

Learning to reach to own body from spontaneous self-touch using a generative model

Valentin Marcel*, J. Kevin O'Regan†, and Matej Hoffmann*

*Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague,
valentin.marcel@fel.cvut.cz, matej.hoffmann@fel.cvut.cz

†Integrative Neuroscience and Cognition Center - UMR 8002 CNRS & Université Paris Descartes,

Abstract—When leaving the aquatic constrained environment of the womb, newborns are thrown into the world with essentially new laws and regularities that govern their interactions with the environment. Here, we study how spontaneous self-contacts can provide material for learning implicit models of the body and its action possibilities in the environment. Specifically, we investigate the space of only somatosensory (tactile and proprioceptive) activations during self-touch configurations in a simple model agent. Using biologically motivated overlapping receptive fields in these modalities, a variational autoencoder (VAE) in a denoising framework is trained on these inputs. The denoising properties of the VAE can be exploited to fill in the missing information. In particular, if tactile stimulation is provided on a single body part, the model provides a configuration that is closer to a previously experienced self-contact configuration. Iterative passes through the VAE reconstructions create a control loop that brings about reaching for stimuli on the body. Furthermore, due to the generative properties of the model, previously unsampled proprioceptive-tactile configurations can also be achieved. In the future, we will seek a closer comparison with empirical data on the kinematics of spontaneous self-touch in infants and the results of reaching for stimuli on the body.

Index Terms—learning body representation, self-touch, embodied cognition, body schema, multimodal variational autoencoder, sensorimotor contingencies

I. INTRODUCTION

The period around birth is perhaps the most confusing for a newborn due to extremely dramatic physical changes, as well as intensive neural development. Physical changes encompass a sudden transition from an aquatic, constrained, largely dark environment to the world outside with stronger gravity, different and strongly varying lighting conditions, and much stimulation beyond the infant's control. At the same time, there is also a massive overall growth in brain volume, the onset of myelination, and the striking development (but also pruning) of gray matter connections, especially in the sensorimotor and visual cortices [1]. Experience, i.e. activity-dependent factors, starts playing a more prominent role in shaping neural circuitry [1], [2]. In this context, the body is likely "the most consistent, the most predictable, and the most verifiable part of the environment" [3].

A candidate mechanism for learning the affordances but also the extent or structure of one's own body, is self-touch [4]. Self-touch is already observed in the fetus [5, p. 29-30, 113-114] but also frequently in the first months after birth [6]. As a good number of such contacts occur out of the infant's

sight and also as vision is not mature at this point, we focus on what can be learned from somatosensory (i.e. tactile and proprioceptive) information alone. Rochat [7] writes: "By 2-3 months, infants engage in exploration of their own body as it moves and acts in the environment. They babble and touch their own body, attracted and actively involved in investigating the rich intermodal redundancies, temporal contingencies, and spatial congruence of self-perception." Following up on our previous, still largely theoretical account [8] drawing on the sensorimotor contingency theory [9], in this work we want to give more concrete contours to this hypothesis through a computational model. The behavior of the model will be compared with what is known about the capacities of the infants in reaching the body. A paradigm to test to what extent infants have developed their "body know-how" is placing vibrotactile stimuli ("buzzers") on different parts of the body and observing their reactions. The experimental results demonstrate how the ability to reach the stimulus is gradually acquired mainly over the period between 2 and 8 months of age [8].

We present a model that learns a multimodal latent representation of self-touch. The interaction with the body is very stable and predictable, especially thanks to the contingencies between touch and proprioception; thus, we can expect the representation to be learned simply through self-exploration. Moreover, learning this latent representation using a variational autoencoder in a denoising framework, we obtain an agent with the ability to internally reconstruct part of its sensorimotor information. This reconstruction capability enables the agent to assimilate any "non-self-touch" interaction with the environment as partial self-touch. In particular, we show that the learned model can generate an entire distribution of self-reaching configurations when stimulated on a part of its body.

The variational auto-encoder (VAE) [10] is a generative network model that learns to represent a data set in a low-dimensional latent space. In the denoising case, such models have the ability to reconstruct missing information. When the input is multimodal, it is possible to reconstruct images from sound, proprioception from vision, and even predict future sensory input from motor commands in a single model [11]. In this work, we implement a similar multimodal model that mixes proprioceptive and tactile (i.e. somatosensory)

afference. It is able to reconstruct touch from proprioception, proprioception from touch, and even generate the desired configurations to reach for an externally stimulated part of the body. More generally, we want to show that, from a latent representation of self-touch, the agent is already able to represent interaction with its environment through internal “mental imagery” of self-touch.

The use of a VAE with its multi-purpose latent representation allowed us also to avoid the need to have explicit maps of the body surface, arm position in space, and motor space, and to avoid learning transformations between these maps. The idea was that such maps are not necessary if we consider that for the agent, for example, “knowing where it is touched” amounts to “knowing how to move to the touched position with its finger”.

Another design constraint that we adopted in our model was the idea that learning by the VAE would be facilitated if the codes used for sensory input and motor output were high-dimensional, multiscale, distributed population codes. This also seemed to be compatible with biology, where both tactile receptors and motor effectors are known to have a wide range of receptive or effector field sizes and are highly redundant.

Finally, we chose to simplify the way we implemented motor control in our approach by assuming that the agent could attain a particular limb position simply by fixing the proprioception values corresponding to that position. This is inspired by the Equilibrium Point Hypothesis (EPH), according to which the brain controls motor behavior by setting equilibrium points, approximately desired postures of the limbs that the musculoskeletal system seeks based on its dynamical properties (e.g., [12], [13]).

The article is organized as follows. In Section III the model for learning a multimodal latent representation is presented. The generative properties of the model are integrated in a control loop that allows the agent to perform a self-reaching task in response to external stimulation. The results are presented in Section IV and compared with developmental evidence and discussed in Section V.

II. RELATED WORK

Several teams have studied early sensorimotor development and learning of the corresponding representations using embodied computational models for agents of different complexity or for robots. Several review articles have appeared [14]–[17]. For the development of the fetus, the work of Kuniyoshi et al. (e.g. [18]) stands out.

Here, we focus on the period immediately after birth and on the self-touch scenario. Roschin et al. [19] and Marcel et al. [20] have approached the problem in a more abstract and mathematical sense. Inspired by Henri Poincaré’s pioneering observations, they studied how an agent consisting of a multi-DoF arm and a spherical body with tactile receptors can infer the dimensionality of its body by performing self-contact. Mannella et al. [21] used a simple planar agent (the same is used here) to study how it can efficiently discover the space of its body surface by internally generating self-contact

goals. Gama et al. [22] used the Nao robot with whole-body skin and tested different active exploration algorithms. Asada, Yoshikawa, Fuke, Hikita, and colleagues developed several models related to learning body models by connecting proprioceptive, tactile, and visual information (e.g., [23], [24]), often by applying Hebbian learning on top of the spaces of individual modalities represented by self-organizing maps.

Connecting or combining different modalities can also be achieved by other neural network architectures. Recently, VAE has gained popularity. One of its assets is its generative nature, that is, the ability to represent complete probability distributions. The denoising VAE is also capable of “filling in” information. Zambelli et al. [11] applied VAE to represent the space of proprioceptive, tactile, visual, auditory, and motor information of the humanoid robot iCub. Blum et al. [25] focused only on proprioception, but set the response properties of neurons that receive proprioceptive afference (in area 2 of the somatosensory cortex) as their target. To this end, they developed a modification of the VAE they call topo-VAE, as it introduces lateral connectivity to enforce topology preservation on the hidden layer. In the present work, we use “vanilla” denoising VAE and train it with tactile and proprioceptive inputs using population coding. Only self-contact configurations are used for training. Our inspiration was from somatosensory representations in primate brain posterior parietal cortex like area 5 (e.g., [26]), where tactile and proprioceptive inputs are combined, possibly giving rise to first implicit representations of the “body in space”. Furthermore, we generate reaching movements for tactile stimuli by iteratively using successive approximate reconstructions of the learned self-contact configurations.

III. METHODOLOGY

A. The agent

The agent, shown in Fig. 2, is a 2D stick figure, with two 3 Degree-of-Freedom (DoF) arms connected through a torso [21]. The arms have the ability to cross each other, generating contacts at the exact position where the arms cross. With respect to the task of generating a contact at one specific location on the body, the platform is kinematically redundant—there are several configurations that result in contact in a specific position. Multiple crossings are allowed so that multiple contacts can be obtained on the arms and torso. In this work agent has limbs of lengths [0.5, 0.8, 0.7, 1.4, 0.7, 0.8, 0.5] respectively for the left hand, left forearm, left arm, torso, right arm, right forearm, right hand and a total length $L = 5.4$.

1) Joint angle space

The agent is made up of 6 actuators encoded by their joint angles θ_k where $k = 1..6$. The joint angles are constrained to mimic the behavior of a biological arm. The limits are decided arbitrarily and correspond more or less to real bio-mechanical constraints: $[-\pi/4, \pi/2]$, $[0, 3\pi/4]$, $[0, 2\pi/3]$, $[0, 2\pi/3]$, $[0, 3\pi/4]$, $[-\pi/4, \pi/2]$, respectively, for left wrist, left elbow, left shoulder, right shoulder, right elbow, right wrist. Outside the joint limits, the model is collision-free.

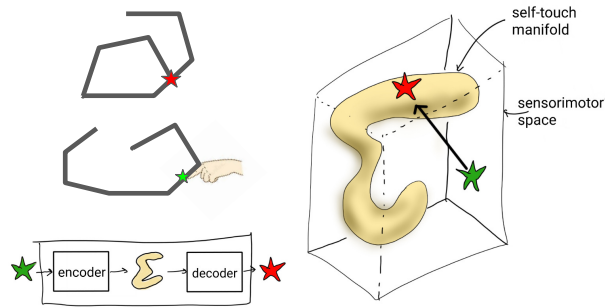


Fig. 1: Conceptual overview. The VAE model learns a latent representation of the self-touch space. The agent exploits the reconstruction properties of the VAE to map an external touch to a self-touch at the external touch location.

2) Proprioceptive space

The agent is endowed with a set of 120 proprioceptive sensors, 20 for each degree of freedom. Each sensor activation depends on the joint angle θ_k of an actuator following the law:

$$a_i = e^{-\frac{(\theta_k - c_i)^2}{2\sigma_i^2}}. \quad (1)$$

The activation of the proprioceptive sensor i is $a_i \in [0, 1]$ and corresponds to the response of a Gaussian receptive field with mean c_i and standard deviation σ_i . The means are drawn uniformly inside joint limits and the standard deviation are chosen uniformly in $[.4, .6]$ of respective joint limit ranges.

3) Tactile space

The agent is endowed with 100 tonic tactile sensors, i.e. responding to sustained pressure, whose activations are a population coding of contact positions on the agent's skin. A contact on the agent body corresponds to a position $x \in [0, L]$ on the agent's 1-dimensional body—cf. Figure 3 where L denotes the length of the agent's body surface ("skin"). The activation of tactile sensors depends on the position of the contact given by:

$$s_j = e^{-\frac{(x - \mu_j)^2}{2\sigma_j^2}}. \quad (2)$$

The activation of tactile sensor j is $s_j \in [0, 1]$ and follow the Gaussian receptive field with mean μ_j and standard deviation σ_j . The receptive field means are drawn uniformly in $[0, L]$ and the standard deviations from $[0.2L, 0.4L]$. When there are multiple contacts on the body, the sensory activation for each sensor is taken to be the activation coming from the closest contact point such that:

$$s_j = \max_{x \in \text{all contact points}} e^{-\frac{(x - \mu_j)^2}{2\sigma_j^2}}. \quad (3)$$

A contact at the exact position of the sensor mean would give an activation of 1 for this sensor, which decreases when the touch is farther away. If no contact, then the activation is 0.

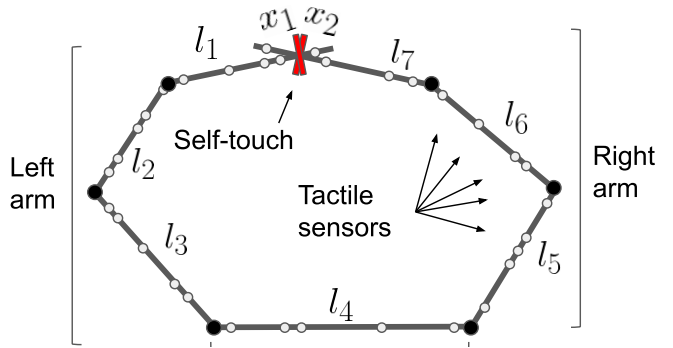


Fig. 2: Structure of the agent in a self-touch configuration.

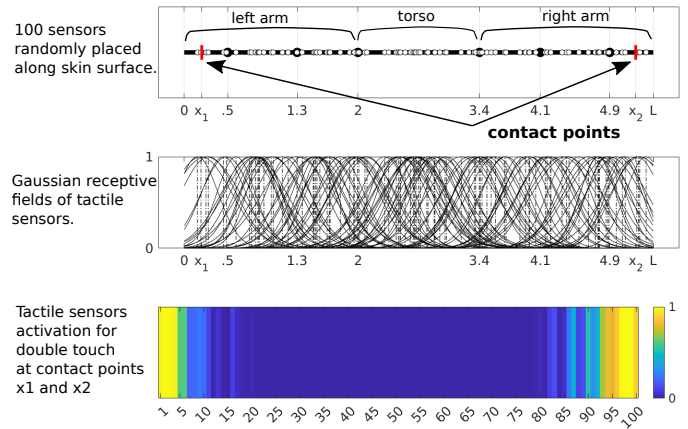


Fig. 3: Tactile population coding. Tactile activities are computed from the distance of touches to the center of Gaussian receptive fields of each tactile sensor on the whole body.

B. Learning self-touch latent space

1) Detecting self-touch

In this work, we will assume that the agent has the intrinsic ability to detect self-touch. Self-touch events have the specific property of simultaneous activation of tactile sensors on different parts of the body. In addition, self-touch occurs in specific configurations of the body. We take inspiration from the "body matching" neurons reported in area 5 of the monkey parietal cortex [26].

2) The Variational Autoencoder Model

The agent learns a latent representation of self-touch using a multimodal variational autoencoder [10]; see Figure 4. A variational auto-encoder is a generative model that consists of two parts: an encoder whose task is to encode the input data $\mathbf{x} = [\mathbf{a}, \mathbf{s}]$ into a latent representation

$$\mathbf{z} = \text{encoder}(\mathbf{x}) \quad (4)$$

usually of lower dimension, and a decoder which attempts to reconstruct the data input from the latent code

$$\hat{\mathbf{x}} = \text{decoder}(\mathbf{z}). \quad (5)$$

Both encoder and decoder are neural networks. In the variational framework, the latent space is stochastic and inputs \mathbf{x} are encoded into a prior probability density, in our case a Gaussian

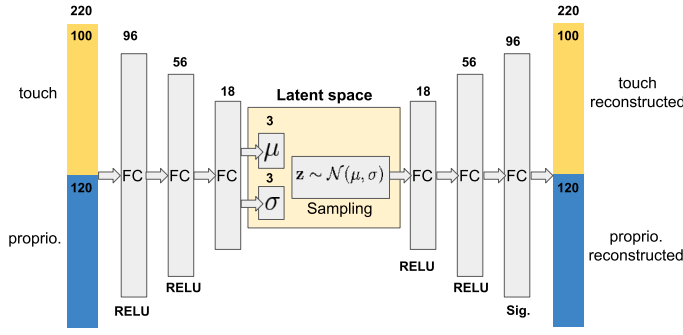


Fig. 4: Variational autoencoder model with 3 encoding layers, a bottleneck layer encoding the mean and variance of the latent code and a decoder with 3 layers.

distribution. For each data point \mathbf{x} , the encoder outputs a mean μ and a standard deviation σ from which multiple latent codes can be sampled $\mathbf{z} \sim \mathcal{N}(\mu, \sigma)$ and then decoded. The proposed VAE model is presented in Figure 4. The somatosensory inputs—proprioceptive and touch activations—are mixed into three successive fully connected layers with RELU activation functions, which lead to a bottleneck layer with two low-dimensional networks encoding the mean and variance of the data distribution in the latent space. The original input is then reconstructed from the second part of the VAE with a similar network structure. The variational autoencoder is trained to optimize a cost function that both penalizes the reconstruction error between input data \mathbf{x} and reconstructed data $\hat{\mathbf{x}}$ as well as the difference between the prior distribution and the actual distribution of encoded input data. The neural network is trained using backpropagation and the cost function is defined as the ELBO loss [27]. Importantly, during motor exploration, training data are collected only during self-touch events—for example, the left arm or the right arm cross the torso or each other.

C. Reconstructing self-touch, the denoising framework

In this article, we want the agent to be able to reconstruct or infer self-touch configurations when it receives an external tactile stimulus. The main hypothesis of this work is the ability of the model to generate self-touch even when the input data do not come from a self-touch configuration; see Figure 1. Because the model is trained on a source domain (self-touch), which differs from the target domain (“incomplete touch”), such a transfer is not straightforward, especially in the unsupervised context of ours. However, an external tactile stimulus touch, or incomplete touch, shares common features with self-touch configurations. In certain configurations, for example, a touch on the left arm can be obtained simply by moving the right hand to the touch area without moving the left arm. Therefore, it would be enough to reconstruct only the part of the sensorimotor input that represents the right arm. In this work, we exploit the reconstructing properties of a VAE in a denoising framework to obtain the ability to generate self-touch from partial touch.

Similarly to the denoising autoencoder, the variational autoencoder can be used to reconstruct missing information in

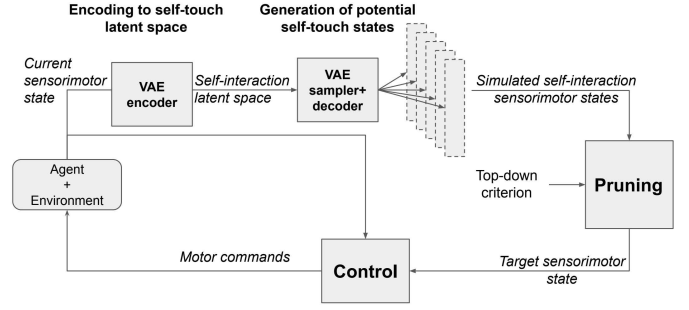


Fig. 5: Online control loop through iterative reconstructions.

a denoising framework. The VAE is trained on synthetically corrupted data and is asked to reconstruct the original version. To obtain the reconstruction property, training is done on both intact inputs and corrupted inputs with artificial added noise. When faced with corrupted data, the network is trained to reconstruct the original intact input. This property is at the core of this work. For the trained network, any configuration which is a single-touch only or even no-touch at all can be considered as a noisy self-touch configuration. When going through the denoising VAE model, such configurations can be expected to be projected to self-touch configurations.

D. Control loop: from external stimulation to self-touch

The denoising VAE performance depends on the level of noise in the input of the model. Although agent configurations close to training data can be successfully projected back in self-touch configurations as shown later in Figure 10, other configurations cannot be directly projected in an acceptable configuration due to their distance from the training distribution. In this work, we propose an active control loop that performs commands in the direction of reconstructed self-touch configurations so as to get closer and closer to it until an actual self-contact. The control loop is presented in Figure 5 and can be described as follows.

The current configuration, which may not be a self-touch configuration, is encoded through the encoder of the VAE, projecting it into a distribution in the latent space. The VAE sampler then generates multiple samples of the latent code according to the encoded distribution, which are then decoded to create a set of possible sensorimotor states. From them, the agent can pick the one it desires according to a top-down criterion that depends on the current goal of the agent. The chosen target sensorimotor state is then fed to a controller to perform a movement in the direction of the predicted self-touch configuration. This process is then repeated until the self-touch goal is achieved.

1) Prediction error as criterion towards self-touch

The current sensorimotor configurations \mathbf{x} , are encoded in the latent space as a posterior distribution $q(\mathbf{z}|\mathbf{x})$ shaped as a Gaussian distribution $\mathcal{N}(\mu, \sigma)$ where μ and σ are output by the two latent layers of VAE. From this distribution, the agent can sample a subset of N potential self-touch latent codes \mathbf{z}_n , $n \in [1, N]$, which are then decoded with

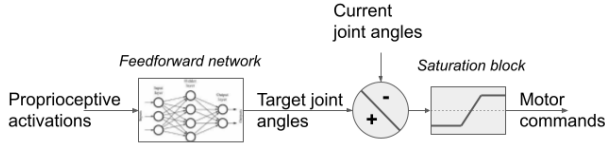


Fig. 6: Simple controller. The controller perform neural decoding of proprioceptive activations into target joint angles which are then compared to the motor configuration. The difference is then saturated to perform commands in the direction of the angle changes.

$$\hat{\mathbf{x}}_n = \text{decoder}(z_n). \quad (6)$$

Therefore, from a current state \mathbf{x} the VAE generates a set of sensorimotor states $\hat{\mathbf{x}}_n$ that are predictions of self-touch configurations. To pick the best reconstructed configuration we will use the standard reconstruction error:

$$C_{\text{reconstruction}}(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2. \quad (7)$$

Given the agent's current state \mathbf{x} , the VAE output minizing the reconstruction error is the next target configuration \mathbf{x}^* for the agent:

$$\mathbf{x}^* = \underset{\hat{\mathbf{x}}_n}{\text{argmin}} \quad C_{\text{reconstruction}}(\mathbf{x}, \hat{\mathbf{x}}_n). \quad (8)$$

By minimizing the reconstruction error, the agent selects the prediction that best matches the sensory input after going through the VAE information bottleneck. By doing so, the agent picks a target configuration that is the closest to the input configuration while being a potential self-touch. If the agent attempts to reach the target configuration, its own state will change and thus the target configuration. Thus, self-touch is performed in an iterative process through a control loop described in Figure 5 and aims towards a configuration that minimizes the reconstruction error.

2) The controller

The motor commands are computed based on the proprioceptive activations that makes up a part of the target sensorimotor state. They can be decoded back to joint angles using a feedforward neural network; cf. Figure 6. Let $\mathbf{x}^* = [\mathbf{a}, \mathbf{s}]$ be the target sensorimotor state; then the target joint angles are

$$\boldsymbol{\theta}^* = \text{NeuralDecoder}(\mathbf{a}). \quad (9)$$

where \mathbf{a} is the vector of proprioceptive activations. To perform neural decoding, we used a feedforward neural network with multiple hidden layers. See Xu et al. [28] for a comparison of neural decoding methods. The feedforward neural network is trained to reconstruct joint angles from proprioceptive encodings.

The target joint angles $\boldsymbol{\theta}^*$ are then compared with the current joint configuration $\boldsymbol{\theta}$. The difference vector is then “saturated” to obtain smoother joint angles commands.

$$\mathbf{u} = \text{sign}(\boldsymbol{\theta}^* - \boldsymbol{\theta}) \cdot \min(|\boldsymbol{\theta}^* - \boldsymbol{\theta}|, \Delta\theta_{\max}) \quad (10)$$

which are then performed by the actuators of the agent.

Notice that the target sensorimotor state may not be an acceptable proprioceptive encoding, in the sense that it does not always correspond to feasible joint configurations. Having a neural decoder that is robust to noise is important to obtain the correct planning properties.

3) Reaching process

Reaching is performed as follows: let $\boldsymbol{\theta}(t)$ be the current joint angles, $\mathbf{s}(t)$ the current tactile activations, $\mathbf{a}(t)$ the current proprioceptive activations, and $\mathbf{x}(t) = [\mathbf{a}(t), \mathbf{s}(t)]$ the current sensorimotor state. At $t = 0$, the agent receives an external touch that activates the tactile sensors. The sensorimotor state is fed to the VAE which generates a set of N potential sensorimotor states $\hat{\mathbf{x}}_n$, $n \in [1, N]$ according to the process in Equation 6, each attempting to reconstruct a self-touch configuration. The target sensorimotor state \mathbf{x}^* is taken following Equation 8 From the target sensorimotor state $\mathbf{x}^* = [\mathbf{a}^*, \mathbf{s}^*]$, the target proprioceptive activation \mathbf{a}^* is extracted and decoded with Equation 9 to obtain the joint angles of the target $\boldsymbol{\theta}^*$. The motor commands $\mathbf{u}(t)$ are then sent to the motors using the current joint angles $\boldsymbol{\theta}(t)$ with the Equation 10. The loop is repeated until the reconstruction error falls under a predefined threshold or after a fixed number of time steps.

E. Experimental setup

1) Motor exploration

The sensorimotor data is acquired along 100000 steps of random babbling following a random walk in the joint space with a random direction and a random step size in $[0, \Delta\theta_{\max}]$ with $\Delta\theta_{\max} = .1$ radians. The same $\Delta\theta_{\max}$ is used for the controller. When the joint angle reaches a joint constraint, it is “bounced back” to the opposite direction. Random babbling generated 22085 self-touch configurations on the full body length as shown in Figure 7. The vector of tactile activations of dimension 100 and the vector of proprioceptive activations of dimension 120 are concatenated into the sensorimotor vector of dimension 220 on which the VAE model is trained.

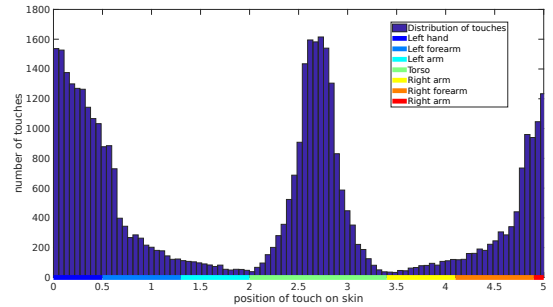


Fig. 7: Distribution of touches after touch random exploration. Some body parts are rarely touched due to joint constraints.

2) VAE model parameters

The VAE architecture is presented in Figure 4. It is composed of an encoder with 3 fully connected hidden layers of size 96, 56, 18 and a 2x3 latent layer for the mean and variance of the posterior distribution $p(\mathbf{z}|\mathbf{x})$. Every layer except the latent layer has a RELU activation function in the output. Following common practice, the decoding layer is a symmetric version of the encoder and outputs the 220 nodes for data reconstruction and a final Logistic activation function normalizing the outputs between $[0, 1]$. The VAE is trained in a denoising framework. The original input is of dimension 22085×220 . Two sets of corrupted data are generated using an additive white Gaussian noise. The amount of noise is computed to obtain a signal-

to-noise ratio of 5. Noise has been added using the `awgn` MATLAB function. The training data set is taken to be 80% of the two sets of original and corrupted data of total size 44170×220 and tested on 20%. The samples presented are randomized and put into minibatches of size 1000. The learning rate is .0001 and was trained for 2000 epochs with the Adam optimizer. Training took 4 hours on a single CPU (Core™ i7-10510U).

3) Neural denoising parameters

The feedforward neural network that decodes proprioceptive activation in joint angles is composed of 4 fully connected hidden layers of sizes 120, 60, 40, 20 with "tansig" activation functions and an output layer of dimension 6 corresponding to the 6 DoF. The outputs are within the $[-1,1]$ interval, which is then scaled to fit the joint limits of the agent. The network is trained on a subset of 10000 randomly chosen proprioception activations using backpropagation and the scaled conjugate gradient method for 10000 epochs using the `train` function in MATLAB on the same CPU.

4) Control loop parameters

The number of latent code samples in the VAE is taken to be $N = 100$. The loop is stopped when the reconstruction criterion is below $\epsilon = .005$; or after 100 time steps.

IV. EXPERIMENTS AND RESULTS

A. Self-touch in the VAE latent space

The VAE learns a latent representation of self-touch configurations in a low-dimensional space—e.g. a 3 dimensional space here. However, it is possible to generate new self-touch postures by generating a latent code, as seen in Figure 8. The

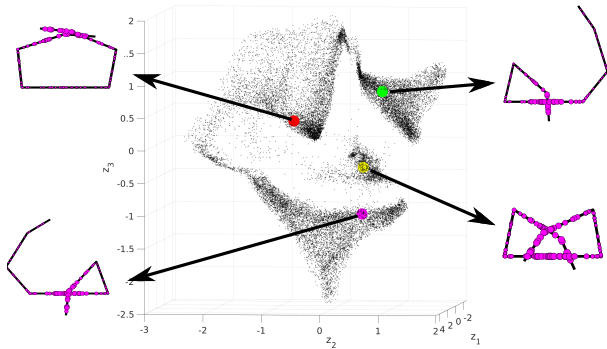


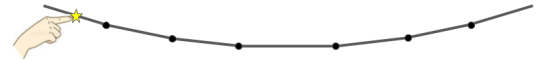
Fig. 8: Self-touch manifold in the latent space. 4 Reconstructed latent VAE postures from the latent code with tactile activations. In the latent space are represented torso touches with the left arm, and the right arm, as well as two hand touches or three touches configurations. To see an animated trajectory in the latent space see the accompanying video <https://youtu.be/IkwBGQ1xBrg>.

latent space is sensitive to the statistics of the data, and thus to the exploration. Therefore, some configurations will be more "expressed" than others in the latent code.

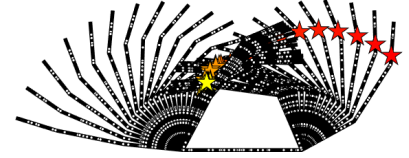
B. Reaching for external touch

The agent is placed in a nontouching configuration and is presented with an external touch on a specific limb in a certain position—similar to the experiments with vibrotactile stimulation of infants [8], [29], [30]—as illustrated in

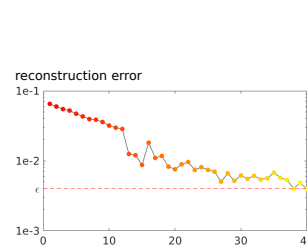
Figure 9a. This stimulation is maintained for the duration of the trial, as would a vibrating mosquito on the skin. Following the control loop, the "best" simulated proprioceptive configuration can be selected according to the online control process presented in Section III-D. During the reaching loop,



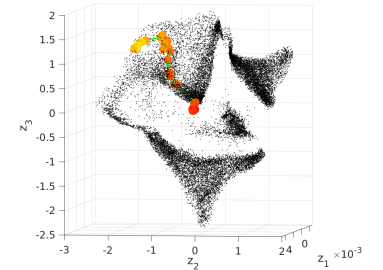
(a) External stimulation.



(b) Example of reaching trajectory. The star is the reaching target.



(c) Reconstruction error during reaching.



(d) Trajectory in latent space.

Fig. 9: Example of reaching for external stimulation.

the target sensorimotor evolves as the agent gets closer to a self-touch configuration. The agent starts with a rough estimate of where it should go and continuously refines its planning. An example of such a process is shown in Figure 10, more self-reaching examples are shown in the accompanying video: <https://youtu.be/IkwBGQ1xBrg>. The reaching trajectory depends on the latent code generated in the VAE so that the final positions are not always identical. Such adaptation can be useful in a context where the agent is blocked with an obstacle and cannot perform certain movements.

In Figure 11 are shown the statistics of reaching on the body for 1000 simulations with random initial postures and random external touch positions. 98% of trials have finished in a self-touch configuration. As expected, reaching is best in areas that have been widely touched, thus mainly trained on, cf. Figure 7. In difficult areas, such as a touch in position 2 or 3.5 on the body, the agent has trouble reaching a configuration generating the correct touch. This is mainly due to the fact that this area is small in the latent code and thus hard to reach. The agent can also be stuck in a local maxima of the reaching criterion. To avoid such problem, it is possible to extend the random search by generating a wider distribution of configurations in the sampling part of the control loop.

V. CONCLUSION, DISCUSSION, AND FUTURE WORK

We have presented a generative model for a simple agent equipped with tactile and proprioceptive sensing that learns from self-touch configurations and develops a latent representation of this combined space. Using a variational autoencoder

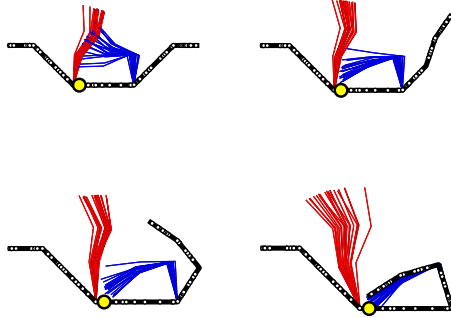


Fig. 10: Variability of generated self-touch configurations during reaching. Given the current skin distance of the contact for 1000 random initial configurations and random external posture (black), 20 generated configurations are shown towards self-touch to the yellow point.

(VAE) in a denoising framework, we have showed how the agent can exploit learned representations to “complete” previously unexperienced stimuli. In particular, by deploying the VAE in an iterative fashion, the agent can generate reaching movements for stimulations on its body.

In this article, we have defined self-touch as the simultaneous activation of touch sensors on different parts of the body. However, such activation can also occur when the agent is exploring the environment, such as touching the ground with multiple fingers or, say, with its back and foot. Such “double touch” events are much less predictable than self-touch events: the infant can touch the environment in widely different ways; different objects will not always be touched with the same fingers, etc. We expect that the proposed model could be applied as a predictive model to distinguish self-touches from double-touches coming from the environment. Due to the lower predictability of these configurations based only on proprioception, they should lead to higher reconstruction errors until the agent is able to use its sight.

An assumption of our model is that our agent learns only in self-touch configurations. That self-touch should be used in learning is compatible with empirical data showing the existence of “body-matching neurons” in the posterior parietal cortex of monkeys (area 5) that respond to specific self-touch configurations and that may drive learning [26]. It may, however, be that non-self-touch configurations are also important, and we will test this in future work by experimenting with less restricted training data.

In robotic models, tactile activation and joint angles are normally treated separately. This model provides a new perspective in that it closely combines tactile and proprioceptive inputs into a single space—inspired by what is known about the brain and the close interactions in somatosensory processing. In fact, this model takes an extreme position in that there is even no separation of the proprioceptive and tactile inputs of the left

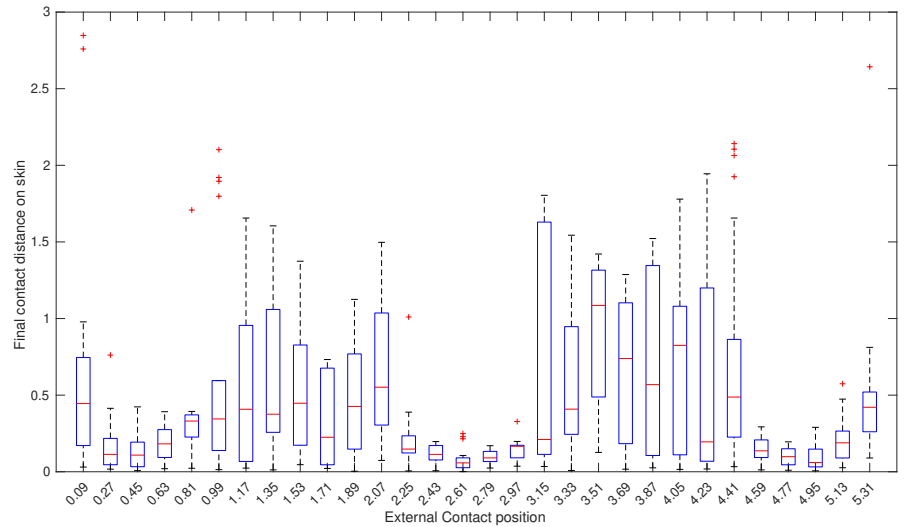


Fig. 11: Statistics of self-reaching error. The box-plots represent the distribution of final touch positions regroupped in bins of width 0.18. The red line is the median and the boxes represent the .25 and .75 percentiles.

and right sides of the body. This departs from what is known about the first stages of somatosensory processing in the brain, but may be in line with representations in the posterior parietal cortex, where lateralization is less pronounced [31].

In the future, self-interaction can be naturally extended to include visual inputs, for example, seeing oneself moving or even seeing oneself being touched (see also [23], [24]). Strategies to reach objects in the field of view could then be obtained with the agent recalling seeing itself in this particular location—such configurations being simulated by the VAE.

The data set on which our model is trained is not ecologically plausible because it is a result of simple motor babbling in joint space. Instead, Blum et al. [25] show that a realistic human movement data set is critical to learning the correct representations of proprioception. In the future, we plan to use movement profiles extracted from infant kinematics during spontaneous activity (see, e.g., [6]). Similarly, the way of generating reaching movements by iterative passes through the VAE is interesting in that the framework is compact and parsimonious and yet it can support reaching behavior. In the future, we will also consider grounding reaching generation in our model in the way done by [32], [33].

Finally, the planar agent used here is still relatively simple. It does have kinematic redundancy, which is an important factor in this context, but overall its morphology is far from an infant. We plan to employ humanoid robots with artificial skin covering their bodies. Self-calibration by self-touch using traditional representations of kinematics and optimization was shown on the iCub [34], with the addition of the visual modality (self-touch and self-observation) in [35]. In [22], the Nao humanoid was used. Using these platforms, we will test the scalability of the approach proposed here and, at the same time, make it possible to use training data, i.e., movement trajectories, extracted from the spontaneous activity of infants (see, e.g., [6]). It is an open question whether

architectures such as VAE will scale or whether some explicit hierarchy or modularization (see [36]) will have to be introduced. Even more biological realism can be introduced by employing musculoskeletal platforms (see [37] for recent work on “musculoskeletal autoencoder”).

ACKNOWLEDGMENT

This work was supported by the Czech Science Foundation (GA CR), project no. 20-24186X. The authors thank Francesco Mannella for sharing and explaining his code related to [21] to us.

REFERENCES

- [1] G. Z. Tau and B. S. Peterson, “Normal development of brain circuits,” *Neuropsychopharmacology*, vol. 35, no. 1, p. 147, 2010.
- [2] R. Khazipov and H. J. Luhmann, “Early patterns of electrical activity in the developing cerebral cortex of humans and rodents,” *Trends in neurosciences*, vol. 29, no. 7, pp. 414–418, 2006.
- [3] A. Stoytchev, “Some basic principles of developmental robotics,” *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 2, pp. 122–130, 2009.
- [4] M. Hoffmann, “The role of self-touch experience in the formation of the self,” in *The Development of the Self Workshop at IEEE ICDL-EpiRob 2017*, 2017.
- [5] A. Piontelli, *Development of Normal Fetal Movements; The Last 15 Weeks of Gestation*. Italy: Springer-Verlag, 2015.
- [6] A. DiMercurio, J. P. Connell, M. Clark, and D. Corbetta, “A naturalistic observation of spontaneous touches to the body and environment in the first 2 months of life,” *Frontiers in Psychology*, vol. 9, p. 2613, 2018.
- [7] P. Rochat, “Self-perception and action in infancy,” *Exp Brain Res*, vol. 123, pp. 102–109, 1998.
- [8] M. Hoffmann, L. K. Chinn, E. Somogyi, T. Heed, J. Fagard, J. J. Lockman, and J. K. O’Regan, “Development of reaching to the body in early infancy: From experiments to robotic models,” in *Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2017, pp. 112–119.
- [9] J. K. O’Regan and A. Noe, “A sensorimotor account of vision and visual consciousness,” *Behavioral and Brain Sciences*, vol. 24, pp. 939–1031, 2001.
- [10] D. P. Kingma and M. Welling, “Stochastic gradient vb and the variational auto-encoder,” in *Second International Conference on Learning Representations, ICLR*, vol. 19, 2014, p. 121.
- [11] M. Zambelli, A. Cully, and Y. Demiris, “Multimodal representation models for prediction and control from partial information,” *Robotics and Autonomous Systems*, vol. 123, p. 103312, 2020.
- [12] E. Bizzi, N. Hogan, F. A. Mussa-Ivaldi, and S. Giszter, “Does the nervous system use equilibrium-point control to guide single and multiple joint movements?” *Behavioral and brain sciences*, vol. 15, no. 4, pp. 603–613, 1992.
- [13] A. G. Feldman, “Once more on the equilibrium-point hypothesis (λ model) for motor control,” *Journal of motor behavior*, vol. 18, no. 1, pp. 17–54, 1986.
- [14] M. Hoffmann, H. Marques, A. Hernandez Arieta, H. Sumioka, M. Lungarella, and R. Pfeifer, “Body schema in robotics: A review,” *Autonomous Mental Development, IEEE Transactions on*, vol. 2, no. 4, pp. 304–324, Dec 2010.
- [15] P. Lanillos, E. Dean-Leon, and G. Cheng, “Enactive self: a study of engineering perspectives to obtain the sensorimotor self through enaction,” in *2017 Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, 2017, pp. 72–78.
- [16] P. D. Nguyen, Y. K. Georgie, E. Kayhan, M. Eppe, V. V. Hafner, and S. Wermter, “Sensorimotor representation learning for an “active self” in robots: a model survey,” *KI-Künstliche Intelligenz*, vol. 35, no. 1, pp. 9–35, 2021.
- [17] G. Schillaci, V. V. Hafner, and B. Lara, “Exploration behaviors, body representations, and simulation processes for the development of cognition in artificial agents,” *Frontiers in Robotics and AI*, vol. 3, 2016.
- [18] Y. Yamada, H. Kanazawa, S. Iwasaki, Y. Tsukahara, O. Iwata, S. Yamada, and Y. Kuniyoshi, “An embodied brain model of the human foetus,” *Scientific Reports*, vol. 6, 2016.
- [19] V. Y. Roschin, A. A. Frolov, Y. Burnod, and M. A. Maier, “A neural network model for the acquisition of a spatial body scheme through sensorimotor interaction,” *Neural computation*, vol. 23, no. 7, pp. 1821–1834, 2011.
- [20] V. Marcel, S. Argentieri, and B. Gas, “Building a sensorimotor representation of a naive agent’s tactile space,” *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 2, pp. 141–152, 2016.
- [21] F. Mannella, V. G. Santucci, E. Somogyi, L. Jacquey, K. J. O’Regan, and G. Baldassarre, “Know Your Body Through Intrinsic Goals,” *Frontiers in Neurobotics*, vol. 12, p. 30, 2018.
- [22] F. Gama, M. Shcherban, M. Rolf, and M. Hoffmann, “Goal-directed tactile exploration for body model learning through self-touch on a humanoid robot,” *IEEE Transactions on Cognitive and Developmental Systems*, 2021.
- [23] S. Fuke, M. Ogino, and M. Asada, “Body image constructed from motor and tactile images with visual information,” *International Journal of Humanoid Robotics*, vol. 4, no. 02, pp. 347–364, 2007.
- [24] Y. Yoshikawa, K. Hosoda, and M. Asada, “Cross-anchoring for binding tactile and visual sensations via unique association through self-perception,” in *Proceedings of the Third International Conference on Development and Learning (ICDL)*, 2004.
- [25] K. P. Blum, M. D. Grogan, Y. Wu, A. J. Harston, L. E. Miller, and A. A. Faisal, “Predicting proprioceptive cortical anatomy and neural coding with topographic autoencoders,” *bioRxiv*, 2021.
- [26] H. Sakata, Y. Takaoka, A. Kawarasaki, and H. Shibutani, “Somatosensory properties of neurons in the superior parietal cortex (area 5) of the rhesus monkey,” *Brain research*, vol. 64, pp. 85–102, 1973.
- [27] D. J. Rezende, S. Mohamed, and D. Wierstra, “Stochastic backpropagation and approximate inference in deep generative models,” in *Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32*, ser. ICML’14. JMLR.org, 2014, p. II–1278–II–1286.
- [28] Z. Xu, W. Wu, S. S. Winter, M. L. Mehlman, W. N. Butler, C. M. Simmons, R. E. Harvey, L. E. Berkowitz, Y. Chen, J. S. Taube, A. A. Wilber, and B. J. Clark, “A comparison of neural decoding methods and population coding across thalamo-cortical head direction cells,” *Frontiers in Neural Circuits*, vol. 13, 2019.
- [29] L. K. Chinn, C. F. Noonan, M. Hoffmann, and J. J. Lockman, “Development of infant reaching strategies to tactile targets on the face,” *Frontiers in psychology*, p. 9, 2019.
- [30] E. Somogyi, L. Jacquey, T. Heed, M. Hoffmann, J. J. Lockman, L. Granjon, J. Fagard, and J. K. O’Regan, “Which limb is it? responses to vibrotactile stimulation in early infancy,” *British Journal of Developmental Psychology*, vol. 36, no. 3, pp. 384–401, 2018.
- [31] A. M. Seelke, J. J. Padberg, E. Disbrow, S. M. Purnell, G. Recanzone, and L. Krubitzer, “Topographic maps within brodmann’s area 5 of macaque monkeys,” *Cerebral Cortex*, vol. 22, no. 8, pp. 1834–1850, 2012.
- [32] M. V. Butz, O. Herbort, and J. Hoffmann, “Exploiting redundancy for flexible behavior: unsupervised learning in a modular sensorimotor control architecture,” *Psychological Review*, vol. 114, no. 4, p. 1015, 2007.
- [33] P. Cisek, “Cortical mechanisms of action selection: the affordance competition hypothesis,” *Philos Trans R Soc Lond B Biol Sci*, vol. 362, no. 1485, pp. 1585–1599, Sep. 2007.
- [34] A. Roncone, M. Hoffmann, U. Pattacini, and G. Metta, “Automatic kinematic chain calibration using artificial skin: self-touch in the iCub humanoid robot,” in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, 2014, pp. 2305–2312.
- [35] K. Stepanova, T. Pajdla, and M. Hoffmann, “Robot self-calibration using multiple kinematic chains – a simulation study on the iCub humanoid robot,” *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1900–1907, 2019.
- [36] S. Ehrenfeld, O. Herbort, and M. V. Butz, “Modular neuron-based body estimation: maintaining consistency over different limbs, modalities, and frames of reference,” *Frontiers in Computational Neuroscience*, vol. 7, p. 148, 2013.
- [37] K. Kawaharazuka, A. Miki, Y. Toshimitsu, K. Okada, and M. Inaba, “Adaptive body schema learning system considering additional muscles for musculoskeletal humanoids,” *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 3459–3466, 2022.