

Robot Self/Other Distinction: Active Inference Meets Neural Networks Learning in a Mirror

Pablo Lanillos¹ and Jordi Pages² and Gordon Cheng³

Abstract. Self/other distinction and self-recognition are important skills for interacting with the world, as it allows humans to differentiate own actions from others and be self-aware. However, only a selected group of animals, mainly high order mammals such as humans, has passed the mirror test, a behavioural experiment proposed to assess self-recognition abilities. In this paper, we describe self-recognition as a process that is built on top of body perception unconscious mechanisms. We present an algorithm that enables a robot to perform non-appearance self-recognition on a mirror and distinguish its simple actions from other entities, by answering the following question: am I generating these sensations? The algorithm combines active inference, a theoretical model of perception and action in the brain, with neural network learning. The robot learns the relation between its actions and its body with the effect produced in the visual field and its body sensors. The prediction error generated between the models and the real observations during the interaction is used to infer the body configuration through free energy minimization and to accumulate evidence for recognizing its body. Experimental results on a humanoid robot show the reliability of the algorithm for different initial conditions, such as mirror recognition in any perspective, robot-robot distinction and human-robot differentiation.

1 Introduction

“Our bodies are our gardens, to the which our wills are gardeners.” Othello 1603. William Shakespeare.

Humans perceive their body in a flexible manner [5] but maintaining their identity [12]. Around the age of two, the majority of infants are able to pass the *self-recognition mirror test* [31], like elephants, dolphins and some non-human primates [2]. This experiment has been argued to be hardly intertwined with self-awareness and social understanding [14] and thus consciousness. Recently, a cleaner wrasse fish reopened the debate about the suitability of this test for evaluating self-awareness [20]. As reported by the authors, it was the first fish to pass the mirror test, waving the notion of consciousness by suggesting different evolutionary paths and showing a broad spectrum of self-awareness.

Self-recognition is commonly understood as a reflective process, i.e. you are aware and able to think about it. However, it can be treated in two different levels [35]: the sensorimotor [25] and the cognitive [2]. Regarding the first level, we hypothesize that the mirror test can be partially passed without being self-conscious, open-

ing the door for artificial agents to self-recognize and to acquire the skill of self/other distinction. Actually, underneath self-recognition, we need body perception and action processes, and the first step to achieve it is contingency-learning [23]. According to [4], humans can learn and predict the consequence of actions and these forward models are used to determine the source of the sensory events. Thus, in this work we extrapolated this concept to robots. The agent should be able to answer the question “is this my body?” by answering “am I generating those effects in the world?”.

We present an algorithm that enabled a robot with non-appearance self-recognition capabilities, by addressing the relation between forward models prediction error and its use for determining the source of the sensory cues. In other words, the artificial agent can self-recognize through accumulating evidence of producing effects in the world by self-generated actions. For that purpose, the robot needs to learn the expected sensory effect of: the changes in the visual field when it moves the body and where its body (e.g. hand) will appear in the visual field when it is in front of the mirror.

Previous works on artificial self-recognition, such as [14, 23, 34], have been hardly criticized [1]: “Merely simulating certain features of self-recognition through training/programming does not mean that the underlying mechanisms are the same, similar, or even remotely related”. Actually, we agree that experimental simplifications make debatable that robots are currently able to self-recognize in the same way that chimpanzees do. Therefore, instead of addressing the self-perception awareness level, we focused on body sensorimotor learning and integration and studying how these mechanisms aid in self-recognition and self/other distinction.

Relation to other approaches

Artificial intelligence approaches for self-recognition are quite far from the biological plausibility [13, 24]. However, some of them are already addressing relevant characteristics of self-perception such as timing, spatial correlations and contingency.

On the one hand, we have methods that do not involve learning. In [14] they used Bayesian inference to accumulate evidence. However, action and perception had to be binary variables and required continuous stable segmentation of the body as an entity, complicating its scalability to real scenarios. In [23] this approach was extended to images and its relation with proprioception. The method depended on the skin accelerometers and the image grid subsampling. On the other hand, contingency is possible to be learnt with spiking neural networks [29] and self/other distinction can be developed by gradually increasing spatiotemporal resolution [26].

Our approach involves both inference and learning and casts self/other distinction and self-recognition to free energy minimization (i.e. variational inference), in order to solve the scalability and

¹ Accepted at European Conference on Artificial Intelligence (ECAI 2020)

¹ Donders Institute for Brain, Cognition and Behaviour, the Netherlands.

² PAL robotics, Pujades 77-79, 4-4 08005 Barcelona, Spain.

³ Institute for Cognitive Systems, Technical University of Munich, Arcisstrasse 21, 80333 Munich, Germany.

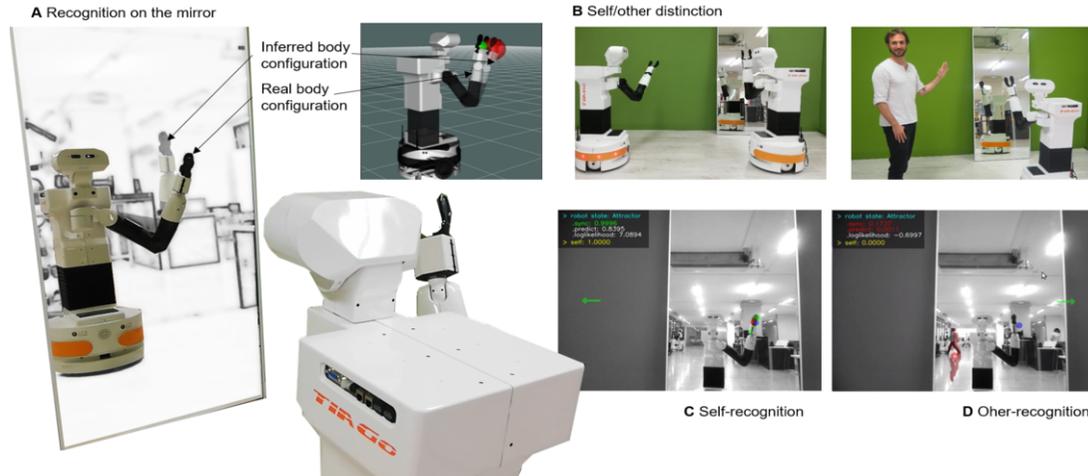


Figure 1. Self-recognition in the mirror and self/other distinction experiments

partial information sensitivity limitations of previous approaches.

Predictive coding approach

Under the predictive coding [30] approach, the robot is able to learn (approximate) generative models of the world, especially the effects of the body in the world. These prediction errors are used for inferring the body state but they are also an indirect measure to distinguish the robot generated actions.

Defining the brain as an inference machine [19] that is able to interpret the world from partial information [9, 15], body perception is presented here as an unconscious process that adapts empirical information to current learned models. Based on the *active inference* paradigm [11, 27], we formalized body perception and action as an optimization problem that minimizes the prediction error, i.e., the error between the expected sensory effect and the observed one [10]. According to active inference, adaptation can be performed by changing the interpretation or belief or by acting to modify the world according to our expectations. Thus, we deployed this adaptation mechanism on a humanoid robot and studied the perceptual response regarding self-recognition when executing movements in front of a mirror or in front of other agents.

Contribution and organization

This work presents a novel active inference computational model that combines free energy optimization from [27] with function learning from [22]. The learning was performed by means of a Mixed Density Networks, MDN [3], to provide scalability and interpolation. The algorithm enables a robot to recognize itself in the mirror using non-appearance body cues and to distinguish its body from other entities. Our approach, as it incorporates learning, is robust to variations in the robot position and the mirror location and angle. On top of this model, we were able to compute the probability that our robot with a given set of parameters would generate a particular observed sensor value. The accumulation of this evidence yielded to self-recognize itself in the mirror and perform self/other distinction.

First, Section 2 describes the experimental setup and the computational model. Section 3 shows the results and Section 4 analyze the relevance of the results with respect to robotics and artificial intelligence, and discusses its limitations regarding animal self-recognition.

2 Methods

2.1 Experimental design

Figure 1 shows the humanoid robot TIAGo [28] and the three different scenarios tested: in front of the mirror, robot-robot interaction and human-robot differentiation. TIAGo is a mobile manipulator by PAL Robotics that has been designed in a modular way. It is composed of a differential drive base; a lifting torso; a 7 DoF arm and a pantilt head. The robot has a laser rangefinder on the base and a RGBD camera on the head. For the experiments, we only used the monocular RGB camera as 3D computations are distorted in the mirror setting. In terms of manipulation capabilities, it has a 7 DoF arm and the lifting torso prismatic joint, being able to reach objects on the floor up to objects on shelves at around 1.75 m. For the self/other distinction experiment, only the arm will be used. Moreover, the robot has stereo microphones speakers and professional text-to-speech software. Figure 2 shows the main modular components of the robot used.

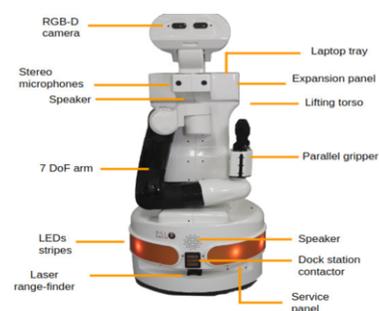


Figure 2. TIAGo robot used for the experiments

2.2 Overall algorithm

The algorithm 1 summarizes the whole process computed in the robot for self-recognition and self/other differentiation. First, the visual input of the system is segmented using optical flow to extract the visual 2D centroid of any visual blob and the histogram of visual movement directions. This data will be used as the visual input s_v and for computing the probability of being a contingent event respectively. The predicted sensor values are computed by means of an MDN and its partial derivatives with respect to the latent space (i.e. joint angles). Then the error between the predicted and the current

sensor input (e.g. visual error e_v) is computed and added into the differential equation weighted by its relevancy (i.e. variance Σ_v). The goal attractor dynamics are also computed for the left and the right condition. Finally, the probability of being generating that sensor input is computed.

If the error between the prediction and the observed values is too high (outlier), a learning process starts, where the robot retrain the MDN with the new input. The outcome depends on three conditions. (1) When there are no contingent effects the robot is not able to learn anything and therefore it is not itself. (2) When it is able to learn an effect but then the model cannot properly approximate the sensory effect then it is not itself either. Finally, (3) if it is able to learn the effect and the error is enough small then it will accumulate evidence until it will properly answer: *it is me*.

Algorithm 1 Self/other distinction algorithm

Require: $\Sigma_p, \Sigma_v, \Sigma_\mu, \Delta_t$ LOGMIN, LOGMAX

```

1:  $\mu \leftarrow$  Initial joints angle estimation
2:  $\mu', v, \mathcal{L} \leftarrow \mathbf{0}$ 
3: while  $\neg$  recognized( $p_{self}$ ) do
4:    $h \leftarrow of(I), I \leftarrow$  camera  $\triangleright$  Optical flow histogram
5:    $p(cont) \leftarrow sigmoid(BCNN.forward(a, h)) \triangleright$  Sec. 2.3.3
6:    $g, \mu^*, \Sigma^* \leftarrow MDN.forward(\mu) \triangleright$  Predictor Sec. 2.3.2
7:    $\partial_\mu g \leftarrow MDN.backward(\mu^*, \Sigma^*) \triangleright$  Sec. 2.3.2
8:    $\dot{\mu} \leftarrow \mathbf{0} \triangleright$  Active Inference. Sec 2.3.1
9:   if  $\exists s_p$  then  $\triangleright$  Proprioception
10:      $e_p = (s_p - \mu)$ 
11:      $\dot{\mu} = \dot{\mu} + e_p / \Sigma_p$ 
12:      $\dot{a} = \dot{a} - (e_p / \Sigma_p) \Delta_t$ 
13:   if  $\exists s_v$  then  $\triangleright$  Vision
14:      $e_v = (s_v - g)$ 
15:      $\dot{\mu} = \dot{\mu} + \partial_\mu g^T e_v / \Sigma_v$ 
16:      $\dot{a} = \dot{a} - (\partial_\mu g^T e_v / \Sigma_v) \Delta_t$ 
17:   if  $\exists s_\rho$  then  $\triangleright$  Goal attractor  $\rho$  dynamics
18:      $e_f = (\mu' - f(\mu, \rho))$ 
19:      $\dot{\mu} = \dot{\mu} + \partial_\mu f^T e_f / \Sigma_\mu$ 
20:    $\mu = \mu + \Delta_t \dot{\mu} \triangleright$  1st order Euler integration
21:    $a = a + \Delta_t \dot{a}$ 
22:   if  $\exists s_v$  then  $\triangleright$  Self-recognition Sec. 2.3.4
23:      $\mathcal{L}_i = -0.5(e_v^T \Sigma_v^{-1} e_v + \ln \det \Sigma_v + n \ln 2\pi)$ 
24:      $\mathcal{L} = \mathcal{L} + \mathcal{L}_i + \ln p(cont)$ 
25:      $p_{norm} = (\exp(\mathcal{L}) - \text{LOGMIN}) / (\text{LOGMAX} - \text{LOGMIN})$ 
26:      $p_{self} = p_{self} + K(p_{norm} - p_{self})$ 
27:   if outlier then
28:     while  $i < N_s \wedge \neg$ timeout do
29:       if  $p(cont) > \delta$  then  $\triangleright$  Learning Sec. 2.3.3
30:          $Qx.append(s_p); Qy.append(s_v)$ 
31:       if  $i > N_s$  then
32:         MDN.train( $Qx, Qy$ )

```

2.3 Detailed computational model

We describe in detail the body perception and action model and then the method for self/other distinction. We define body perception as inferring joint angles of the robot arm μ by means of different sensory modalities. The robot approximates the belief of his body by means of minimizing the free energy [10], computing the most plausible solution of his arm location using the error between the predicted sensory input and the observed one. Mathematically, perception and action are interpreted by the following two equations

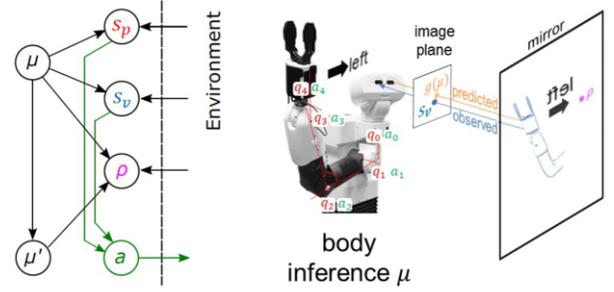


Figure 3. Body perception model and notation. (Left) Generative model of the robot perception and action. (Right) Mirror and robot setup. q are the joint measurements and μ, μ' is the inferred angle of the joints and its first-order dynamics. v is the visual location of the hand and ρ the goal. a_i is the action exerted in joint i .

adapted from the free energy principle approach model proposed in [11]:

$$\text{Perception} \quad \mu = \arg \min_{\mu} F \rightarrow \dot{\mu} = \mu' - \frac{\partial F(\mu, \rho)}{\partial \mu} \quad (1)$$

$$\text{Action} \quad a = \arg \min_a F \rightarrow \dot{a} = -\frac{\partial F(\mu, \rho)}{\partial a} \quad (2)$$

2.3.1 Active inference model

We adapted the free energy model presented in [27] to the TIAGO robot. Using joint encoders s_p and visual information s_v , the graphical representation of the inference process is described in Fig. 3. The agent encodes its belief of its body μ with the following density [6]:

$$p(\mu, s, \rho) = p(s_p | \mu) p(s_v | \mu) p(\mu' | \mu, \rho) \quad (3)$$

Where the first two likelihoods are the sensor modality contributions (assumed independent between each other) and the third term is the prior density for the first-order dynamics, which also depends on the causal variables ρ . We will control the robot in joint velocities by defining actions a as the velocity of each joint, thus only first-order dynamics are needed.

Under the mean-field Laplace approximation, the free energy bound can be approximated as⁴:

$$F \approx -\ln p(\{s, \rho\}, \mu) - \left[\frac{1}{2} \ln |\Sigma| + n \ln 2\pi \right] \quad (4)$$

where n is the size of μ . In order to solve Eq. 1 and 2, we computed the gradient of F with respect to the latent variables μ, a .

$$\frac{\partial F}{\partial \mu} = \frac{(s_p - \mu)}{\Sigma_p} + \frac{\partial g(\mu)^T (s_v - g(\mu))}{\partial \mu \Sigma_v} + \quad (5)$$

$$+ \frac{\partial f(\mu, \rho)^T (\mu' - f(\mu, \rho))}{\partial \mu \Sigma_\mu} \quad (6)$$

$$\frac{\partial F}{\partial a} = \frac{\partial s_p^T (s_p - \mu)}{\partial a \Sigma_p} + \frac{\partial s_v^T (s_v - g(\mu))}{\partial a \Sigma_v} \quad (7)$$

where $g_v(\mu)$ is the visual generative model that is learnt (Sec. 2.3.2) and $f(\mu, \rho)$ is the model dynamics of a perceptual attractor [27].

⁴ In some formulations of the free energy the sensor inputs s and other causal variables are the described by the same variable. However, in this paper, we split them for clarity purposes.

2.3.2 Generative models learning

To solve Eq. 5 and 7 the generative models for the sensor $g(\mu)$ and its partial derivative $\partial g/\partial \mu$ should be known. Conversely to previous works on active inference we enabled the robot to learn those models taking inspiration from our latest contributions on body perception with function learning [21, 22]. However, in this case, we used a probabilistic neural network model to provide scalability and interpolation. The visual generative model is learnt by means of a Mixture Density Network (MDN) [3]. In this way, we encode not only the mapping but the uncertainty associated. In principle, the forward model of the robot end-effector with respect to the joint latent space can be learnt using just one Gaussian kernel. However, we allowed the robot to learn more complex visual patterns generated by the actions or due to the lens non-linearity. The same model can generalize to forward model learning using more complex probabilistic deep models [7] within the free energy framework. The predicted sensation is computed through the forward pass of the MDN and outputs the Gaussian mixture model parameters: Π_i coefficients, \bar{s}_i mean and Σ_i variance. To compute the prediction of the sensation given the state $s = g(\mu)$ we select the maximum probable kernel and get the mean and the variance as follows:

$$i^* = \arg \max_i \text{softmax}(W_{\Pi_i}^T h(\mu) + b_{\Pi}) \quad (8)$$

$$g(\mu) = W_{\bar{s}}^T h(\mu) + b_{\bar{s}} \quad (9)$$

where $h(\mu)$ is the hidden layer output and W, b are the network weights and bias respectively. Although for this specific visual location learning only one hidden layer is sufficient, more layers can be added for more complex non-linear function learning.

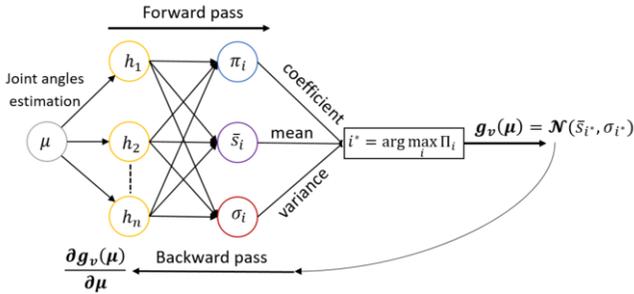
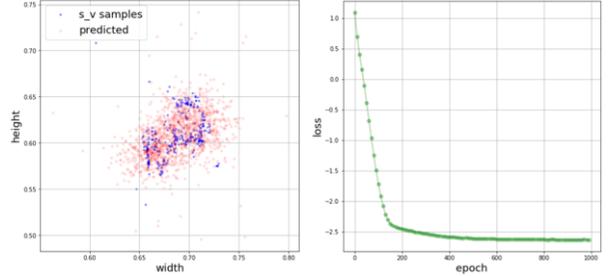


Figure 4. Forward model learning with MDN. The forward pass gives the prediction and the backward pass gives the partial derivative with respect to the body configuration μ .

Figure 4 describes the MDN to obtain the forward and partial derivatives through the backward pass. Figure 5 shows an instance of the visual forward model training for the robot in front of the mirror. The negative log-likelihood converges rapidly in 400 epochs. The blue dots correspond to the visual samples acquired when the robot moves the arm and the red dots are generated with the learnt predictor $g(\mu)$ for random joint angles within the set of the waving left and right movement. The interpolation power of the network shown is really crucial for the stability of free energy optimization.

The inverse mapping or partial derivative with respect to the latent variable is computed through the backward pass of the MDN of the most plausible Gaussian kernel i^* .

$$\frac{\partial g(\mu)}{\partial \mu} = P_{i^*} \frac{(\bar{s}_{i^*} - s_v)}{\sigma_{i^*}^2} \frac{\partial z_{i^*}}{\partial \mu} \quad (10)$$



(a) Training samples and $g(\mu)$ (b) MDN training loss

Figure 5. Learned mapping with the MDN from the joints positions to the 2D visual point (end-effector) computed through optical flow. Visual samples (blue) used for training and $g(\mu)$ predictor learned (red). Axis ranges that correspond to the image dimensions have been normalized between 0 and 1 dividing by the width (640px) and the height (480px) respectively.

where z_i is the neuron output, \bar{s}_i is the mean value output of the selected kernel, s_v is the current observation of the end-effector in the visual field and $P_{i^*} = \Pi_{i^*} \mathcal{N}(\bar{s}_{i^*}, \sigma_{i^*}^2) / \sum_j \Pi_j \mathcal{N}(\bar{s}_j, \sigma_j^2)$ - See [3].

In order to train the neural network we use as the input the joints estimated angles μ and as the output the 2D centroid location of the biggest moving optical flow blob. The points that are contingent (see Sec. 2.3.3) are buffered into a batch that is then used to modify network weights through backpropagation [18]. The loss function used is the standard negative log-likelihood of the Gaussian mixture model: $\mathcal{L} = -\ln \sum_i \Pi_i \mathcal{N}(\bar{s}_i, \sigma_i^2)$.

2.3.3 Contingency learning

We assume that the robot has a system that is able to classify whether it is a contingent effect or not. This prior knowledge can be justified by stating that the robot has learnt the potential effect of moving the hand left and right in front of him. In essence, the robot has never looked in a mirror but has performed actions with the arm that result in the following contingent effect learning: when the robot moves the hand to the left produces a change in the visual field with a specific direction. We computed the effect in the visual field given the action (i.e. velocity) by means of an optical flow algorithm [33]. To learn to classify this effect we used a deep learning classifier with input the state of the robot (i.e. estimated joint states μ) and the velocity being exerted (i.e. a) and the output the histogram of the optical flow velocity directions. Figure 6 shows the input, the output and the architecture of the neural network.

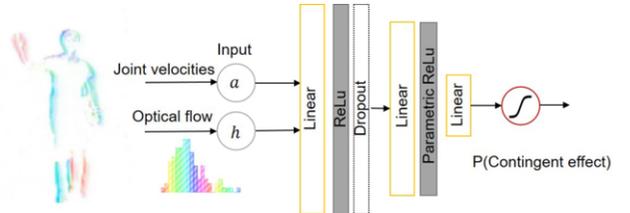


Figure 6. Contingent sensory event classifier. Given the histogram of the optical flow velocities of the visual field (left), the joint states and the velocities being exerted, the deep binary classifier computes whether the effects are contingent for the left and right action.

Figure 7 shows the input data, the learning curve and the classification of samples of the test set. The joint velocities (coloured

lines) for the hand wave movement show the profiles of left and right movements. The associated optical flow histogram (coloured dots) varies depending on the movement. The orientations are split into 10 bins represented with different colours and the y-axis is the normalized frequency. The loss function used is the binary cross-entropy with logistic output. Then we used a sigmoid function to compute the probability. The optimizer was Adam with 0.01 learning rate. It is important to highlight that we had to include the non-movement case to deal with the sensor noise. We tested, in Fig. 7(c), 100 random points from the test set where we included: self-generated actions, Gaussian random noise and non-movement with visual input. The contingency probability (y-axis) is shown for all those cases. Furthermore, some random samples have some probability of being classified as self-generated.

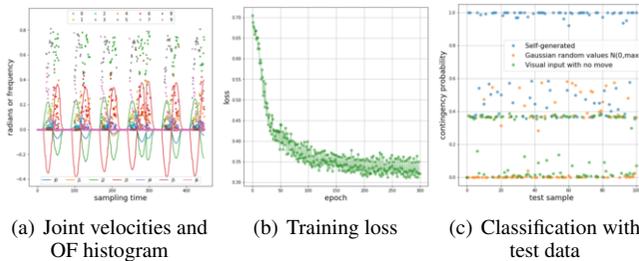


Figure 7. Contingency classifier learning. The input are the 7 joint velocities (lines) and the optical flow histogram with 10 bins (dots). The output is the probability of a image velocity vector generated by an action on the joints.

2.3.4 Self-recognition: Is that me?

In order to provide non-appearance self-recognition, we transform the question of *is that me?* into *did I generate those sensors values?* The robot accumulates evidence of generating the sensed effects in the world. This is computed through the marginal likelihood or sensory surprise. Highly surprising events (effects in the visual field) drastically drop the probability of being self. Conversely, in order to increase the probability of being self, the effect should continuously match the expected value generated by the learnt forward model.

We compute the probability that our model with a given set of parameters would generate a particular observed cue in the visual field as:

$$\mathcal{L}_i = -\frac{1}{2}(e_v^T \Sigma_v^{-1} e_v + \ln \det \Sigma_v + n \ln 2\pi) \quad (11)$$

$$\mathcal{L} = \mathcal{L} + \mathcal{L}_i + \ln p(\text{contingent}) \quad (12)$$

where $p(\text{contingent})$ is computed with the deep net classifier of Sec. 2.3.3. Inspired by [17], the probability of being itself is computed by normalization and smoothing as follows:

$$p_{norm} = (\exp(\mathcal{L}) - \text{LOGMIN}) / (\text{LOGMAX} - \text{LOGMIN}) \quad (13)$$

$$p_{self} = p_{self} + K(p_{norm} - p_{self}) \quad (14)$$

where K is a smoothing constant working as a Kalman gain.

3 Experimental results

Two experiments were designed: self-recognition in the mirror, depicted in Fig. 1A, and self/other distinction with other similar robot and a human, represented in Fig. 1B. A video describing the experiments is here: <https://youtu.be/319N972xjD8>. The parameters of the MDN network were fixed for all experiments: $\mu \in$

\mathbb{R}^7 , one hidden layer with 20 units, two-dimensional output $s_i \in \mathbb{R}^2$ (i.e. hand 2D position in the image). The kernels were Gaussian distributions \mathcal{N} with mean \bar{s}_i and variance σ_i . We assumed that the output variables (2D location in the image) have the same noise distribution, thus we only used one variance parameter for each Gaussian kernel. We fixed the number of Gaussian mixtures to four ($i \in (0, 4)$). The optimizer used was Adam [18] and the training used batches of 200 samples and 1000 epochs.

3.1 Self-recognition in the mirror

The robot was placed in 10 random positions in front of a body size mirror (Fig. 1A) and we ran the experiment 10 times for each position. The arm was always visible in the mirror and we did not include any prior information about the mirror. All experiments were successfully classified as itself.

The procedure went as follows. The robot moved the hand left and right by setting an attractor of the end-effector in the visual field. In other words, the desired location in the image changed from left to right of the image centre. The attractors were active for 2 seconds each. Then, the delay between initiating the left and right waving changed every time randomly in the range of 3 seconds. This enforced that actions were not periodic. Finally, the training finished if the number of samples was more than 200 or it reached 20 seconds.

Figure 8(a) shows the prediction error dynamics for proprioception e_p and visual s_v cues for one trial⁵. The bottom plot shows the mean and standard deviation of the probability of being self in all trials with no prior training with the mirror. The first time that the robot looks at the mirror and starts moving the arm, the robot is not able to predict its hand position producing a big prediction error. Before the training stage, the robot was not able to recognize itself. Thus, it starts learning the relation between its body state and the visual input by means of the MDN (Sec. 2.3.2). After learning for 10 seconds, the robot continues moving the hand left and right until it decides whether it is him or not. After learning the forward model the error in the visual sensation drops allowing the system to accumulate evidence of being itself (the model is correctly generating the visual sensations). Still, there are small prediction errors related to the adaptation of the current learnt model and real observations. Errors in proprioception are generated by the perceptual attractor created when it has the goal to move to the left or to the right. Any disturbance in the expected visual sensation (surprise) dropped the probability of being self.

3.2 Self/other distinction

For the self/other experiment, we performed the same amount of trials. We faced two similar robots doing the same movements but due to the randomness of the delays between actions they were asynchronous (Fig. 1B). As well as in the previous experiment, figure 8(b) shows the error in proprioceptive and visual cues of one of the trials. The bottom plot of Fig. 8(b) shows the probability of being self for all trials. In this experiment, the robot could not learn anything as there was no causality between the actions (joint velocities) and the observed sensory information. This yielded into the big visual prediction errors after the training. Therefore the robot inferred that it is not itself.

We further tested the model with an experimenter in front (Fig. 1B) and we observed that even when trying to perform the

⁵ Other trials have similar profiles but with different delays. Thus, they don't bring any further information.

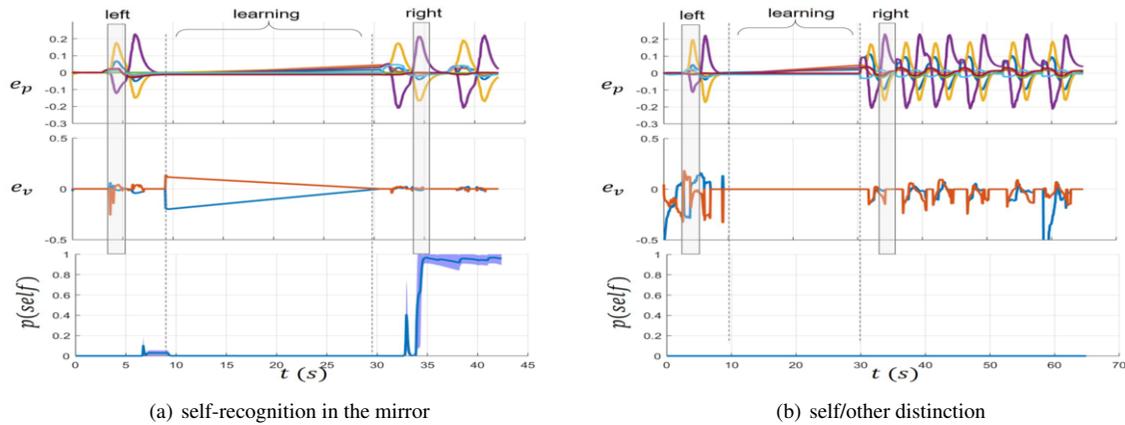


Figure 8. Self-recognition and self/other distinction experiments variable dynamics. e_p and e_v are the proprioceptive and visual errors. During learning, the active inference update is disabled. The probability of being self, $p(\text{self})$, rises when the predicted visual effect is contingent and matches the observation.

same movements the robot was able to properly distinguish between self-generated and other generated visual cues. In all cases, the robot correctly identified the agent in front as not itself. However, as expected, in the case of perfect synchronization the robot will identify it as itself. From the cognition point of view interpreting another robot as itself seems wrong, but also humans will perceive agency when seeing an arm that moves at every instant with correlation with their commands to the muscles.

4 Machines able to recognize themselves

We pointed out that unconscious body perception, action and learning is underneath self-recognition. Experiments showed that contingent non-appearance cues can be learned through interaction and be used to perform self/other distinction. The proposed algorithm enabled the robot to robustly identify itself in the mirror just by learning the forward model, i.e., the expected visual input given the inferred body state. As long as the arm appeared in the mirror, the initial conditions were irrelevant. This is in line with the experiments with non-human primates that needed learning with the mirror before achieving the self-recognition ability.

The robot was also able to differentiate itself from other entities. The algorithm was sensitive to any unexpected sensations as we are using the minimization of surprise as the key mechanism for the differentiation. Even when trying to deceive the algorithm (e.g. robot-human experiment), any small delay or difference dropped down the probability of being itself. Interestingly, the variance parameters (relevance of the sensory modality for inferring the body) was crucial for dealing with the robot own noise but differentiating from other entities cues. We further observed an interesting side effect of using the action to minimize the prediction error. The human or robot hand in the self/other distinction experiment acted sometimes as an attractor for the body inference, generating mirroring actions similar to imitation.

In summary, with these experiments, we are showing that self/other distinction and self-recognition is possible to be implemented in artificial agents. Although it will not be a conscious cognitive process as described in humans or elephants, it will have some common ground in the sensorimotor action-perception level.

Double comparator model

Within the self-perception cognitive science literature it is com-

mon knowledge the comparator model [8]. This approach determines that humans are aware of producing an effect in the world through their actions (i.e. agency [16]) by means of the prediction error congruence. Some researchers criticized the simplistic vision of the model and proposed that contingency detection and causal inference should be also involved as processes on top of congruence [37]. This is similar to our model. We use both prediction error and contingency learning. In essence, we propose the “double comparator” model inspired by the work of [36]. The first comparator is the prediction error mechanism that approximate the models learnt to reality for improving interpretation. The second comparator takes into account other processes like spatio-temporal contingency and performs the proper distinction. In our model, we engineered this abstract level using a probabilistic approach, i.e., by calculating the marginal likelihood to accumulate temporal evidence and then computing the probability of generating that observed effect in the sensory space.

Towards passing the original mirror test

We showed how non-appearance cues can be used for self-recognition. However, robots are still not able to pass the original self-recognition test like animals do [2]. For instance, the robot here did not learn the mapping between the tactile and the visual reflection to produce the motor response. In order to fully provide robots with a full-fledged self-recognition similar to humans, there are a set of technical challenges that should be further addressed:

- *Complex visual segmentation.* The algorithm was designed for only one moving blob segmentation using optical flow. Complex inputs with multiple causal sources need improved segmentation.
- *Large scale input data.* Forward models learning of the end-effector location in the image should be scaled to real sensory input data, such as images.
- *Whole body interactions.* We used the arm for the experiments as a proof-of-concept. The same approach can be extended to whole-body self/other distinction.
- *Other cues.* The proprioceptive-visual tandem is just one of the sensory inputs involved in self-recognition. We did not model appearance, tactile or auditory cues. We considered movement contingency as the first level of self-distinction but then an appearance model similar to the body image should be included.
- *Biological plausibility.* Although the mathematical abstraction of body perception has strong neuroscience-inspired notions, using the marginal likelihood for accumulating evidence does not have

a direct relation with the mechanism observed in the brain. However, it has roots into the model evidence approach under the Bayesian brain assumption.

5 Conclusion

We proposed an algorithm for self/other distinction using non-appearance sensory cues by combining free energy optimization with neural network learning. The approach showed the importance of not only learning the models of the body but also approximating those models to the observed reality. After learning the humanoid robot was able to accumulate evidence during the generated movements to discern whether it was itself or other entity in different scenarios. This reliability is due to the explicit handling of the sensorimotor uncertainty in the proposed model, while still performing robust self-differentiation. When the expected sensation was similar to the sensor values, i.e. the prediction error was low, the evidence of being itself was accumulated. Conversely, any prediction error dropped down the probability of being itself yielding to self/other distinction. This work stresses that self-recognition and self/other distinction grounds on the body perception and action model and that it is possible to give artificial agents, such as robots, these capabilities. Further work will focus on scaling to raw images [32] and including appearance cues.

ACKNOWLEDGEMENTS

This work was supported by the MSCA SELFCEPTION project (www.selfception.eu) EU H2020 grant no. 741941 and PAL robotics.

REFERENCES

- [1] James R Anderson and Gordon G Gallup, 'Mirror self-recognition: a review and critique of attempts to promote and engineer self-recognition in primates', *Primates*, **56**(4), 317–326, (2015).
- [2] James R Anderson and Gordon G Gallup Jr, 'Which primates recognize themselves in mirrors?', *PLoS Biology*, **9**(3), e1001024, (2011).
- [3] Christopher M Bishop, 'Mixture density networks', Technical report, (1994).
- [4] Sarah-Jayne Blakemore and Jean Decety, 'From the perception of action to the understanding of intention', *Nat. Rev. Neurosci.*, **2**(8), 561, (2001).
- [5] Matthew Botvinick and Jonathan Cohen, 'Rubber hands 'feel' touch that eyes see', *Nature*, **391**(6669), 756, (1998).
- [6] Christopher L Buckley, Chang Sub Kim, Simon McGregor, and Anil K Seth, 'The free energy principle for action and perception: A mathematical review', *Journal of Mathematical Psychology*, (2017).
- [7] Kurtland Chua, Roberto Calandra, Rowan McAllister, and Sergey Levine, 'Deep reinforcement learning in a handful of trials using probabilistic dynamics models', in *Advances in Neural Information Processing Systems*, pp. 4754–4765, (2018).
- [8] Nicole David, Albert Newen, and Kai Vogeley, 'The "sense of agency" and its underlying cognitive and neural mechanisms', *Consciousness and cognition*, **17**(2), 523–534, (2008).
- [9] Peter Dayan, Geoffrey E Hinton, Radford M Neal, and Richard S Zemel, 'The helmholtz machine', *Neural computation*, **7**(5), 889–904, (1995).
- [10] Karl Friston, 'The free-energy principle: a unified brain theory?', *Nature reviews neuroscience*, **11**(2), 127, (2010).
- [11] Karl J Friston, Jean Daunizeau, James Kilner, and Stefan J Kiebel, 'Action and behavior: a free-energy formulation', *Biological cybernetics*, **102**(3), 227–260, (2010).
- [12] Shaun Gallagher, 'Philosophical conceptions of the self: implications for cognitive science', *Trends Cogn. Sci.*, **4**(1), 14–21, (2000).
- [13] Yasmin Kim Georgie, Guido Schillaci, and Verena Vanessa Hafner, 'An interdisciplinary overview of developmental indices and behavioral measures of the minimal self', in *2019 Joint IEEE 9th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, pp. 129–136, IEEE, (2019).
- [14] Kevin Gold and Brian Scassellati, 'Using probabilistic reasoning over time to self-recognize', *Robotics and Autonomous Systems*, **57**(4), 384–392, (2009).
- [15] Hermann von Helmholtz, *Handbuch der physiologischen Optik*, volume 9, Leopold Voss, 1867.
- [16] Bernhard Hommel, 'Action control and the sense of agency', *The sense of agency*, 307–326, (2015).
- [17] Sebastian Kahl and Stefan Kopp, 'A predictive processing model of perception and action for self-other distinction', *Frontiers in psychology*, **9**, 2421, (2018).
- [18] Diederik P Kingma and Jimmy Ba, 'Adam: A method for stochastic optimization', *arXiv preprint arXiv:1412.6980*, (2014).
- [19] David C Knill and Alexandre Pouget, 'The bayesian brain: the role of uncertainty in neural coding and computation', *TRENDS in Neurosciences*, **27**(12), 712–719, (2004).
- [20] Masanori Kohda, Takashi Hotta, Tomohiro Takeyama, Satoshi Awata, Hirokazu Tanaka, Jun-ya Asai, and Alex L Jordan, 'If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals?', *PLoS biology*, **17**(2), e3000021, (2019).
- [21] Pablo Lanillos and Gordon Cheng, 'Active inference with function learning for robot body perception', *International Workshop on Continual Unsupervised Sensorimotor Learning, IEEE Developmental Learning and Epigenetic Robotics (ICDL-EpiRob)*, (2018).
- [22] Pablo Lanillos and Gordon Cheng, 'Adaptive robot body learning and estimation through predictive coding', in *2018 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, pp. 4083–4090, IEEE, (2018).
- [23] Pablo Lanillos, Emmanuel Dean-Leon, and Gordon Cheng, 'Yielding self-perception in robots through sensorimotor contingencies', *IEEE Transactions on Cognitive and Developmental Systems*, **9**(2), 100–112, (2016).
- [24] Pablo Lanillos, Emmanuel Dean-Leon, and Gordon Cheng, 'Enactive self: a study of engineering perspectives to obtain the sensorimotor self through enaction', in *Developmental Learning and Epigenetic Robotics, Joint IEEE Int. Conf. on*, (2017).
- [25] Robert W Mitchell, 'Kinesthetic-visual matching and the self-concept as explanations of mirror-self-recognition', *Journal for the theory of social behaviour*, **27**(1), 17–39, (1997).
- [26] Yukie Nagai, Yuji Kawai, and Minoru Asada, 'Emergence of mirror neuron system: Immature vision leads to self-other correspondence', in *Development and Learning (ICDL), IEEE Int. Conf. on*, volume 2, pp. 1–6, (2011).
- [27] Guillermo Oliver, Pablo Lanillos, and Gordon Cheng, 'Active inference body perception and action for humanoid robots', *arXiv preprint arXiv:1906.03022*, (2019).
- [28] Jordi Pages, Luca Marchionni, and Francesco Ferro, 'Tiago: the modular robot that adapts to different research needs', in *Proc. of the IROS Workshop on Robot Modularity*, (2016).
- [29] Alexandre Pitti, Ganna Pugach, Philippe Gaussier, and Sotaro Shimada, 'Spatio-temporal tolerance of visuo-tactile illusions in artificial skin by recurrent neural network with spike-timing-dependent plasticity', *Scientific reports*, **7**, (2017).
- [30] Rajesh PN Rao and Dana H Ballard, 'Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects', *Nature neuroscience*, **2**(1), 79–87, (1999).
- [31] Philippe Rochat, 'Five levels of self-awareness as they unfold early in life', *Consciousness and cognition*, **12**(4), 717–731, (2003).
- [32] Cansu Sancaktar and Pablo Lanillos, 'End-to-end pixel-based deep active inference for body perception and action', *arXiv preprint arXiv:2001.05847*, (2019).
- [33] Eero P Simoncelli, 'Bayesian multi-scale differential optical flow', *Handbook of Computer Vision and Applications*, 397–422, (1999).
- [34] Alexander Stoytchev, 'Self-detection in robots: a method based on detecting temporal contingencies', *Robotica*, **29**(01), 1–21, (2011).
- [35] Matthis Synofzik, Gottfried Vosgerau, and Martin Voss, 'The experience of agency: an interplay between prediction and postdiction', *Frontiers in psychology*, **4**, 127, (2013).
- [36] Daniel M Wegner, *The illusion of conscious will*, MIT press, 2017.
- [37] Lorijn Zaadnoordijk, Tarek R Besold, and Sabine Hunnius, 'A match does not make a sense: on the sufficiency of the comparator model for explaining the sense of agency', *Neuroscience of consciousness*, **2019**(1), niz006, (2019).