A Bayesian Robot That Distinguishes “Self” from “Other”

Kevin Gold (kevin.gold@yale.edu) and **Brian Scassellati (scaz@cs.yale.edu)**
Department of Computer Science, Yale University
New Haven, CT 06511 USA

Abstract

A Bayesian kinesthetic-visual matching model allows a humanoid robot to perform mirror self-recognition without social understanding. The robot learns the relationship between its own motor activity and perceived motion by observing the movements of its arm for four minutes. Over this time, it builds a simple dynamic Bayesian model that relates these two events, and uses this self model to build a model of “animate others” that is a copy of its self model but with the motor state hidden. Presented with a mirror, the robot then judges its mirror image to match its “self” model, while people are judged to be “animate others.”

Keywords: Self-recognition; robot; mirror test; Bayesian; animacy; contingency.

Introduction

There are two kinds of theories that seek to explain how 2-year-old human infants, higher primates, and other intelligent animals discover that their mirror images are their own reflections. The first kind of theory proposes that mirror self-recognition requires social understanding, as the learner must have some idea of how it appears to others (G. Gallup, 1982). The second kind of theory proposes that mirror self-recognition comes from matching kinesthetic experience to visual feedback, and that therefore no social knowledge is necessary for self-recognition (Mitchell, 1997). The present study is meant to add some concreteness to the latter hypothesis of kinesthetic-visual matching, and show that it is computationally feasible and robust to identify one's mirror image using only a Bayesian self-model that relates motor activity to motion.

The “mirror test” of G. G. Gallup (1970) is typical of tests for mirror self-recognition. In the original test, a chimpanzee was given ten days to acclimate itself to a mirror in its cage. A spot of red dye was then applied to a chimpanzee's eyebrow and ear while it was unconscious. On seeing its reflection in the mirror, the chimpanzee would use its mirror reflection to reach for the marks on its own head. Similar tests have found mirror self-recognition in humans at 2 years (Amsterdam, 1972), orangutans (Suarez & Gallup, 1981), elephants (Plotnik, Waal, & Reiss, 2006), and dolphins (Reiss & Marino, 2001), while macaques, gorillas, and lesser primates appear to be unable to self-recognize (G. Gallup, 1982).

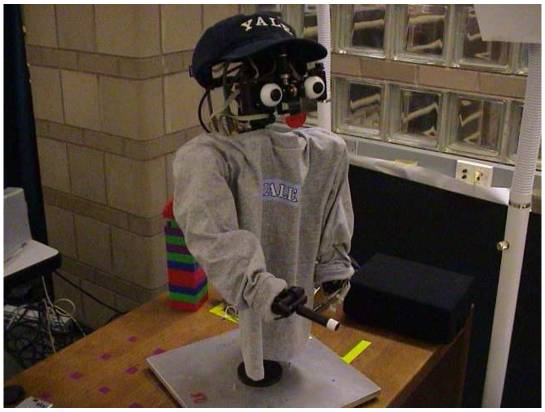

Figure 1: Nico is an upper-torso humanoid robot with the arm and head kinematics of a one-year-old.

The ability to recognize oneself in the mirror appears in human infants within a few months of some other critical social abilities, such the development of sympathy at others' distress (Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). The emergence of this ability at about the same time in human infants has been interpreted by some to be more than a coincidence, suggesting that “theory of mind” skills are necessary for mirror self-recognition (G. Gallup, 1982; Plotnik et al., 2006). However, it is difficult to determine from phylogenetic and developmental observations alone whether this link is more than coincidental.

The present work demonstrates that it is possible to perform robust self-recognition on the basis of matching kinesthetic experience to visual motion alone. The method consists of the robot comparing three models for each object in its visual field. The first model is that of random noise, generated with no structure over time. The second model consists of an observed internal state of motor activity that generates the external feedback of motion; thus, the consistency of the match between motor activity and motion dictates the likelihood of this model. The third model is that of motion generated by somebody else; it is identical to its own self-motion model, only the motor state is hidden and must be reasoned about probabilistically. The likelihood of each model is updated with every new observation, such that the robot always has a best guess as to whether

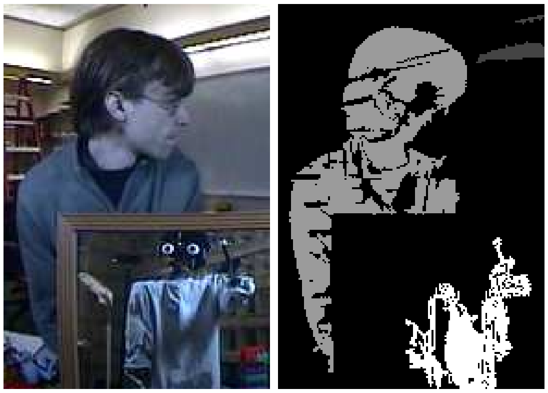

Figure 2: Output from the self/other algorithm while the robot views the experimenter and the robot's mirror image. The algorithm classifies objects found through background subtraction as "self" (white, bottom), "animate other" (light gray, left), or "inanimate" (dark gray, upper right) in real time.

an object in its visual field is itself, someone else, or an inanimate object.

Using techniques from Bayesian reasoning, these models can be computed in real-time, requiring only a few calculations between observations. In the case of the self-model and the noise model, the models can even perform unsupervised learning in real-time with no iteration over the robot's observation history. Finally, the "Animate Other" model, for motion generated by other people in the vicinity, presents an interesting case because the robot can use its own self-model as an approximation, thus allowing the whole learning process to occur online in real-time.

The robot learns a basic self-model by observing its own actions. Presented with a mirror, the robot can then use its learned self-model to determine the likelihood that the motion of its mirror image is its own motion. After a few movements, the probability that the mirror arm is "self" goes to nearly 1, while the other probabilities become insignificant.

Note that while our research group has previously published a different method for motion-based self-recognition (Gold & Scassellati, 2006), the Bayesian models described below are new to this paper, and are considerably less susceptible to error when a human moves at the same time as the robot. Another group has published work on learning a slightly more complex forward model than the one described here using Bayesian methods (Dearden & Demiris, 2005), but that group did not adapt the method to the problem of judging whether an entity was the self or not.

Mathematical Background and Models

Our method compares three models for every object in the robot's visual field to determine whether it is the robot itself, someone else, or neither. The use of Bayesian networks allows the robot to calculate at each time t the likelihoods λ_t^ν, λ_t^σ, and λ_t^ω, corresponding to the likelihoods of the evidence given the noise model, the self model, and the "animate other" model, respectively. Normalizing these likelihoods then gives the probability that each model is correct, given the evidence. We shall first discuss how the models calculate their probabilities under fixed parameters, then explain how the parameters themselves are adjusted in real-time.

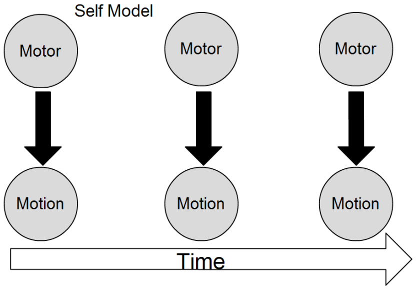

Figure 3: The robot's self model, in graphical model notation. Darkened circles represent observations, and arrows represent conditional dependence relations.

The "inanimate" model is the simplest, as we assume inanimate objects only appear to have motion due to sensor noise or when they are dropped. If we characterize the occurrence of either of these events as the event r, then this model is characterized by a single parameter: the probability $P(r)$ that random motion is detected at an arbitrary time t. Observations of this kind of motion over time are assumed to be independent, such that the overall likelihood λ_t^ν can be calculated by simply multiplying the likelihoods at each time step of the observed motion.

The robot's second model for an object is the "self" model, in which the motor actions of the robot generate the object's observed motion. The model is characterized by two probabilities: the conditional probability $P(m|\phi)$ of observing motion given that the robot's motors are moving, and the conditional probability $P(m|\neg\phi)$ of observing motion given that the robot's motors are not moving. (Henceforth, m and $\neg m$ shall be the observations of motion or not for motion event M, and ϕ and $\neg\phi$ shall serve similarly for motor event Φ. Note that these probabilities need not sum to 1.)

Figure 3 represents the robot's "self" model graphically, using standard graphical model conventions. Each circle corresponds to an observation of either the robot's own motor action (top circles) or the observed motion of the object in question (bottom circles), with time t increasing from left to right. The circles are all shaded to indicate that these event outcomes are all known to the robot. The arrows depict conditional dependence; informally, this corresponds to a notion of causality. Thus, the robot's motor action at time t causes the perception of motion at time t. Though an observer would know that the robot's motor activity at time $t+1$ can be predicted from its activity at time t, the robot does not "know" this in its model; it calculates only the likelihood of the motion evidence given the motor evidence, and not the likelihood of the motor evidence itself.

To determine the likelihood of this model for a given object, the robot must calculate the probability that its sequence of motor actions would generate the observed motion for the object. The relevant calculation at each time step is the probability $P(M_t|\Phi_t)$ of motor event Φ_t generating motion observation M_t. These probabilities calculated at each time step can then be simply multiplied together to get the overall likelihood of the evidence, because the motion observations are conditionally independent given the robot's motor actions.

The likelihood λ_t^σ that the motion evidence up to time t was generated by the robot's own motors is then:

$$\lambda_t^\sigma = \prod_t P(M_t|\Phi_t) \tag{1}$$

where, in our simple Boolean implementation,

$$P(M_t|\Phi_t) = \begin{cases} P(m_t|\phi_t) & \text{if } m_t \text{ and } \phi_t \\ 1 - P(m_t|\phi_t) & \text{if } \neg m_t \text{ and } \phi_t \\ P(m_t|\neg\phi_t) & \text{if } m_t \text{ and } \neg\phi_t \\ 1 - P(m_t|\neg\phi_t) & \text{if } \neg m_t \text{ and } \neg\phi_t \end{cases} \tag{2}$$

Under this model, updating the likelihood at time $t+1$ is simply a matter of multiplying by the correct value of $P(M_{t+1}|\Phi_{t+1})$:

$$\lambda_{t+1}^\sigma = P(M_{t+1}|\Phi_{t+1})\lambda_t^\sigma \tag{3}$$

Note that equation 1 and the graphical model presented in Figure 3 are much more general than the simple Boolean model implementing them that is described by equation 2. For more advanced models, M_t could be a complete reading of joint angles, Φ_t could describe a trajectory through space, and $P(M_t|\Phi_t)$ could be an arbitrary distribution on motion trajectories given the motor readings. The current implementation, however, chooses simplicity over expressive power.

The third and final model is that of another person (or other animate agent) in the visual field. This model is identical to the self model, but now the motor states are hidden to the robot, leaving it to infer the other person's motor states. Removing the motor information turns the model into a Hidden Markov Model (HMM). The Bayesian network in Figure 4 represents this by leaving the motor event nodes unshaded, indicating that they have not been directly observed. To assess the likelihood that motion was generated by another animate agent, the robot must now infer the underlying motor states that generated the motion. Performing this calculation requires two transition probabilities, $P(\phi_{t+1}|\neg\phi_t)$ and $P(\neg\phi_{t+1}|\phi_t)$, corresponding to the probabilities that the person begins motor activity from rest or ceases its motor activity, respectively.

To find the likelihood that an object is an animate other, the robot must keep track of the probabilities of each motor state at each time step. This can be updated online in a constant amount of time at each time step using the forward algorithm (Rabiner, 1989). The forward algorithm can be described by the following equation for calculating the new likelihood λ_{t+1}^ω given an observation of motion M_{t+1} and motor state probabilities $\overrightarrow{P(\Phi)}$:

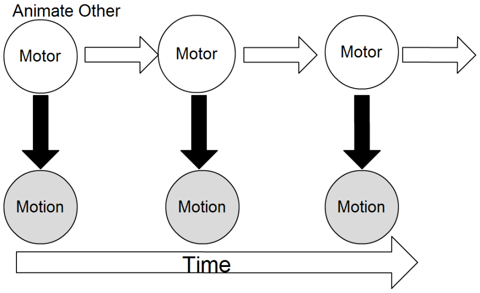

Figure 4: The model for an "animate other," in graphical model notation. The model the robot uses is a copy of its self model, but with the motor information unobserved and inferred (unshaded circles).

$$\lambda_{t+1}^\omega = \sum_{\Phi_{t+1}} P(M_{t+1}|\Phi_{t+1}) \sum_{\Phi_t} P(\Phi_{t+1}|\Phi_t)P(\Phi_t) \tag{4}$$

Notice that this equation necessarily entails calculating probabilities for the entity's hidden state as a subroutine. In our simple implementation, this only decides whether the agent is engaging in motor activity or not. But, as with the self model, equation 4 and figure 4 are more general than the simple boolean model we are using here. A more complex model relating the entity's motor states to its motion would allow action recognition as a pleasant side effect of the likelihood calculation.

The forward algorithm requires prior probabilities on the possible motor states to propagate forward. Since the robot has no particular information about the state of the "other" at time 0, it arbitrarily sets these to 0.5. However, both the "self" and "other" models are more complex and more rare than the assumption of noise, so the final likelihoods are weighted by the priors of $P(\text{inanimate}) = 0.8$, $P(\text{self}) = 0.1$, $P(\text{other}) = 0.1$. As constants, these priors do not matter much in the long run, but they can reduce false classifications when first encountering an object.

The other parameters to the model, the conditional probabilities, do matter in the long term; luckily, they can be learned online rather than set arbitrarily. For the "inanimate" and "self" models, the robot does this by counting its observations of each event type ($\{m, \neg m\} \times \{\phi, \neg\phi\}$) and weighting each observation by the probability that the object truly belongs to the model for which the parameters are being calculated. Thus:

$$P(r) = \frac{\sum_{it} P_{it}(\text{noise})M_{it}}{\sum_{it} P_{it}(\text{noise})} \tag{5}$$

$$P(m|\phi) = \frac{\sum_{it} P_{it}(\text{self})M_{it}\Phi_t}{\sum_{it} P_{it}(\text{self})\Phi_t} \tag{6}$$

$$P(m|\neg\phi) = \frac{\sum_{it} P_{it}(\text{self}) M_{it}(1-\Phi_t)}{\sum_{it} P_{it}(\text{self})(1-\Phi_t)} \quad (7)$$

where $P_{it}(\text{noise})$ is the robot's best estimate at time t of the probability that object i is noise, and M_{it} is 0 or 1 depending on whether object i is moving. This strategy is a kind of expectation maximization, because it alternates between fitting a model to data and classifying the data with a model. Normally, expectation maximization requires iterating over all previous observations in updating the model, but this model is simple enough that the robot can update it in real time without going back to revise its previous probability estimates, without too much loss of accuracy.

Since this method is iterative, the robot must begin with some estimate of each of the probabilities. The robot begins with a guess for each parameter as well as a small number of "virtual" observations to support that guess. These guesses function as priors on the parameters, as opposed to the priors on classifications described earlier. The choice of priors here does not matter much, since the system can adapt to even bad priors (see "Experiments," below). Any prior will smooth the model's development of the correct parameters, by reducing its reliance on its first few observations.

Using the expectation maximization strategy on the "animate other" model would not work quite as well, because it contains unobserved states. Technically, to perform expectation maximization on a Hidden Markov Model requires the forward-backward algorithm (Baum & Petrie, 1966) to obtain *a posteriori* estimates of the hidden states, which would require iterating over all the data repeatedly as the robot gained more data. However, we can finesse this problem, reduce the size of the hypothesis space, and prove an interesting point about self-models all at the same time if the robot *uses its own self model to generate the "animate other" model.* The probabilities $P(m|\phi)$ and $P(m|\neg\phi)$ are set to exactly the same values as the self model; this is equivalent to the assumption that the robot has about the same chance of perceiving motion if either itself or someone else is actually moving. The transitional probabilities $P(\phi_{t+1}|\neg\phi_t)$ and $P(\neg\phi_{t+1}|\phi_t)$ are based on the robot's own motor activity by counting its own action transitions of each type. Though the human's motions are likely to be quite different from those of the robot in their particulars, the general fact that "objects in motion tend to stay in motion" is true of both, and the real discrimination between the two hinges on the contingency with the robot's own motor actions, and not the particular transition probabilities of the "animate other" model.

Robotic Implementation

The experiments described below were performed on Nico, a small humanoid robot built to match the proportions and kinematics of a one-year-old child (Figure 1). Nico possesses an arm with 6 degrees of freedom, corresponding to the degrees of freedom of a human arm up to the wrist. The arm made sweeping gestures roughly 1 second in length to a randomly chosen position roughly every five seconds. Feedback from the motors in the form of optical encoder readings indicated to the robot whether each motor had stopped moving.

For vision, Nico used 320 × 240 images pulled from the wide-angle CCD camera in Nico's right eye at roughly 30 frames per second. Images from Nico's camera were then passed through a background subtraction filter, leaving only objects that had moved since the experiment began and scattered noise. Connected regions that did not exceed 100 pixels (roughly 0.1% of the image) were discarded as noise.

Objects were tracked over time by matching each region R_i in frame F with the region in frame $F-1$ that shared the largest number of pixels with R_i. If more than one connected region in the same frame attempted to claim the same object identity, as frequently happened when joined regions separated, a new identity was generated for the smaller region. An object with area A was judged to be moving if $4\sqrt{A}$ of its pixels had changed their region label from one frame to the next. This formula was chosen to be roughly proportional to the length of the object's perimeter, while taking into account that background subtraction tended to produce "fuzzy" borders that are constantly changing.

The final output of vision processing was an image of labeled regions that could be tracked over time and judged at each time step to be moving or not moving. This output was made available at a rate of roughly 9 frames per second. For each segmented region, the probabilities of the three models described above were calculated and updated in real time using the algorithms described earlier. Figure 2 shows output typical of the self-other algorithm after learning, with image regions grayscale-coded by maximum likelihood classification. Note that because the background subtraction algorithm blinds the robot to objects that have not moved since the start of the experiment, the robot cannot actually classify its body, but only its arm. A segmentation algorithm based on depth would join the arm to the full body and classify the whole assembly based on the movement of the arm, but this was not implemented.

Experiments

Methodology

The robot was given 4 minutes to observe the movements of its own arm, starting with $P(r)$, $P(m|\phi)$, and $P(m|\neg\phi)$ all set to the implausible value of 0.5. These starting values were given the weight of 30 observations, or roughly 3 seconds of data. The robot made its observations in the absence of a mirror and without any explicit feedback. Distractors from the arm included both inanimate objects (light fixtures that background subtraction had failed to remove) and animate others (students passing in the hall adjacent to the lab). To automatically collect data on the robot's hypotheses about its arm without hand-labeling each frame, the probabilities generated for the largest object within its field of view were recorded as data; this object was the arm in most instances.

The robot's parameters were then frozen at the four minute mark for testing, to ensure the robot's perfor-

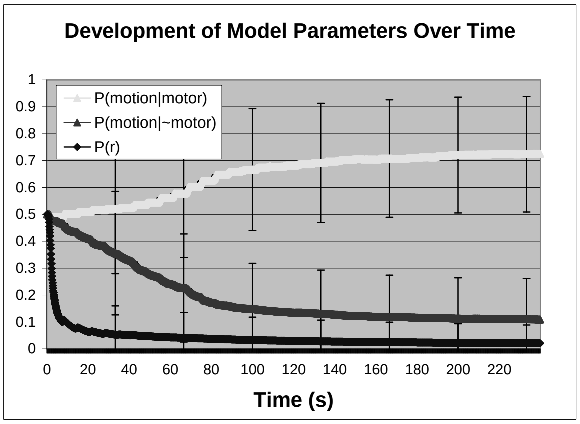

Figure 5: Average model parameters over four minutes of unsupervised learning on the robot's visual and motor feedback, with 95% confidence intervals.

mance was based solely on its observations of its own unreflected arm. Using the parameters it learned during the previous four minutes, the robot then was presented with a mirror, and the robot continued its random movements in front of the mirror. The robot's hypotheses for the largest object within the area covered by the mirror were recorded automatically, to avoid the chore of hand-labeling frames.

Using the same parameters, the robot then judged one of the authors (K.G.) for two minutes. Again, the robot's hypotheses for the largest object located within the area of interest were recorded automatically. The author's actions varied from near inactivity (sitting and checking his watch) to infrequent motion (drinking bottled water) to constant motion (juggling), while the robot continued to make periodic movements every 5 seconds. The experimenter was blind to the robot's classifications during this time.

These experiments were repeated 20 times, to verify the robustness of both the learning mechanism and the classification schemes that it generated. Each learning trial reset all model probabilities to 0.5, and each test trial used the parameters generated in the corresponding learning trial.

Results

Within four minutes of learning, the robot's model probabilities consistently changed from 0.5 to roughly $P(r) = 0.020$, $P(m|\phi) = 0.73$, and $P(m|\neg\phi) = 0.11$. Figure 5 shows the mean values over time for these parameters, along with 95% confidence intervals calculated using Student's t for 19 degrees of freedom. The robustness of the model even under such terrible starting conditions was quite surprising; one would expect that the models would require some initial notion that motion was more likely with motor activity, but this fact was entirely learned.

This learning also occurred in the presence of many tracking failures, caused by the failure of background subtraction to correctly segment the image or the arm passing out of the field of view. Over each trial, the arm passed out of the field of view or was incorrectly segmented many times, only to have the "new" object quickly reclassified correctly from scratch.

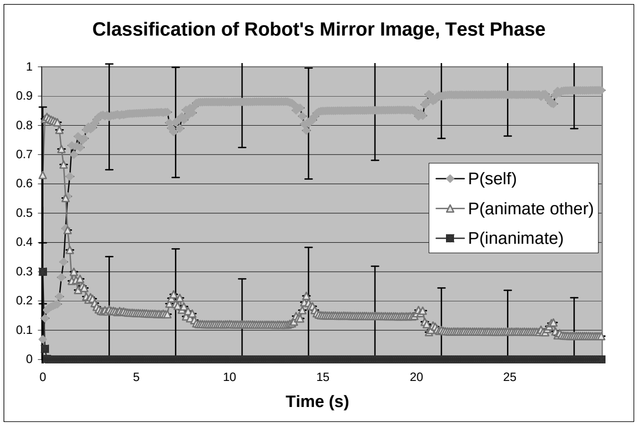

Figure 6: Robot's judgements of its mirror image after 4 minutes of observing its own unreflected movements, with 95% confidence intervals.

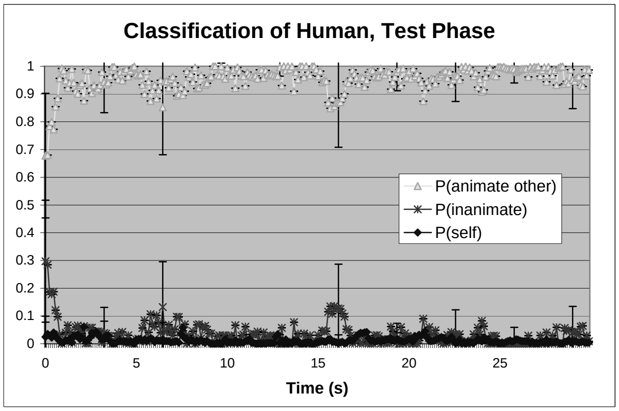

Figure 7: Robot's judgments of the experimenter after 4 minutes of observing its own unreflected movements, with 95% confidence intervals.

When confronted with a mirror after the learning phase, the robot consistently judged the mirror image to be "self" after a single complete motion of its arm, and remained confident in that estimate over time. Figure 6 shows the robot's estimates of its mirror image over just the first 30 seconds, so as to better highlight the rapid change in its hypothesis during its first movement. The periodic dips in its confidence were not significant, but were probably caused by the slight lag between the robot sensing its motor feedback and seeing the visual feedback, as its motors needed time to accelerate to a detectable speed.

The robot's judgments of the experimenter as "animate other" were similarly quick, and its confidence remained high throughout every test trial. Figure 7 again shows only the first 30 seconds, to better highlight the changes in the first two seconds. The data for the next minute and a half was quite similar.

In short, the robot correctly classified both its mirror image and the experimenter quickly, persistently, and in all trials, using only the parameters it learned from watching its unreflected arm for four minutes.

What Does the Mirror Test Prove?

The experiments here illustrate that social understanding is not necessarily a prerequisite for mirror self-recognition. This would tend to lend support to the arguments of Mitchell (1997) that kinesthetic-visual matching is on the whole a more coherent theory than social explanations of the mirror test. In other words, the mirror test may not be about "self-awareness" or "theory of mind" at all; it may merely be a test of an organism's ability to adapt to new kinds of visual feedback.

Granted, the robot here did not go through a complete "mirror test." The traditional mirror test requires a mapping from a location on the mirror image to one on the real body, since the organism must reach for its own face in response to seeing the spot of rouge. Our system merely classifies the physical arm and the mirror arm as two instances of the same "self" category. In addition, the original mirror test requires that the organism learn about its appearance, so that it can detect a change when the rouge is applied.

However, these do not seem to be terribly difficult aspects to the test. Learning the appearance of the self is simply a matter of applying machine learning to the image regions identified as "self." Moreover, even pigeons can learn (with training) to use mirrors to find blue dots on their bodies that are otherwise not visible to them (Epstein, Lanza, & Skinner, 1981), so it would be unwise to put too much emphasis on the physical mapping aspect of the test. The hardest part of replicating human understanding of the mirror image is probably that humans can understand the mirror image as an illusion or representation of the self; but it is not clear to what degree any other species understand this.

It interesting that Bayesian self-recognition systems can be used for action recognition, and vice versa – particularly because of the recent interest in "mirror neurons" that can identify either self-generated actions or others' actions (Rizzolatti, Fogassi, & Gallese, 2001). The presence of mirror neurons alone cannot explain mirror test performance, since monkeys possess such neurons but do not in fact pass the mirror test. Still, mirror neurons may be necessary but not sufficient, or higher primates may possess more sophisticated versions of the same systems.

Though our simple robotic model suggests that it is *possible* to perform mirror self-recognition using only the presence or absence of motor activity and motion, it does not match well with the time scale of self-recognition in primates. Our model took only 4 minutes to learn, but mirror self-recognition takes nearly 2 years to develop in humans (Amsterdam, 1972), and adult chimpanzees typically require a few days with a mirror before they cease to direct social behaviors toward it (G. G. Gallup, 1970). One possible reason for this discrepancy is that these species use more complicated models relating their motor activity to their motion. A larger parameter space would require more time to learn, and a more complicated model may also produce spatial location expectations that are violated by the mirror image. It may be the case that some species fail the mirror test simply because their short developmental periods do not leave time to train the kind of self-recognition system that is complex enough to provide spatial expectations, but flexible enough to recognize self-generated motion in a mirror.

Acknowledgments

Support for this work was provided by a National Science Foundation CAREER award (#0238334) and award #0534610 (Quantitative Measures of Social Response in Autism). Some parts of the architecture used in this work was constructed under NSF grants #0205542 (ITR: A Framework for Rapid Development of Reliable Robotics Software) and #0209122 (ITR: Dance, a Programming Language for the Control of Humanoid Robots) and from the DARPA CALO/SRI project. This research was supported in part by a grant of computer software from QNX Software Systems Ltd.

References

Amsterdam, B. (1972). Mirror self-image reactions before age two. *Developmental Psychobiology*, *5*(4).

Baum, L. E., & Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *Annals of Mathematical Statistics*, *41*.

Dearden, A., & Demiris, Y. (2005). Learning forward models for robots. In *Proc. ijcai* (pp. 1440–1445). Edinburgh, Scotland.

Epstein, R., Lanza, R. P., & Skinner, B. F. (1981). "Self-awareness" in the pigeon. *Science*, *212*, 695–696.

Gallup, G. (1982). Self-awareness and the emergence of mind in primates. *American Journal of Primatology*, *2*, 237–248.

Gallup, G. G. (1970). Chimpanzees: self-recognition. *Science*, *167*(3914), 86-87.

Gold, K., & Scassellati, B. (2006). Learning acceptable windows of contingency. *Connection Science*, *18*(2), 217–228.

Mitchell, R. W. (1997). Kinesthetic-visual matching and the self-concept as explanations of mirror-self-recognition. *Journal for the Theory of Social Behavior*, *27*(1).

Plotnik, J. M., Waal, F. B. M. de, & Reiss, D. (2006). Self-recognition in an asian elephant. *PNAS*, *103*(45).

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–296.

Reiss, D., & Marino, L. (2001). Mirror self-recognition in the bottlenose dolphin: A case of cognitive convergence. *Proceedings of the National Academy of Sciences*, *98*(10).

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, *2*, 661–670.

Suarez, S., & Gallup, G. G. (1981). Self-recognition in chimpanzees and orangutans, but not gorillas. *Journal of Human Evolution*, *10*, 157–188.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28*(1), 126–136.