal{O}$ can be plotted in phase space, as illustrated in Figure 4. The left panel in Figure 4 depicts the grasp dynamics for a typical two-fingered grasp of an irregular triangle. Configuration changes under π_c are represented by a path in the phase plane. Initially, the contacts are in configuration (a). π_c drives the system to an intermediate grasp configuration (b), and converges to configuration (c) – a minimum in the squared wrench residual ϵ where the velocity $\dot{\epsilon}$ is zero. Many other paths lead to the same attractor; in fact, the shaded region in Figure 4 represents the set of all states leading to the attractor corresponding to configuration (c). This region is termed the *basin of attraction* of the grasp attractor.

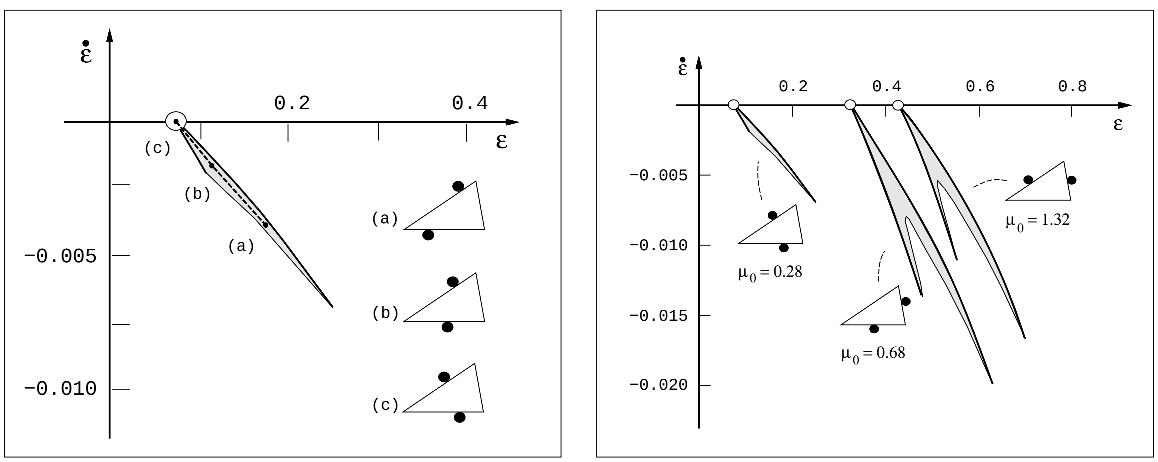


Fig. 4. Left panel depicts the evolution of a two fingered grasp trial from configuration (a) to (b) and to (c), the convergent configuration. The complete, two-fingered phase portrait for the irregular triangle is shown on the right.

The evolution of the grasp state (and the basin of attraction itself) can be represented by a set of paths, each of which captures a characteristic dynamic response. Paths representing the same environmental context can be combined to form a model of prototypical system behavior. For example, an illustration of all dynamic models for the irregular triangle is depicted in Figure 4 (right panel). It has three attractors and basins of attraction, corresponding to the three possible combinations of two contacts and three distinct edges. The convergent grasp configurations and respective quality indices μ_0 are also shown in Figure 4. The index μ_0 is the minimum friction coefficient required for a null space in the grip Jacobian, G, with rank ≥ 1 – it is a performance oriented label for each attractor and it is associated with all precursor states in the basin of attraction.

Figure 5 depicts a set of hypothetical models corresponding to policy π_i. If each model is given a discrete label (A, B, C, D, E), one can describe the transitions between subsets of models in terms of a discrete graph, shown in the right panel of Figure 5. As depicted, the regions in phase space in which two or more models overlap are identified with the labels corresponding to the overlapping models; (B, C) is one example. The resulting representation defines a discrete state space that describes the evolution of information in this grasping process.

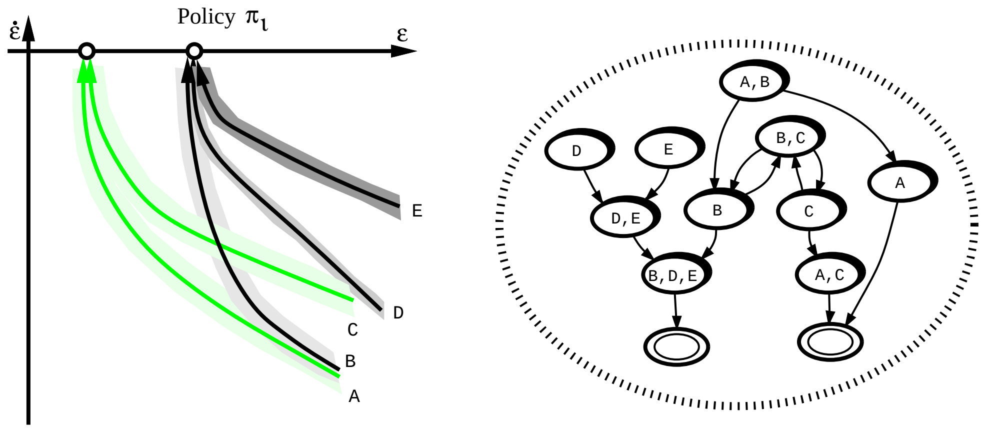


Fig. 5. Diagrams depict the phase portrait $[\epsilon \;\; \dot{\epsilon}]$ for policy π_i (left) and all possible context transitions (right).

5 Context-Dependent Grasp Policies

The transformation from a set of continuous models to a graph of discrete states can be carried out for each control policy $\pi_i \in \Pi$. The discrete state space allows the system to experiment with sequences of control policies, within the reinforcement learning framework. After convergence, the system will be able to employ the best policy for each state, and reduce the variability and uncertainty introduced by the many possible objects and overcome (to a certain degree) the locality of the component policies π_i.

Figure 6 illustrates how policy switching may lead to improved performance. Initially the system adopts policy $\pi_{T,1}$ to grasp an object of unknown geometry. The phase space coordinate is found and the information state is identified in the discrete context transition graph. Now, suppose that prior experience recommends switching to policy $\pi_{T,1,2}$, followed by $\pi_{T,2}$. This sequence has caused the grasp configuration to transform from a trajectory toward a suboptimal attractor to one headed toward the best two-fingered attractor through and intermediate three-fingered control context. Switching policies, as in this hypothetical example, forms finger gaits toward optimal contact configurations.

5.1 Pilot Data – Interaction Dynamics

Figure 7 illustrates a typical instant in the process of forming a haptically-guided grasp with the GE-P50 robot arm and Salisbury hand. The system attempts to identify sequences of control engagements that pass through robust haptic landmarks toward good grasps and in the process learns a great deal about the coupled dynamics of the hand/object/control system.

Dynamic programming-based Reinforcement Learning (RL) [3] is a natural paradigm for programming these systems since RL does not require external supervision and encodes policies as *sequences* of actions with associated rewards. In general, these rewards can be rare, occurring infrequently and only after extended sequences of actions. In the pilot study presented here, a simulation of our Stanford/JPL robot hand with Brock tactile sensors interacts with three simulated object types. We used a family of cylinders, and rectangular and triangular prisms with random variations in geometric parameters. The identity and orientation of the object are unknown at the beginning of each trial. Grasp policies are expressed as sequences of the four grasp controllers discussed earlier and RL is used to solve the temporal credit assignment problem for an optimal policy. The experiment involved 35 grasps using each of the four grasp controllers (exclusively) on each of three objects yielding $4 \times 3 \times 35 = 420$ data sets from which 61 separate models were retained.

The grasp controller makes a control decision every time the pattern of membership in the dynamic models changes – this event signals the fact that extra information has been acquired in the grasp experiment. Q-learning was used to derive the optimal switching policy using a Boltzmann exploration. The total number of training trials was 1600; in each trial a new object type with new geometric parameters was chosen randomly. The utility of pursuing a different control law at a decision point was evaluated after each state transition – the system has the choice of terminating the trial or invoking a different controller. If a grasp trial generates 50 contact movements, it times out and the current grasp configuration is scored. Terminal grasp configurations receive a score of $(1 - \mu_0)$,

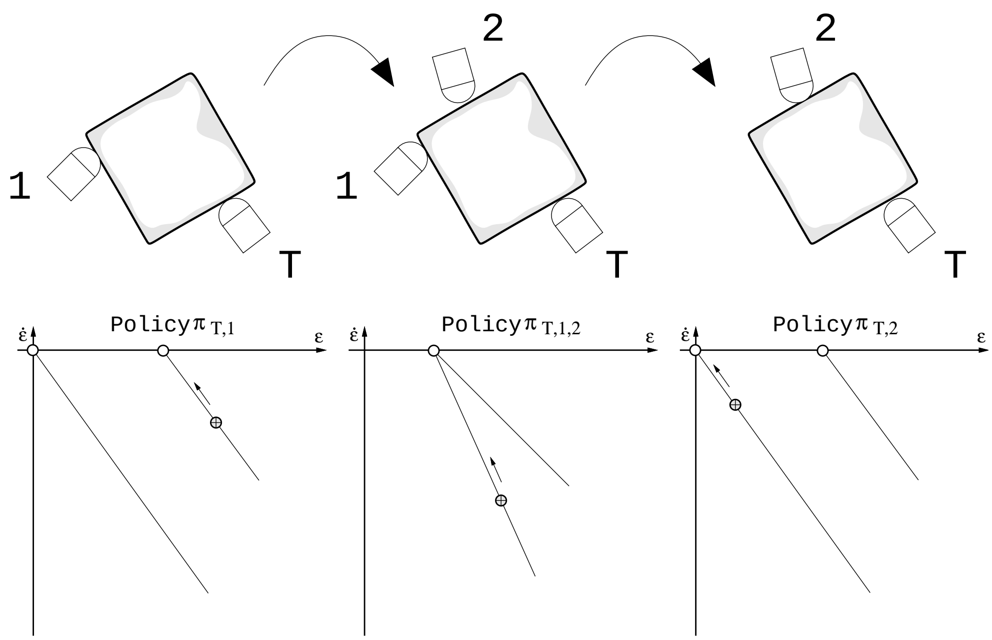


Fig. 6. A hypothetical context-dependent grasp of a cube (top view). Policy sequence $\pi_{T,1}$, $\pi_{T,1,2}$, $\pi_{T,2}$ accomplishes the best available two-fingered grasp configuration by using an intermediate three-fingered control context.

where μ_0 is the minimum coefficient of friction required to build a null space in the grip Jacobian with rank ≥ 1 in the final grasp configuration [12].

Figure 8 depicts a typical learning curve (curve labeled $\triangle\Box\bigcirc$, top left curve). The curve is an average of the ten most recent data points. Each point is the grasp score of a terminal grasp configuration (normally an attractor in the set of fixed points of the control basis) for a randomly chosen object. The data corresponding to each object are presented as well, the resulting learning curves are labeled $\triangle$, $\Box$, and $\bigcirc$. Because the control terminates with $\dot{\epsilon} < \delta$; $|\delta| > 0$, it will not be the case that the average score will go to 1. The curves for the individual objects are close to the optimal, within the limitations of the Q-learning algorithm.

Figure 9(a) shows a performance histogram over 100 grasp trials. In each trial, a random element of the control basis was applied to a randomly chosen object from the set. Figure 9(b) illustrates the distribution of results achieved on 100 trials with the acquired grasp policy. The grasp policy suppresses the majority of low quality solutions; 93% of the solutions have scores higher than 0.7, compared to 56% for the native controllers. The variance associated with solution quality is also substantially smaller. The same is true if we examine performance object-by-object (Figure 10).

5.2 Development and Incremental Robot Programming

Closed-loop behavioral primitives lead to models of the characteristic dynamics of grasp control interactions with the open grasping domain. Control activations may be considered in a symbolic state space for which we may derive an explicit system model (in the transition probabilities). The Discrete Event Dynamic Systems (DEDS) supervisor in Figure 2 can incorporate logical constraints on the outcome of actions and can, therefore, direct exploration and is a useful mechanism for *shaping* policy formation [32]. Time dependent sets of axioms in the DEDS specification can focus exploration on a sequence of computationally tractable sub-problems. We view this intervention as a developmental bias in which important control knowledge is accumulated over time.

Conjecture 1 (Reflexive Basis for Motor Development). Reflexes are scheduled in a manner consistent with other developmental mechanisms, in a sequence that leads an agent through a progression of incremental and environmentally-mediated learning tasks. These tasks acquire critical knowledge structures in an appropriate order.

Certain aspects of development appear to proceed through distinct resource constraints: from proximal to distal kinematic chains; from head to tail; from simple to complex tasks; from quasi-static to dynamic strategies; and from effects observed late in a behavioral sequence to prospective causes.

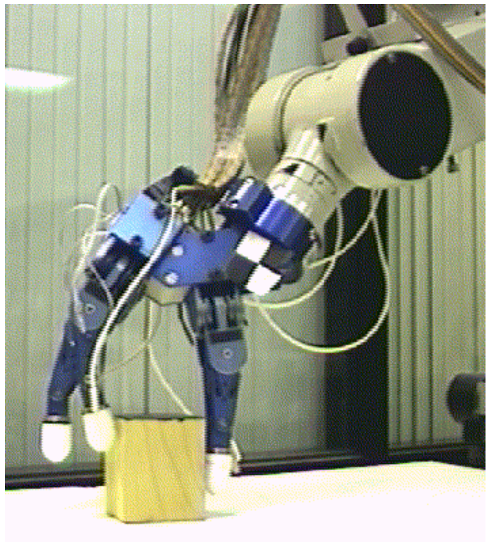


Fig. 7. The haptic grasping system.

Distinctions in behavioral utility late in a chain of actions (distal actions) can provide metrics for making early discriminations (proximal categories), For example, McCarty *et al.* studied the initial reach to a spoon laden with applesauce and presented to infants in left and right orientations [44]. The developmental trajectory observed is summarized in Figure 11. Initial policies are biased toward dominant hand strategies which work well when the spoon is oriented with its handle to the dominant side. However, when it is not, there is significantly less value to the dominant hand reach. Variations in the applesauce reward forms a discrimination metric space with which to distinguish important categories in this process – dominant-side and non-dominant-side presentations of the spoon. One hypothesis holds that this process involves a search for perceptual features that distinguish classes of behavioral utility. When this happens, *new* perceptual features have been *learned* that were not present in the original, native representation. They have been selected from a possibly infinite set of alternatives because they form a valuable distinction in the stream of percepts – valued for its ability to increase the reward derived from one's interaction with the task. One may view this process as one in which properties and constraints imposed by the task are incorporated into a policy incrementally starting with the latter (distal) actions and gradually propagating back through the action sequence to early (proximal) actions. There are parallels to so-called "pick-and-place" constraints studied in robotics [36].

6 Visual Context Recovery

When a mature human subject reaches for an object, the hand is oriented and shaped appropriately in anticipation of the grasp. This anticipatory pre-shape takes place before contact with the object is made, and is informed by visual cues. There is no conclusive evidence regarding what visual information is extracted and how it is used to inform the reaching process. At least, the applesauce experiment sheds some light on the developmental trajectory that leads toward sophisticated pre-shaping behavior. For our humanoid grasping system, and eventually for the integrated Magilla platform, we would like to develop an incremental learning system that produces skilled vision-based anticipatory behavior, and that parallels the developmental trajectory observed in humans.

The preceding sections described how sophisticated haptic grasping skills can be acquired through exploratory interaction with the environment. Experience produces models of the interaction dynamics between the hand and the grasped objects. These dynamic models provide relevant haptic context for robust, closed-loop grasping strategies. Haptic information provides powerful motor guidance for discovering high-quality grasps, which are relatively rare regions in the parameter space. However, the utility of haptic information for broader context recovery is limited due to its sequential and myopic nature.

Broader context can be provided by vision. Once learned haptic policies have been acquired for a task, the haptic context recovery component can be subsumed by a high-bandwidth visual modality that associates appropriate grasp parameters with visual features. If adequate visual features are found that robustly identify the control context, then these are equivalent in information content to the haptic dynamic models. Importantly, some haptic

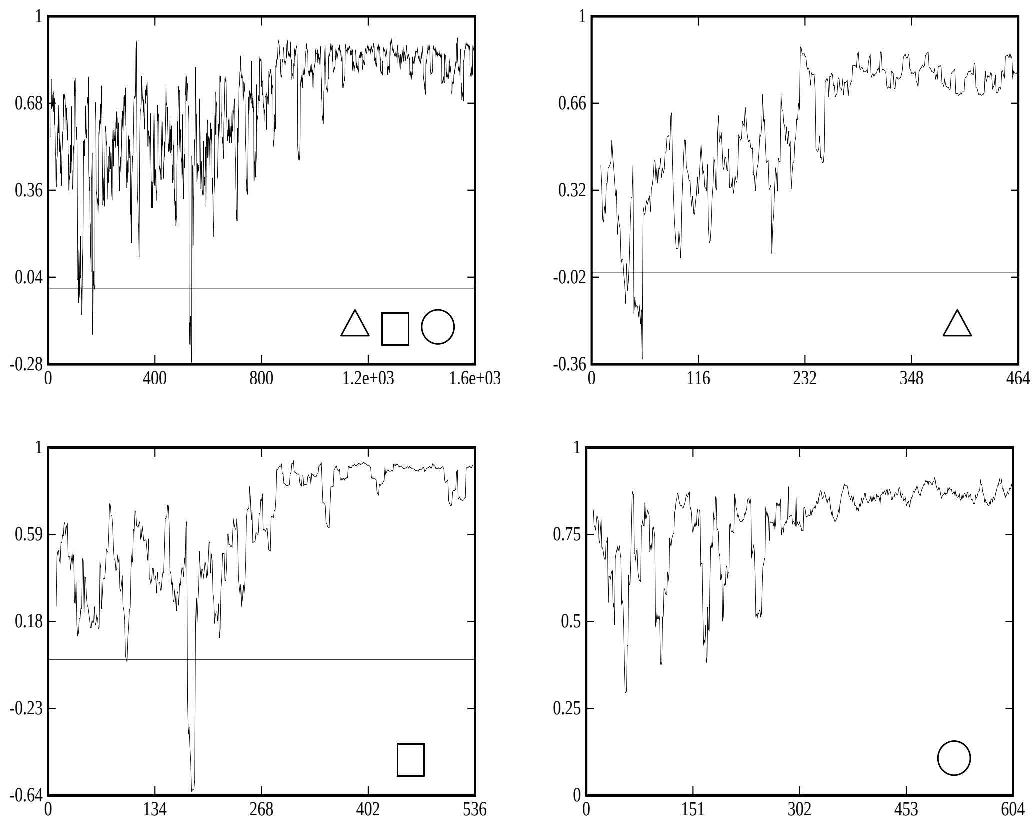
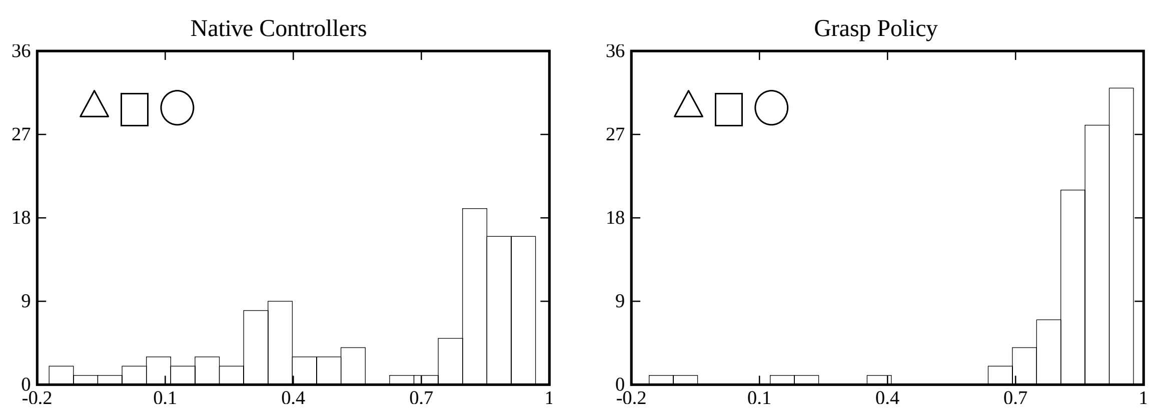


Fig. 8. A typical learning curve for the data set and learning curves for individual object types. Vertical axes are grasp scores $(1 - \mu_0)$, and horizontal axes are the trial number.



(a) (b)

Fig. 9. Distribution of grasp scores, over 100 trials. Left panel shows the average performance of the native controllers; right panel shows the result of the grasp policy whose state is derived from the dynamic models.

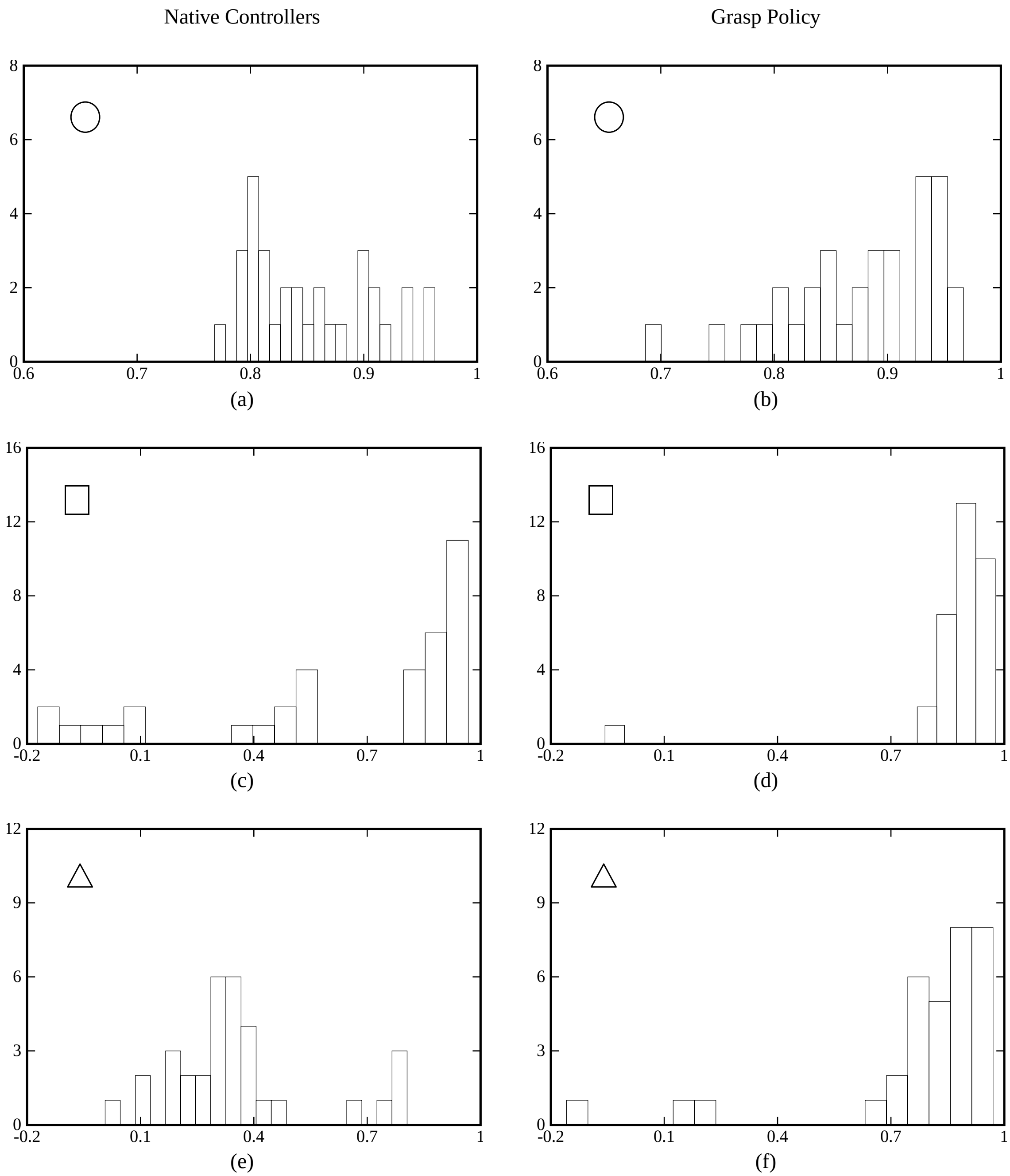


Fig. 10. Distribution of grasp scores by object type; cylinder (a and b), rectangular prism (c and d), and triangular prism (e and f). The left column shows the average performance of the native controllers, and the right column shows performance using interaction dynamics to provide context.

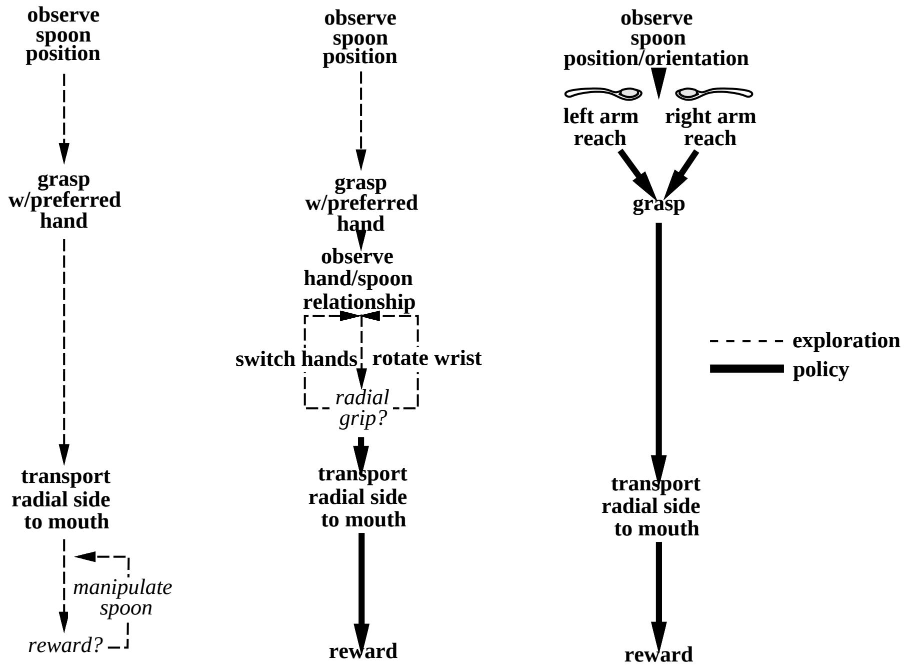


Fig. 11. Prospective Behavior revealed in the Applesauce Experiment.

context must exist before such visual features can be identified, because vision in and of itself cannot provide the haptic information required for grasping:

Conjecture 2 (Haptic-then-Visual Development). A haptic subsystem is employed first to discover useful grasping strategies at the expense of perceptual acuity and efficiency. Having identified such, haptic models form the basis for the acquisition of high-precision, efficient visual operators that recover important control contexts. This results in a powerful associative multi-modal model of interaction with objects in which haptic experience can be predicted by visual features and vice versa.

We now describe our current work on visual context recovery in support of the grasping system discussed above [56]. As a first step, the objective is to develop a plausible scheme for learning visual features that robustly correlate with the orientation of the hand during a successful grasp. Then, these features can be used to recommend a hand orientation *and* a native grasp controller for a two- or three-fingered grasp that should be engaged before the first tactile contact occurs, bootstrapping the haptically-driven grasp and eliminating the need for expensive and inefficient haptic context recognition. Each type of object may require a dedicated visual feature to fully capture the haptic context. Object identities are not known to the system, so the need for dedicated features must be discovered by grasping experience. Visual learning is entirely driven by the utility of the features to the haptic system.

6.1 Learning Visual Features that Predict Haptic Utility

To represent visual context, we employ local appearance-based features. Oriented derivatives of 2D Gaussian functions are used to form a steerable basis. This permits the efficient synthesis of features at arbitrary orientations, as well as the measurement of feature orientations [59]. Two types of *primitive* features are used: A *texel* is a vector consisting of filter responses from Gaussian-derivative operators of the first three orders; an *edgel* uses uses an orthogonal pair of first-order derivatives only [55]. Spatial *combinations* of these primitives can express a wide variety of shape and texture characteristics at various degrees of specificity. An incremental, on-line learning procedure assembles such compound features in a simple-to-complex manner, as the need for increasingly distinctive features arises. Figure 12 illustrates a geometric arrangement of oriented primitives that has been generated by such an approach to form a useful distinction. Visual distinctions in this framework need not be universal in the sense that they are tagged to particular states and tasks in the behavior of the system – sometimes inexpensive constellations of features are adequate for discriminating locally between important visual contexts.

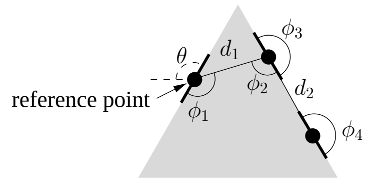


Fig. 12. A geometric feature of order 3, composed of three primitives. The feature is defined by the angles ϕ and the distances d, and the orientation of this specific instance is denoted by θ. Each primitive is either an edgel or a texel.

Each feature $\mathbf{f}$ is present at a pixel location l to a degree $s_{\mathbf{f}}(l) \in [0, 1]$, which is the normalized inner product of the vector of applicable filter responses at l with the pattern vector defining $\mathbf{f}$. A feature is present in an image I to the degree $s_{\mathbf{f}} = \max_{l \in I} s_{\mathbf{f}}(l)$. For more detail on these features, see our earlier work [55].

The vision system observes an object as it is presented and subsequently records the hand orientations associated with the best grasp for each object (as measured by the μ_0 metric; see Section 4.2). Assuming that these features respond to the object itself, their image-plane orientation $\theta_{\mathbf{f}}$ should be related to the robotic hand orientation θ_h by a constant additive offset $\Delta\theta$. A given feature, measured during many grasping tasks, hence generates data points that lie on straight lines on the toroidal surface spanned by the hand and feature orientations (Fig. 13). There may be more than one straight line because a given visual feature may respond to more than one specific object orientation (e.g., due to object symmetries), or to several distinct objects that differ in shape. To use these data for predicting hand orientations given a feature orientation, one needs to find the offsets $\Delta\theta$. This is an instance of the K-Means problem in one-dimensional circular (angular) space, with K unknown.

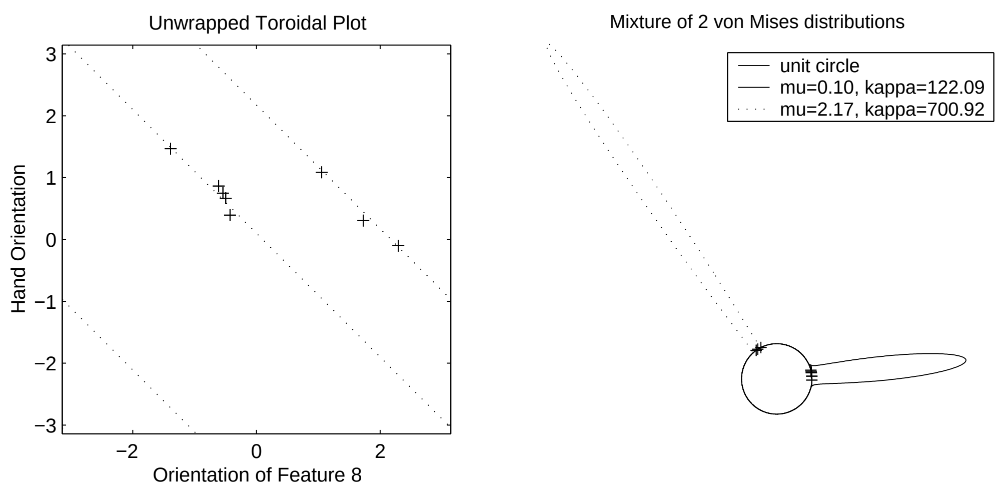


Fig. 13. Left: Data points induced by a given feature on various images of an object form straight lines on a torus (two in this case). Right: A mixture of two von Mises distributions was fit to these data. The probability density at an angle is visualized by the distance of the line from the unit circle.

To solve this problem, we assume the $\Delta\theta$ are drawn independently from a mixture of *von Mises* distributions. The von Mises distribution can be regarded as a circular equivalent of the linear Gaussian distribution, and has the probability density function [22]

$$f_{\mathrm{vM}}(\theta|\mu, \kappa) = \frac{e^{\kappa \cos(\theta-\mu)}}{2\pi I_0(\kappa)}, \quad 0 \leq \kappa < \infty$$

where $I_0(\cdot)$ is the modified Bessel function of order zero. The mean direction of the distribution is given by μ, and κ is a concentration parameter with $\kappa = 0$ giving a uniform circular distribution, and $\kappa = \infty$ to a point distribution. The mixture distribution (see Fig. 13) is defined by its density function

$$f_{\mathrm{mix}}(\theta) = \sum_{k=1}^{K} p_k f_{\mathrm{vM}}(\theta|\mu_k, \kappa_k)$$

with mixture proportions $0 < p_k < 1$ and $\sum_k p_k = 1$. For all plausible numbers of clusters K, a $(3K-1)$-dimensional non-linear optimization problem is solved to find the μ_k, κ_k and p_k. The objective function to be maximized is the log-likelihood of the observed data Θ given a parameterization $\mathbf{a}$ consisting of the μ_k, κ_k and p_k:

$$\log P(\Theta|\mathbf{a}) = \sum_i \log \sum_{k=1}^{K} p_k f_{\mathrm{vM}}(\theta_i|\mu_k, \kappa_k) \tag{2}$$

The most probable model can then be found using Bayes' Rule. In the case of uniform prior probabilities over all possible model parameterizations $\mathbf{a}_m$, the model $\mathbf{a}$ maximizing $P(\mathbf{a}|\Theta)$ is simply the one that maximizes $\log P(\Theta|\mathbf{a})$ (Eqn. 2). The appropriate number of clusters K is determined according to the Integrated Completed Likelihood criterion [6].

To recommend a hand orientation, the system selects from all features $\mathbf{f}$ that respond more strongly than a threshold $t_\mathbf{f}$ the feature with highest prediction potential $\mathrm{KSD}_\mathbf{f}$, introduced shortly. If the mixture model corresponding to this feature has more than one mode k that is supported by at least three data points, the mode with maximal κ_k is selected. The potential of each feature to make a useful recommendation is measured by the Kolmogorov-Smirnoff distance $\mathrm{KSD}_\mathbf{f}$ between the distributions of correct and wrong recommendations made in the past. The threshold $t_\mathbf{f}$ is selected such as to maximize $\mathrm{KSD}_\mathbf{f}$, under the premise that the feature $\mathbf{f}$ is not consulted if its response $s_\mathbf{f}$ in an image is less than $t_\mathbf{f}$. The result is that based on previous experience of the system, the Bayes-optimal feature (i.e., the feature with least expected misprediction rate) is selected from among all super-threshold features. The recommended hand orientation is then given by the orientation of the strongest occurrence of the selected feature $\mathbf{f}$ in the present image, and its associated $\Delta\theta$.

The system also determines whether to use a two- or a three-fingered grasp. For this purpose, separate feature sets (visual context models) are learned for two- and three-fingered grasps, and statistics are maintained of the grasp utilities (μ_0) associated with each feature. To form a grasp parameter recommendation, the best hand orientations are derived, as described above, separately for two- and three-fingered grasps. Of these two candidate recommendations, the one with the lower expected friction coefficient μ_0 is chosen.

Features are learned as follows. Given an image, the responses of all features are measured. The best feature is selected as described above, and is used to recommend a hand orientation. The robot then executes the grasp, starting with the recommended hand orientation. If the hand orientation turns out not to be appropriate, i.e. it needs to be corrected by more than a given threshold, then all mixture models are re-estimated based on a case list of previous experiences. A new prediction is made based on the new models. If this new prediction is still wrong, then two new features are generated: A primitive feature is randomly sampled from the image, and a new compound feature is generated by randomly expanding an existing feature by adding a new point as illustrated in Fig. 12. If a feature performs well, its KSD will increase over time, and it will increasingly be employed. If it performs poorly, its KSD will decrease, and it will eventually cease to be used at all. Unused features are discarded periodically.

6.2 Pilot Data – Hand Pre-Shaping Using Learned Visual Features

A series of pilot experiments was performed in simulation, using data generated by the real grasping system, and photo-realistically rendered, noise-degraded images. Three object types were used (Fig. 14). Lacking the ability to perform large numbers of grasps on the real robot, the recommended grasps were simulated by comparing the recommended hand orientation with the actually executed hand orientation associated with the training image, modulo the known rotational symmetry properties of the object. Since cylinders have infinite-fold rotational symmetry, no features were ever learned for cylinders.

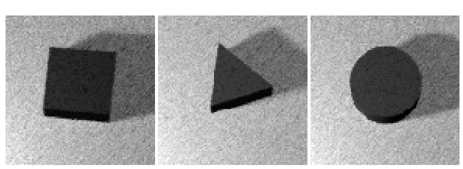


Fig. 14. Example views of objects used to test the system.

Our pilot studies indicate that the system learns to make useful recommendations (Figure 15). All results were computed in 2-fold cross-validation. If the training set contains a single object class and little noise in the training

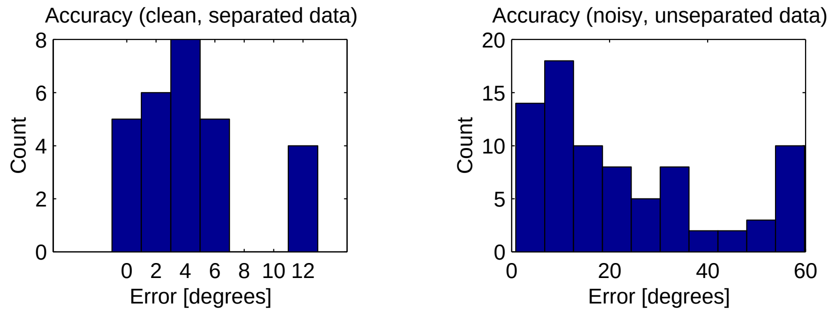


Fig. 15. Quantitative results of hand orientation prediction.

signal (the actual hand orientation during the grasp), the training set is typically learned during a single iteration. Performance on an independent test set is almost always excellent, with prediction error magnitudes on the order of the variation in the training signal. If the training set contains outliers, i.e. hand orientations that produced a poor grasp, then the training set is harder to learn because the system expends substantial effort trying to learn these outliers. However, performance degrades gracefully because features are selected by Kolmogorov-Smirnoff distance, which prefers reliable features modeling the the majority of useful training examples. On a noisy test set, most poor recommendations occur on outliers. Notably, two-fingered grasps of the triangular object are inherently unstable and unpredictable.

Figure 16 demonstrates the utility of the learned visual context to the haptic grasping system. The bottom row illustrates that neither two- nor three-fingered native controllers alone are sufficient to execute high-quality grasps reliably. The two-fingered native controller works well on rectangular but poorly on triangular prisms; for the three-fingered controller the opposite is true. If the recommendation of the visual system is followed, the achieved grasp quality is consistently high. Moreover, the proportion of extremely fast single-probe grasps increases drastically, and very long trials (more than about 20 probes) are practically eliminated (cf. the two-fingered native controller on the left). This visual/native-haptic policy performs about equally well as the “blind” learned policy described in Section 5 (cf. Figure 9b). Thus, visual context has almost subsumed haptic context in that it provides equivalent information *before* the onset of the grasp. We are currently evaluating the performance of the learned policy primed by the visual system – a cross-modal, redundant compound policy that corresponds quite closely to human grasping behavior.

7 Conclusion

We have presented a philosophy and motivation for studying humanoid robots and a perspective that aims to exploit insight from the social and behavioral sciences. We have also introduced our humanoid “Magilla” and reported preliminary results regarding the incremental acquisition of reaching and grasping skills. In our model, closed-loop haptic control models are acquired first, and are later augmented by visual context. A critical limitation of our present, simplified model is that the haptic and visual learning stages are largely separate. In order to develop inherently cross-modal associative models of interaction, a tighter integration of the haptic and visual modalities is required.

Magilla will shortly see the full-scale integration of these ideas with the objective of producing a “normally-on” robot whose internal representations are only indirectly controlled by the human programmer and the range and frequency of tasks submitted to it. We hope to construct a robot with clearly discernible preferences for engaging sensory and motor resources and an intrinsic incentive for understanding the world around it. We are encouraged by the inherently cross-modal and explicitly associative models that result from this paradigm.

Acknowledgments

This work was supported in part by the National Science Foundation under grants CISE/CDA-9703217, IRI-9704530 and IRI-9503687, and by the Air Force Research Labs, IFTD (via DARPA) under grant F30602-97-2-0032.

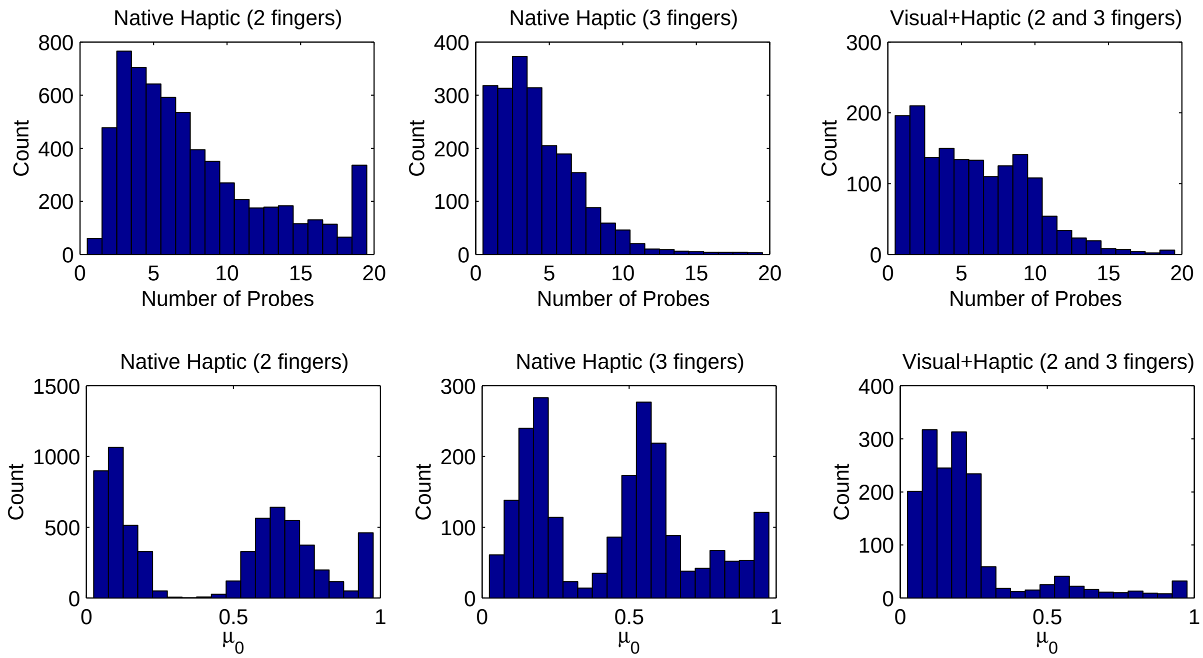


Fig. 16. Utility of the learned visual context to the haptic system when grasping rectangular and triangular prisms. The two columns on the left show the performance of the two- and three-fingered native controllers. The rightmost column shows the performance achieved if the visual system determines the initial hand orientation, and which of the two native controllers to employ.

References

1. Elizeth G. Araujo and Roderic A. Grupen. Learning control composition in a complex environment. In Pattie Maes, Maja Mataric, Jean-Arcady Meyer, Jordan Pollack, and Stewart W. Wilson, editors, *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 333–342, Cambridge, Massachusetts, September 1996. MIT Press.
2. A.E. Aronson et. al. *Clinical Examinations in Neurology*. W.B. Saunders Co., Philadelphia, PA, 1981.
3. A. G. Barto, S. J. Bradtke, and S. P. Singh. Learning to act using real-time dynamic programming. Technical Report 93-02, University of Massachusetts, Amherst, MA, 1993.
4. N. E. Berthier, R. K. Clifton, V. Gullapalli, D. McCall, and D. J. Robin. Visual information and the control of reaching. *Journal of Motor Behavior*, 28:187–197, 1996.
5. N. E. Berthier, S. P. Singh, A. G. Barto, and J. C. Houk. Distributed representations of limb motor programs in arrays of adjustable pattern generators. *Journal of Cognitive Neuroscience*, 5:56–78, 1993.
6. Christophe Biernacki, Gilles Celeux, and Gérard Govaert. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, to appear.
7. R.A. Brooks. A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14–23, March 1986.
8. Randy C. Brost. Automatic grasp planning in the presence of uncertainty. In *Proceedings 1986 IEEE Conference on Robotics and Automation*, volume 3, pages 1575–1581, 1986.
9. J.S. Bruner. Organization of early skilled action. *Child Development*, 44:1–11, 1973.
10. D. Chang and M. Cutkosky. Rolling with deformable fingertips. In *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems*, pages 194–199, 1995.
11. J. Coelho and R. Grupen. Constructing effective multifingered grasp controllers. In *Proceedings of the 1994 Conference on Robotics and Automation*, San Diego, CA, May 1994. IEEE.
12. Jefferson A. Coelho, Jr. and Roderic A. Grupen. A control basis for learning multifingered grasps. *Journal of Robotic Systems*, 14(7):545–557, 1997.
13. Jefferson A. Coelho, Jr. and Roderic A. Grupen. Learning in non-stationary conditions: A control theoretic approach. In *Proceedings of the Seventeenth International Conference on Machine Learning*. Morgan Kaufmann, 2000.
14. A. Cole, P. Hsu, and S. Sastry. Dynamic regrasping by coordinated control of sliding for a multifingered hand. In *Proceedings of the 1989 Conference on Robotics and Automation*, pages 781–786, Scottsdale, AZ, May 1989. IEEE.
15. C. Connolly. A robotics perspective on motor programs and path planning. *Behavioral and Brain Sciences*, 15(4):728–729, December 1992.

16. C. Connolly and R. Grupen. On the applications of harmonic functions to robotics. *Journal of Robotics Systems*, 10(7):931–946, 1993.
17. M.R. Cutkosky and P.K. Wright. Friction, stability and the design of robotic fingers. *Journal of Robotics Research*, 5(4), Winter 1986.
18. T. Dean and M. Wellman. *Planning and Control*. Morgan Kaufmann, 1991.
19. B. Faverjon and J. Ponce. On computing two-finger force-closure graps of curved 2d objects. In *Proceedings of the IEEE Conference on Robotics and Automation*, volume 1, pages 424–429, April 1991.
20. R.S. Fearing. Simplified grasping and manipulation with dextrous robot hands. *IEEE Journal of Robotics Research*, 2(4):188–195, January 1983.
21. C. Ferrari and J. Canny. Planning optimal grasps. In *Proc. 1992 IEEE Int. Conf. Robotics Automat.*, volume 3, pages 2290–2295, Nice, FRANCE, May 1992.
22. N. I. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1993.
23. A.M. Fraser and A. Dimitriadis. Forecasting probability densities by using hidden markov models with mixed states. In A.S. Weigend and N.A. Gershenfeld, editors, *Time Series Prediction: Forecasting the Future and Understanding the Past*, pages 265–281. Santa Fe Institute/Addison-Wesley, 1993.
24. Luiz Marcos Garcia Gonçalves, Gilson A. Giraldi, A. F. Oliveira Antonio, and Roderic A. Grupen. Learning policies for attentional control. In *Proceedings of the 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA '99)*. IEEE, 1999.
25. R. Grupen. Designing integrated sensorimotor systems. In *AAAI Spring Symposium —Toward Physical Interaction and Manipulation*, Palo Alto, CA, 1994. AAAI.
26. R. Grupen and K. Souccar. Manipulability-based spatial isotropy: A kinematic reflex. In *Workshop on Mechatronical Computer Systems for Perception and Action*, Halmstad, SWEDEN, June 1-3 1993.
27. Roderic A. Grupen, Jefferson A. Coelho, Jr., Justus H. Piater, Manfred Huber, and Michael T. Rosenstein. Structure and growth: A model of development for robot systems. Technical report, University of Massachusetts, Amherst, MA, in preparation.
28. Roderic A. Grupen, Manfred Huber, Jefferson A. Coelho, Jr., and Kamal Souccar. A basis for distributed control of manipulation tasks. *IEEE Expert*, 10(2):9–14, 1995.
29. L. Han and J.C. Trinkle. Dextrous manipulation by rolling and finger gaiting. In *Proceedings of the 1998 Conference on Robotics and Automation*, pages 730–735, 1998.
30. W. Howard and V. Kumar. On the stability of grasped objects. *IEEE Transactions of Robotics and Automation*, 12(6):904–917, 1996.
31. M. Huber and R. Grupen. Learning to coordinate controllers - reinforcement learning on a control basis. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI)*, Nagoya, JP, August 1997. IJCAII.
32. M. Huber and R. Grupen. Prior structure for on-line learning. In *Proceedings of the CIRA'97*, Monterey, CA, 1997. IEEE.
33. M. Huber and R. Grupen. A control structure for learning locomotion gaits. In *Proceedings of the Seventh International Symposium on Robotics and Applications*. TSI Press, 1998.
34. M. Huber, W. MacDonald, and R. Grupen. A control basis for multilegged walking. In *Proceedings of the Conference on Robotics and Automation*, Minneapolis, MN, April 1996. IEEE.
35. Manfred Huber and Roderic A. Grupen. A feedback control structure for on-line learning tasks. *Robots and Autonomous Systems*, 22(3/4):303–315, 1997.
36. J.L. Jones and T. Lozano-Pérez. Planning two-fingered grasps for pick-and-place operations on polyhedra. In *Proceedings of 1990 Conference on Robotics and Automation*, pages 683–688. IEEE, May 1990.
37. I. Kao and M. R. Cutkosky. Quasistatic manipulation with compliance and sliding. *International Journal of Robotics Research*, 11(1):20–40, 1992.
38. Daniel E. Koditschek. The application of total energy as a lyapunov function for mechanical control systems. In *Dynamics and Control of Multibody Systems*, volume 97 of *Contemporary Mathematics*, pages 131–157. American Mathematical Society, 1989.
39. Daniel E. Koditschek. The control of natural motion in mechanical systems. *Journal of Dynamic Systems, Measurement, and Control*, 113:547–551, 1991.
40. Willard S. MacDonald and Roderic A. Grupen. Building walking gaits for irregular terrain from basis controllers. In *Proceedings of the International Conference on Robotics and Automation*, pages 481–486, Albuquerque, New Mexico, 1997.
41. X. Markenscoff, L. Ni, and C Papadimitriou. The geometry of grasping. *International Journal of Robotics Research*, 9(1):61–74, 1990.
42. M.T. Mason and J.K. Salisbury. *Robot Hands and the Mechanics of Manipulation*. The MIT Press, Cambridge, MA, 1985.
43. Maja J Matarić. Reinforcement learning in the multi-robot domain. *Autonomous Robots*, 4(1):73–83, 1997.
44. M. E. McCarty, R. K. Clifton, and R. R. Collard. The beginnings of tool use by infants and toddlers. *submitted to Journal of Motor Behavior*, 1999.
45. M. Meila and Michael I. Jordan. Markov mixtures of experts. In R. Murray-Smith and T.A. Johansen, editors, *Multiple Model Approaches to Modelling and Control*. Taylor and Francis, 1997.
46. B.W. Mel. Combining color, shape, and texture histogramming in a neurally inspired approach to visual object recognition. *Neural Computation*, 9:77–804, 1997.

47. D. Michie and R.A. Chambers. Boxes: An experiment in adaptive control. In E. Dale and D. Michie, editors, *Machine Intelligence 2*. Edingburgh, 1968.
48. D. J. Montana. The kinematics of contact and grasp. *International Journal of Robotics Research*, 7(3):17–32, 1988.
49. H. Murase and S.K. Nayar. Visual learning and recognition of 3d objects from appearance. *International Journal of Computer Vision*, 14:5–24, 1995.
50. R. M. Murray, Z. Li, and S. S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, Boca Raton, CA, 1994.
51. Roderick Murray-Smith and Tor Arne Johansen (Eds.). *Multiple Model Approaches to Nonlinear Modelling and Control*. Taylor and Francis, London, 1997.
52. R.C. Nelson and A. Selinger. A cubist approach to object recognition. In *Proceedings of the International Conference on Computer Vision*, 1998.
53. V.D. Nguyen. The synthesis of stable grasps in the plane. In *Proceedings of the 1986 Conference on Robotics and Automation*, volume 2, pages 884–889, San Francisco, CA, April 1986. IEEE.
54. J. Piaget. *The Origins of Intelligence in Childhood*. International Universities Press, 1952.
55. J.H. Piater and R.A. Grupen. Toward learning visual discrimination strategies. In *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, Fort Collins, CO, 1999. IEEE.
56. Justus H. Piater. Learning visual features to predict hand orientations. Computer Science Technical Report 2000-40, University of Massachusetts, Amherst, MA, July 2000. Paper presented at the ICML-2000 Workshop on Machine Learning of Spatial Knowledge, in conjunction with the Seventeenth International Conference on Machine Learning, Stanford University, June 29 – July 2, 2000.
57. A.D. Pick. Improvement of visual and tactual form discrimination. *Journal of Experimental Psychology*, 69:331–339, 1965.
58. M.H. Raibert. *Legged Robots that Balance*. MIT Press, Cambridge, MA, 1986.
59. Rajesh P. N. Rao and Dana H. Ballard. An active vision architecture based on iconic representations. *Artificial Intelligence*, 78:461–505, 1995.
60. E. Rimon and D. Koditschek. The construction of analytic diffeomorphisms for exact robot navigation on star worlds. In *Proceedings of the 1989 Conference on Robotics and Automation*, volume 1, pages 21–26, Scottsdale, AZ, May 1989. IEEE.
61. A.A. Rizzi, L.L. Whitcomb, and D.E. Koditschek. Distributed real-time control of a spatial robot juggler. *IEEE Computer Magazine*, 25(5), May 1992.
62. J.K. Salisbury. *Kinematic and Force Analysis of Articulated Hands*. PhD thesis, Stanford University, May 1982.
63. Stefan Schaal and D. Sternad. Programmable pattern generators. In *International Conference on Computational Intelligence in Neuroscience (ICCIN'98)*, pages 48–51, 1998.
64. B. Schiele and J.L. Crowley. Object recognition using multidimensional receptive field histograms. In *Proceedings of the 4th European Conference on Computer Vision*, Cambridge, UK, 1996.
65. Theodore J. Schnackertz and Roderic A. Grupen. A control basis for visual servoing tasks. In *International Conference on Robotics and Automation*, Nagoya, Japan, 1995. IEEE.
66. P.G. Schyns and L. Rodet. Categorization creates functional features. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(3):681–696, 1992.
67. J.R. Silver and H.A. Rollins. The effects of visual and verbal feature-emphasis on form dicrimination in pre-school children. *Journal of Experimental Child Psychology*, 16:205–216, 1973.
68. K. Souccar and R. Grupen. Distributed motion control for multiple robotic manipulators. In *Proceedings of the Conference on Robotics and Automation*, Minneapolis, MN, April 1996. IEEE.
69. Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, Massachusetts, 1998.
70. E. Thelen. Time-scale dynamics and the development of an embodied cognition. In R.F. Port and T. van Gelder, editors, *Mind as Motion: Explorations in the dynamics of cognition*, chapter 3, pages 69–100. The MIT Press, Cambridge, MA, 1995.
71. E. Thelen and L. Smith. *A Dynamic Systems Approach to the Development of Cognition and Action*. MIT Press, Cambridge, MA, 1994.
72. Esther Thelen, Daniela Corbetta, Kathi Kamm, John P. Spencer, K. Schneider, and R. Zernicke. The transition to reaching: Mapping intention and intrinsic dynamics. *Child Development*, 64(4):1058–1098, August 1993.
73. J. C. Trinkle. On the stability and instantaneous velocity of grasped frictionless objects. *IEEE Transactions of Robotics and Automation*, 8(5), 1992.
74. J. C. Trinkle and R. P. Paul. Planning for dexterous manipulation with sliding contacts. *International Journal of Robotics Research*, 9(3):24–48, 1990.
75. F.R. Wilson. *Hand : How Its Use Shapes the Brain, Language, and Human Culture*. Vintage Books, New York, NY, first edition, 1999.
76. T. Zelniker and L. Oppenheimer. Effects of different training methods on perceptual learning in impulsive children. *Child Development*, 47:492–497, 1976.