Title
A Probabilistic Framework for Model-Based Imitation Learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 26(26)

ISSN
1069-7977

Authors
Shon, Aaron P.
Grimes, David B.
Baker, Chris L.
et al.

Publication Date
2004

Peer reviewed

A Probabilistic Framework for Model-Based Imitation Learning

Aaron P. Shon, David B. Grimes, Chris L. Baker, and **Rajesh P.N. Rao**
{aaron, grimes, clbaker, rao}@cs.washington.edu
CSE Department, Box 352350 University of Washington Seattle WA 98195 USA

Abstract

Humans and animals use imitation as a mechanism for acquiring knowledge. Recently, several algorithms and models have been proposed for imitation learning in robots and humans. However, few proposals o er a framework for imitation learning in a stochastic environment where the imitator must learn and act under real-time performance constraints. We present a probabilistic framework for imitation learning in stochastic environments with unreliable sensors. We develop Bayesian algorithms, based on Meltzo and Moore's AIM hypothesis for infant imitation, that implement the core of an imitation learning framework, and sketch basic proposals for the other components. Our algorithms are computationally efficient, allowing real-time learning and imitation in an active stereo vision robotic head. We present results of both software simulations and our algorithms running on the head, demonstrating the validity of our approach.

Imitation learning in animals and machines

Imitation is a common mechanism for transferring knowledge from a skilled agent (the *instructor*) to an unskilled agent (or *observer*) using direct demonstration rather than manipulating symbols. Various forms of imitation have been studied in apes [Visalberghy and Fragaszy, 1990, Byrne and Russon, 2003], in children (including infants only 42 minutes old) [Meltzoff and Moore, 1977, Meltzoff and Moore, 1997], and in an increasingly diverse selection of machines [Fong et al., 2002, Lungarella and Metta, 2003]. The attraction for machine learning is obvious: a machine with the ability to imitate has a drastically lower cost of reprogramming than one which requires programming by an expert. Imitative robots also offer testbeds for cognitive researchers to test computational theories, and provide modifiable agents for contingent interaction with humans in psychological experiments.

Few previous efforts have presented biologically plausible frameworks for imitation learning. Bayesian imitation learning has been proposed to accelerate Markov decision process (MDP) learning for reinforcement learning agents [Price, 2003]; however, this framework chiefly addresses the problem of learning a forward model of the environment [Jordan and Rumelhart, 1992] via imitation (see below), and the correspondence with cognitive findings in humans is unclear. Other frameworks have been proposed for imitation learning in machines [Breazeal, 1999, Scassellati, 1999, Billard and Mataric, 2000], but most of these are not designed around a coherent probabilistic formalism such as Bayesian inference. Probabilistic methods, and Bayesian inference in particular, are attractive because they handle noisy, incomplete data, can be tuned to handle realistically large problem sizes, and provide a unifying mathematical framework for reasoning and learning. Our approach is unique in combining a biologically inspired approach to imitation with a Bayesian framework for goal-directed learning. Unlike many imitation systems, which implement only software simulations, this paper demonstrates the value of our framework through both simulation results and a real-time robotic implementation.

Components of an imitation learning system

The observer must surmount a number of problems in attempting to replicate the behavior of the instructor. Although described elsewhere [Schaal et al., 2003, Rao and Meltzoff, 2003], we briefly reformulate them as follows:

1. **State identification:** Ability to classify high-dimensional sensor data into a lower-dimensional, relevant state robust to sensor noise. State identification should differentiate between the internal state of the observer (proprioceptive feedback, etc.) and the state of the environment, including the states of other agents, particularly the instructor.

2. **Action identification:** Ability to classify sequences of states in time.

3. **State mapping:** Transformation from the egocentric coordinate system of the instructor to the egocentric coordinate system of the observer.

4. **Model learning:** Learning forward and inverse models [Blakemore et al., 1998] to facilitate interaction with the environment.

5. **Policy learning:** Learning action choices that maximize a reward function, as observed from the actions selected by the instructor in each given state.

6. **Sequence learning and segmentation:** Ability to memorize sequences of key states needed to complete

an imitation task; ability to segment imitation tasks, and to divide tasks into subtasks with particular subgoal states.

A Bayesian framework for goal-directed imitation learning

Imitation learning systems that learn only state and action mappings (without modeling the environment or the instructor's goals) ignore the separability of the instructor's intent from the actions needed to accomplish that intent. Systems that use deterministic models rather than probabilistic ones ignore the stochastic nature of realistic environments. We propose a goal-directed Bayesian formalism that overcomes both of these problems. The notation s_t denotes the state (both internal and external to an agent) at time t, and a_t denotes the action taken by an agent at time t. s_G denotes a special "goal state" that is the desired end result of the imitative behavior. The key to viewing imitation learning as a model-based, goal-directed Bayesian task is to identify:

Forward model: Predicts a distribution over future states given current state(s), action(s), and goal(s)—$P(s_{t+1}|a_t, s_t, s_G)$. Models how different actions affect environmental state.

Inverse model: Infers a distribution over actions given current state(s), future state(s), and goal(s)—$P(a_t|s_t, s_{t+1}, s_G)$. Models which action(s) should be selected to transition from one environmental state to another.

Prior model: Infers a distribution over actions given current state(s) and goal(s)—$P(a_t|s_t, s_G)$. Models the policy (or preferences) followed by a particular instructor in transitioning through the environment to achieve a particular goal.

Thus the prior model involves learning an MDP (or a partially observable MDP), while the forward model involves learning a "simulator" of how the environment (possibly including other agents) reacts to actions performed within it. Learning inverse models is a notoriously difficult task [Jordan and Rumelhart, 1992], not least because multiple actions could have mapped from s_t to s_{t+1}. However, using Bayes' rule, we can infer the distribution returned by the inverse model using the forward and prior models:

$$P(a_t|s_t, s_{t+1}, s_G) \propto P(s_{t+1}|a_t, s_t, s_G)\,\Pr(a_t|s_t, s_G) \qquad (1)$$

Equation 1 can be used to either select the maximum a posteriori action to complete a state transition, or to sample over a distribution of alternatives, refining the model (and representing an exploration-exploitation tradeoff reminiscent of reinforcement learning). Sampling from the distribution over actions is also called *probability matching*. Evidence exists that the brain employs probability matching in at least some cases [Herrnstein, 1961, Krebs and Kacelnik, 1991].

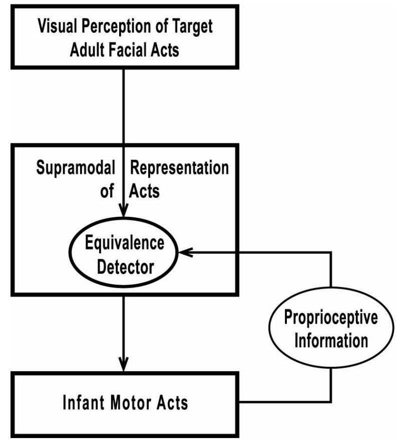

Figure 1: **AIM hypothesis model for infant imitation:** The AIM hypothesis of Meltzo and Moore [Meltzo and Moore, 1997] argues that infants match observations of adults with their own proprioceptions using a modality-independent representation of state. Our computational framework suggests an efficient, probabilistic implementation for this hypothesis.

Fig. 1 graphically represents Meltzoff and Moore's Active Intermodal Mapping (AIM) hypothesis [Meltzoff and Moore, 1997]. According to this cognitive model, imitation begins with an infant (or other agent) forming a representation of features in the outside world. Next, this representation is transformed into a "supra-modal," or modality-independent, representation of those features. An equivalence detector matches the current modality-independent representation of the instructor's state with a modality-independent representation of the infant observer's state. Proprioceptive feedback guides the infant's motor output toward matching the instructor's state. Our framework for Bayesian action selection using learned models captures this idea of imitation as a "matching-to-target" process.

Fig. 2 depicts a block diagram of our architecture. Like AIM, our system begins by running several feature detectors (skin detectors, face trackers, etc.) on sensor inputs from the environment. Detected features are monitored over time to produce state sequences. In turn, these sequences define actions. The next step is to transform state and action observations into instructor-centric values, then map from instructor-centric to observer-centric coordinates. Observer-centric values are employed to update probabilistic forward and prior models in our Bayesian inference framework. Finally, combining distributions from the forward and prior models as in Eqn. 1 yields a distribution over actions. The resulting distribution over actions is converted into a single motor action the observer should take next, with an efference copy conveyed to the feature detectors to cancel out the effects of self-motion.

State and action identification

Deriving state and action identity from sensor data involves task- and sensor-specific functions. Although it is impossible to summarize the extensive

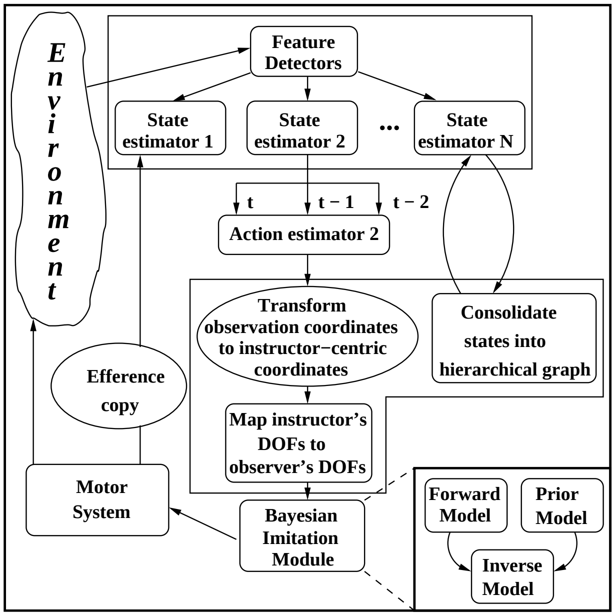

Figure 2: **Overview of model-based Bayesian imitation learning architecture:** As in AIM, the initial stages of our model correspond to the formation of a modality-independent representation of world state. Mappings from instructor-centric to observer-centric coordinates and from the instructor's motor degrees of freedom (DOFs) to the observer's motor DOFs play the role of equivalence detector in our framework, matching the instructor's motor output to the motor commands of the observer. E erence copy provides proprioceptive feedback to close the motor control loop.

body of work in action and state identification here, we note recent progress in extracting actions from laser rangefinder and radio [Fox et al., 2003] and visual [Efros et al., 2003] data. In most cases, computational expediency necessitates employment of dimensionality reduction techniques such as principal components analysis, Isomap [Tenenbaum et al., 2000], or locally linear embedding [Roweis and Saul, 2000]. Saliency detection algorithms [Itti et al., 1998] may also help reduce high-dimensional visual state data to tractable size.

Learning state mappings

A prerequisite for any robotic imitation task is to determine a mapping from the instructor's state to the observer's [Nehaniv and Dautenhahn, 2002]. We view this state mapping problem as an instance of subgraph isomorphism, where the goal is to match subgraphs from the instructor (corresponding to effectors, e.g. limbs) to their corresponding graphs in the observer. In the simulation and robotic head results shown below, the mappings are trivial; developing detailed graph-theoretic approaches to mapping from instructor states to observer states remains an ongoing topic of investigation.

Learning forward models

Numerous supervised and unsupervised approaches (see e.g. [Jordan and Rumelhart, 1992, Todorov and Ghahramani, 2003]) have been proposed to learn models of the environment, and to discover policies to maximize rewards obtained from the environment. Evidence demonstrates that infants learn forward models of how their limbs, facial muscles, and other body parts react to motor commands, a process referred to by Meltzoff and Moore [Meltzoff and Moore, 1997] as "body babbling." Such forward model learning could occur both prenatally and during infancy. We anticipate using well-established supervised algorithms to acquire forward models of environmental dynamics. Unsupervised learning of forward and inverse models to generate motor policies is a well-known problem in the reinforcement learning community (see [Kaelbling et al., 1996] for a survey). In reinforcement learning, an agent's internal reward signal alone is used to learn models of the environment, rather than relying on examples provided by a teacher as in imitation learning.

Sequence learning and segmentation

Realistic imitation learning systems must be able to learn sequences of states that define actions, and to segment these sequences into meaningful chunks for later recall or replay. Part of our ongoing work is to define how semantically meaningful chunks can be defined and recalled in real time. Recent developments in concept learning (e.g., [Tenenbaum, 1999]) suggest how similar environmental states might be grouped together, enabling development of hierarchical state and action representations in machine systems.

A Bayesian algorithm for inferring intent

Being able to determine the intention of others is a crucial requirement for any social agent, particularly an agent that learns by watching the actions of others. Recent studies have revealed the presence of "mirror neurons" in monkey cortex that fire both when an animal executes an action and when it observes others performing similar actions. These findings suggest a neurological substrate for intent inference in primates [Rizzolatti et al., 2000]. One appealing aspect of our framework is that it suggests a probabilistic algorithm for determining the intent of the instructor. That is, an observer can determine a distribution over goal states based on watching what actions the instructor executes over some period of time. This could have applications in machine learning systems that predict what goal state the user is attempting to achieve, then offer suggestions or assist in performing actions that help the user reach that state. The theory could lead to quantitative predictions for future cognitive studies to determine how humans infer intent in other intelligent agents.

Our algorithm for inferring intent uses applications of Bayes' rule to compute the probability over goal states given a current state, action, and next state obtained by the instructor, $P(s_G|s_{t+1}, a_t, s_t)$. This probability distribution over goal states represents the instructor's intent. One point of note is that $P(s_{t+1}|a_t, s_t, s_G) \equiv P(s_{t+1}|a_t, s_t)$; i.e., the forward model does not depend on the goal state s_G, since the environment is indifferent to the desired goal. Our derivation proceeds as follows:

$$P(s_{t+1}|a_t, s_t, s_G) = \frac{P(s_{t+1}, s_t, a_t, s_G)}{P(a_t, s_t, s_G)} \quad (2)$$

$$P(s_{t+1}|a_t, s_t, s_G) = \frac{P(s_G|s_{t+1}, a_t, s_t)}{P(s_G|a_t, s_t)} \frac{P(s_{t+1}, a_t, s_t)}{P(a_t, s_t)} \quad (3)$$

Because $P(s_{t+1}|a_t, s_t, s_G) \equiv P(s_{t+1}|a_t, s_t)$, and since $\frac{P(a_t,s_t)}{P(s_{t+1},a_t,s_t)} = \frac{1}{P(s_{t+1}|a_t,s_t)}$:

$$P(s_G|a_t, s_t) = P(s_G|s_{t+1}, a_t, s_t) \quad (4)$$

$$\frac{P(s_G, a_t, s_t)}{P(a_t, s_t)} = P(s_G|s_{t+1}, a_t, s_t) \quad (5)$$

$$\frac{P(a_t|s_G, s_t) P(s_G, s_t)}{P(a_t, s_t)} = P(s_G|s_{t+1}, a_t, s_t) \quad (6)$$

$$P(a_t|s_G, s_t) P(s_G, s_t) \propto P(s_G|s_{t+1}, a_t, s_t) \quad (7)$$

$$P(s_G|s_{t+1}, a_t, s_t) \propto P(a_t|s_G, s_t) P(s_t|s_G) P(s_G) \quad (8)$$

The first of the terms in Eqn. 8 represents the prior model. The second term represents a distribution over states at time t, given a goal state s_G. This could be learned by, e.g., observing the instructor manipulate an object, with a known intent, and recording how often the object is in each state. Alternatively, the observer could itself "play with" or "experiment with" the object, bearing in mind a particular goal state, and record how often each object state is observed. The third term is a prior over goal states; it can be derived by modeling the reward model of the instructor. If the observer can either assume that the instructor has a similar reward model to itself (the "like-me" hypothesis [Meltzoff, 2002]), or model the instructor's desired states in some other way, it can infer $P(s_G)$.

Interestingly, these three terms roughly match the three developmental stages laid out by Meltzoff [Meltzoff, 2002]. According to our hypothesis, the first term in Eqn. 8 corresponds to a distribution over actions as learned during imitation and goal-directed actions. This distribution can be used if all the observer wants to do is imitate body movements (the first step in imitation that infants learn to perform according to Meltzoff's theory of development). The second term in Eqn. 8 refers to distributions over states of objects given a goal state. Because the space of actions an agent's body can execute is presumably much less than the number of state configurations objects in the environment can assume, this distribution requires collecting much more data than the first. Once this second term is learned, however, it becomes easier to manipulate objects to a particular end—an observer that has learned $P(s_t|s_G)$ has learned which states of an object or situation "look right" given a particular goal. The complexity of this second term could explain why it takes babies much longer to learn to imitate goal-directed actions on objects than it does to perform simple imitation of body movements (as claimed in Meltzoff's theory). Finally, the third term, $P(s_G)$, is the most complex term to learn. This is both because the number of possible goal states s_G is huge, and the fact that the observer must model the instructor's distribution over goals indirectly (the observer obviously cannot directly access the instructor's reward model). The observer must rely on features of its own reward model, as well as telltale signs of desired states (e.g., states that the instructor tends to act to remain in, or that cause the instructor to change the context of its actions, could be potential goal states) to infer this prior distribution. The difficulty of learning this distribution could explain why it takes so long for infants to acquire the final piece of the imitation puzzle, determining the intent of others. We did not explicitly design the terms in our intent inference algorithm to match childhood developmental stages; rather, the derivation follows from the inverse model formulation in Eqn. 1 and straightforward applications of Bayes' rule.

Simulation results

Fig. 3 demonstrates imitation results in a purely simulated environment. The task is to reproduce observed trajectories through a maze containing three different goal states (maze locations marked with ovals). This simulated environment simplifies a number of the issues mentioned above: the location and value of each goal state is known by the observer a priori; the movements of the instructor are observed free from noise; the forward model is restricted so that only moves to adjacent maze locations are possible; and the observer can detect when it is next to a wall (although it does not know a priori that it cannot move through walls).

The observer first learns a forward model by interacting with the simulated environment for 500 simulation steps. The instructor then demonstrates 4 different trajectories to the observer (1 to the white goal, 2 to the light gray goal, 1 to the dark gray goal), allowing the observer to learn a prior model. Fig. 3(a) shows the maze environment used in our simulations. Fig. 3(b) shows a sample training trajectory (black arrows) where the instructor moves from location (1,1) to the goal state at (3,3). The solid white line (over arrows) demonstrates the observer reproducing the same trajectory after learning. The observer's trajectory varies somewhat from the instructor's due to the stochastic nature of the environment. Fig. 3(c) shows another training trajectory, comprising 47 steps, where the instructor moves toward the white goal (goal 1). The observer's task for this trajectory is to estimate, at each time step of the trajectory, a distribution over which goal state the instructor is headed toward. During the inference process, the observer does not have direct knowledge of the actions selected by the instructor; it must infer these by monitoring state changes in the environment. The graph in Fig. 3(d) shows this distribution over goals, where data points represent inferred intent averaged over epochs of 8 simulation steps each (i.e., the first data point on the graph represents inferred intent averaged over simulation steps 1-8, the second data point spans simulation steps 9-17, etc., with the last epoch spanning 7 simulation steps). Note that the estimate of the goal is correct over all epochs. The algorithm is particularly confident once the ambiguous section of the trajectory, where the instructor could be moving toward the dark gray or the light gray goal, is passed. Performance of the algorithm would be enhanced by more training; only 4 sample tra-

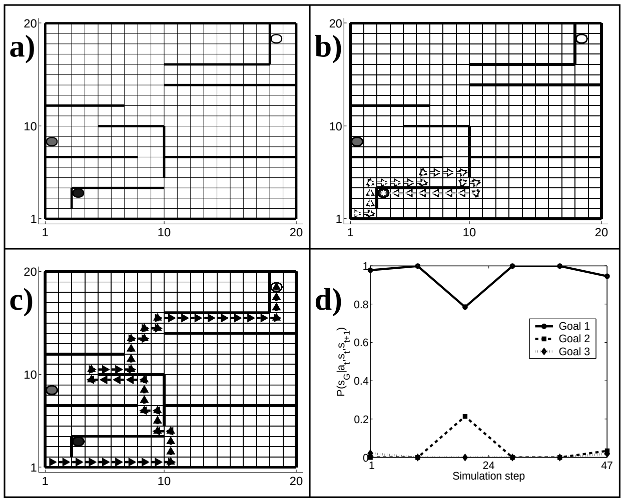

Figure 3: **Simulated environment for imitation learning:** (a) Maze environment used to train observer. Thick black lines denote walls; ovals represent goal states. Lightness of ovals is proportional to the probability of the instructor selecting each goal state (re ecting, e.g., relative reward value experienced at each state). (b) Example trajectory (black arrows) from the instructor, ending at the second goal. Reproduction of the trajectory by the observer is shown as a solid white line overlying the arrows; inference is performed as in Eqn. 1. The instructor required 23 steps to reach the goal; the observer required a slightly larger number of steps due to both the stochastic nature of the environment and imperfect learning of the forward and prior models. (c) Instructor's trajectory in the intention inference task. (d) Graph showing a distribution over instructor's goal states, as inferred by the observer at di erent time points in the simulation. Note how the actual goal state, goal 1, maintains a high probability relative to the other goal states throughout the simulation. Goal 2 brie y takes on a higher probability due to limited number of training trajectories.

jectories were presented to the algorithm, meaning that its estimates of the distributions on the right hand side of Eqn. 8 were extremely biased.

Real-time application in a robotic head

We have also implemented our probabilistic approach in a Biclops active stereo vision head (Fig. 4(a)). The head follows the gaze of a human instructor, and tracks the orientation of the instructor's head to determine where to look next. Gaze following [Brooks and Meltzoff, 2002, Scassellati, 1999] (Fig. 4(b)) represents a key step in the development of shared attention, in turn bootstrapping more complicated imitation tasks. Our system begins by identifying an image region likely to contain a face (based on detecting skin tones and bounding box aspect ratio). We employ a Bayesian pose detection algorithm [Wu et al., 2000] that matches an elliptical model of the head to the human instructor's face. Our algorithm then transforms the estimated gaze into the Biclops' egocentric coordinate frame, causing the Biclops to look toward the same point in space as the human instructor. We trained the pose detector on a total of 13 faces, with each training subject looking at 36 different targets; each target was associated with a different pan and tilt angle relative to pan 0, tilt 0 (with the subject

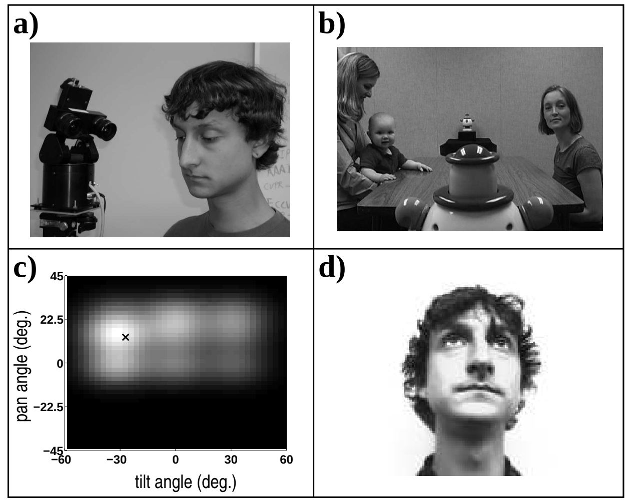

Figure 4: **Gaze tracking in a robotic head:** (a) Biclops active stereo vision head from Metrica, Inc. (b) Infants as young as 9 months can detect gaze based on head direction; older infants ($\geq$ 12 months) use opened eyes as a cue to detect whether they should perform gaze tracking (from [Brooks and Meltzo , 2002]). (c) Likelihood surface for the face shown in (d), depicting the likelihood over pan and tilt angles of the subject's head. The region of highest likelihood (the brightest region) matches the actual pan and tilt angles (black X) of the subject's face shown in (d).

looking straight ahead).

Fig. 4(c) depicts a likelihood surface over pan and tilt angles of the instructor's head in the pose shown in Fig. 4(d). Our system generates pan and tilt motor commands by selecting the maximum a posteriori estimate of the instructor's pan and tilt, and performing a simple linear transform from instructor-centric to egocentric coordinates. Out of 27 out-of-sample testing images using leave-one-out cross-validation, our system is able to track the angle of the instructor's head to a mean error of ± 4.6 degrees. [1] Our previous efforts [Shon et al., 2003] demonstrated the ability of our system to track the gaze of an instructor; ongoing robotics work involves learning policy models specific to each instructor, and inferring instructor intent based on object saliency.

Conclusion

This paper describes a Bayesian framework for imitation learning, based on the AIM model of imitation learning by Meltzoff and Moore. The framework emphasizes imitation as a "match-to-target" task, and promotes separation between the dynamics of the environment and the policy a particular teacher chooses to employ in reaching a goal. We have sketched the basic components for any imitation learning system operating in realistically large-scale environments with stochastic dynamics and noisy sensor observations. Our model naturally leads to a Bayesian algorithm for inferring the intent of other

[1] We de ne error as:

$$\mathcal{E} = \sqrt{\left(\theta_{pan} \quad \hat{\theta}_{pan}\right)^2 + \left(\theta_{tilt} \quad \hat{\theta}_{tilt}\right)^2}$$

where θ is the true angle, and $\hat{\theta}$ is our system's estimate of the angle.

agents. We presented preliminary results of applying our framework to a simulated maze task and to gaze following in an active stereo vision robotic head. We are currently investigating the ability of the framework to scale up to more complex robotic imitation tasks in real-world environments. We are also exploring the connections between our probabilistic framework and findings from developmental psychology.

Acknowledgements

We thank Andy Meltzo for extensive discussions and feedback, and for supplying Figs. 1 and 4(b). APS was funded by a NSF Career grant to RPNR, and DBG was funded by NSF grant 133592 to RPNR. We also thank the anonymous reviewers for their helpful comments.

References

[Billard and Mataric, 2000] Billard, A. and Mataric, M. J. (2000). A biologically inspired robotic model for learning by imitation. In Sierra, C., Gini, M., and Rosenschein, J. S., editors, *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 373–380, Barcelona, Catalonia, Spain. ACM Press.

[Blakemore et al., 1998] Blakemore, S. J., Goodbody, S. J., and Wolpert, D. M. (1998). Predicting the consequences of our own actions: the role of sensorimotor context estimation. *J. Neurosci.*, 18(18):7511–7518.

[Breazeal, 1999] Breazeal, C. (1999). Imitation as social exchange between humans and robots. In *Proc. AISB99*, pages 96–104.

[Brooks and Meltzo , 2002] Brooks, R. and Meltzo , A. (2002). The importance of eyes: How infants interpet adult looking behavior. *Developmental Psychology*, 38:958–966.

[Byrne and Russon, 2003] Byrne, R. W. and Russon, A. E. (2003). Learning by imitation: a hierarchical approach. *Behavioral and Brain Sciences*.

[Efros et al., 2003] Efros, A. A., Berg, A. C., Mori, G., and Malik, J. (2003). Recognizing action at a distance. In *ICCV '03*, pages 726–733.

[Fong et al., 2002] Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2002). A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3–4):142–166.

[Fox et al., 2003] Fox, D., Hightower, J., Liao, L., Schulz, D., and Borriello, G. (2003). Bayesian ltering for location estimation. *IEEE Pervasive Computing*.

[Herrnstein, 1961] Herrnstein, R. J. (1961). Relative and absolute strength of responses as a function of frequency of reinforcement. *J. Exp. Anal. Behaviour*, 4:267–272.

[Itti et al., 1998] Itti, L., Koch, C., and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE PAMI*, 20(11):1254–1259.

[Jordan and Rumelhart, 1992] Jordan, M. I. and Rumelhart, D. E. (1992). Forward models: supervised learning with a distal teacher. *Cognitive Science*, 16:307–354.

[Kaelbling et al., 1996] Kaelbling, L. P., Littman, L. M., and Moore, A. W. (1996). Reinforcement learning: A survey. *J. Artificial Intelligence Res.*, 4:237–285.

[Krebs and Kacelnik, 1991] Krebs, J. R. and Kacelnik, A. (1991). Decision making. In Krebs, J. R. and Davies, N. B., editors, *Behavioural Ecology (3rd edition)*, pages 105–137. Blackwell Scienti c Publishers.

[Lungarella and Metta, 2003] Lungarella, M. and Metta, G. (2003). Beyond gazing, pointing, and reaching: a survey of developmental robotics. In *EPIROB '03*, pages 81–89.

[Meltzo , 2002] Meltzo , A. N. (2002). Elements of a developmental theory of imitation. In Meltzo , A. N. and Prinz, W., editors, *The imitative mind: Development, evolution, and brain bases*, pages 19–41. Cambridge: Cambridge University Press.

[Meltzo and Moore, 1977] Meltzo , A. N. and Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198:75–78.

[Meltzo and Moore, 1997] Meltzo , A. N. and Moore, M. K. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6:179–192.

[Nehaniv and Dautenhahn, 2002] Nehaniv, C. and Dautenhahn, K. (2002). The correspondence problem. In *Imitation in Animals and Artifacts*. MIT Press.

[Price, 2003] Price, B. (2003). *Accelerating Reinforcement Learning with Imitation*. PhD thesis, University of British Columbia.

[Rao and Meltzo , 2003] Rao, R. P. N. and Meltzo , A. N. (2003). Imitation learning in infants and robots: Towards probabilistic computational models. In *Proc. AISB*.

[Rizzolatti et al., 2000] Rizzolatti, G., Fogassi, L., and Gallese, V. (2000). Mirror neurons: intentionality detectors? *Int. J. Psychol.*, 35:205.

[Roweis and Saul, 2000] Roweis, S. and Saul, L. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326.

[Scassellati, 1999] Scassellati, B. (1999). Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. *Lecture Notes in Computer Science*, 1562:176–195.

[Schaal et al., 2003] Schaal, S., Ijspeert, A., and Billard, A. (2003). Computational approaches to motor learning by imitation. *Phil. Trans. Royal Soc. London: Series B*, 358:537–547.

[Shon et al., 2003] Shon, A. P., Grimes, D. B., Baker, C. L., and Rao, R. P. N. (2003). Bayesian imitation learning in a robotic head. In *NIPS (demonstration track)*.

[Tenenbaum, 1999] Tenenbaum, J. (1999). Bayesian modeling of human concept learning. In Kearns, M. S., Solla, S. A., and Cohn, D. A., editors, *Advances in Neural Information Processing Systems*, volume 11. MIT Press, Cambridge, MA.

[Tenenbaum et al., 2000] Tenenbaum, J. B., de Silva, V., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323.

[Todorov and Ghahramani, 2003] Todorov, E. and Ghahramani, Z. (2003). Unsupervised learning of sensory-motor primitives. In *Proc. 25th IEEE EMB*.

[Visalberghy and Fragaszy, 1990] Visalberghy, E. and Fragaszy, D. (1990). Do monkeys ape? In *Language and intelligence in monkeys and apes: comparative developmental perspectives*, pages 247–273.

[Wu et al., 2000] Wu, Y., Toyama, K., and Huang, T. (2000). Wide-range, person- and illumination-insensitive head orientation estimation. In *AFGR00*, pages 183–188.