Building the Representation of an Agent Body from its Sensorimotor Invariants

Valentin Marcel, Boris Garcia, Sylvain Argentieri and Bruno Gas

Abstract— **Making systems able to autonomously adapt themselves to changes in their own body or in their environment is still a challenging task questioning a lot of different scientific communities. Many works propose either sophisticated adaptive model-based or learning-based techniques, as a solution. Most of them rely on the traditional perceive/decide/act framework, inspired by our human intuition about how we perceive the world. But recent contributions have shown that it is possible for an agent to discover the structure of its interaction with the environment or its own body via the so-called sensorimotor flow. This work is rooted in this paradigm, and a method for the building of an internal representation of the agent body is proposed. Importantly, it does not require any a priori knowledge nor model. A careful mathematical formalization is outlined, together with simulations demonstrating the effectiveness of the approach.**

I. Introduction

Can a mobile agent perceive its environment without any model nor on its environment neither on itself? More and more researchers from different scientific fields (psychophysics, artificial intelligence, philosophy, ...) raise this question, which is of particular importance when dealing with autonomous mobile robotics. Indeed, using *a priori* models of the environment can make robotic systems really fast, efficient and robust when executing complex predefined tasks. But in the same time such systems may have difficulties to behave autonomously, namely to adapt themselves to unknown environments that have not been modeled or learned before –except by considering that an universal model could be obtained from a learning procedure, as proposed by the Bayesian Perception Theory [1]–.

The proposed approach follows a different path, paved by Poincaré more than 100 years ago [2], [3]. In this line of research, what is called *perception* is not an innate capacity. It is something which is learned and can not be separated from motor action. Indeed, this *sensorimotor* flow carries fundamental informations about the *external space* and its geometry [4]. For instance, Poincaré states that the external space dimension can be extracted from it by an agent endowed with several motor degrees of freedom and sensors whose outputs depend on the 3D position of the system. More generally, the sensorimotor flow indeed reveals some correlations, i.e. *sensorimotor contingencies* [5], which are invariant and carry information about the external environment of the agent. Philipona et al. proposed a mathematical formulation of this idea, but not by using proprioceptive signals as suggested by Poincaré, but by exploiting sensor outputs signals instead [6]. The demonstration is based on the study of the sensory manifold and on its dimension. On this basis, the authors have been able to prove Poincaré's intuitions but only when using infinitesimal movement amplitudes. Such a limitation –originating from the use of standard linear mathematical tools– does not allow any experimental validation of the approach. Bootstrapping methods have been proposed to solve this problem [7]. For instance, Laflaquière et al. succeeded in extending the approach to much more realistic movement amplitudes by coupling a motor bootstrap technique with a Curvilinear Component Analysis (CCA) [8]. But all these approaches were criticized by Frolov in [9], claiming that they require a stable external environment during the exploration. Indeed, these works postulate some *a priori* on the environment that cannot be obtained by the sensory system only. As a solution, Frolov introduced a body endowed with tactile sensors and a mobile arm. While proprioceptive signals still encode the arm movements, the arm end-effector has to touch the body itself so as to ensure a stable perception, thus obtained without any prior hypothesis on the environmental state. Following another way, Laflaquière et al. have shown that, beyond the dimension of space, it is also possible to obtain an external space representation by using appropriate partitions of the motor space [10], resulting in a much more *motor oriented* than *sensor oriented* framework. In this contribution, each end-effector position of an arm is represented by the subset of the motor configurations letting invariant the sensory state, the so-called *kernel manifolds*. But the limitations pointed out by Frolov apply here again, as a stable environmental state is still required. More generally, the learning of sensorimotor relationships –be it for obtaining a representation of the agent body and/or the environment– has received more attention recently. For instance, an online learning framework for the building of low-dimensional sensorimotor maps is proposed in [11]. Sensorimotor embedding is also proposed in [12] as a new approach to the dimensionality reduction problem faced by many works when trying to extract spatial/geometric knowledge from the raw sensorimotor flow.

This paper is focused on the extension of Laflaquière approach to the building of an internal representation of an agent body, in the vein of Frolov's previous work. The kernel manifolds used in [10] will be precisely mathematically formalized, and some proofs relying on the isomorphism theorems will demonstrate the relevance of this framework for the building of such a representation. It will then be shown that an internal representation of an agent body can be obtained via the matching of proprioceptive and tactile signals using a CCA technique. The paper is organized as follows. §II is devoted to the mathematical foundations of the kernel manifolds defined in [10]. Next, §III extends

The authors are with Sorbonne Universités, UPMC Univ. Paris 06, UMR 7222, ISIR, and CNRS, UMR 7222, ISIR, F-75005 Paris, France.
Email: {marcel,argentieri,gas}@isir.upmc.fr

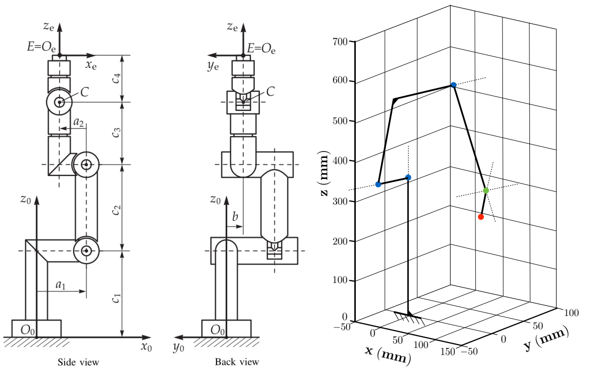

Fig. 1: Used robotic system. It is made of $M = 5$ degrees of freedom, each of them being controlled via a motor command m_i, $i = \{1, \ldots, M\}$ and moving the end-effector in the $3D$ space. Blue dots represent revolute joints, green dot depicts spherical joint, and red dot sketches for the end-effector.

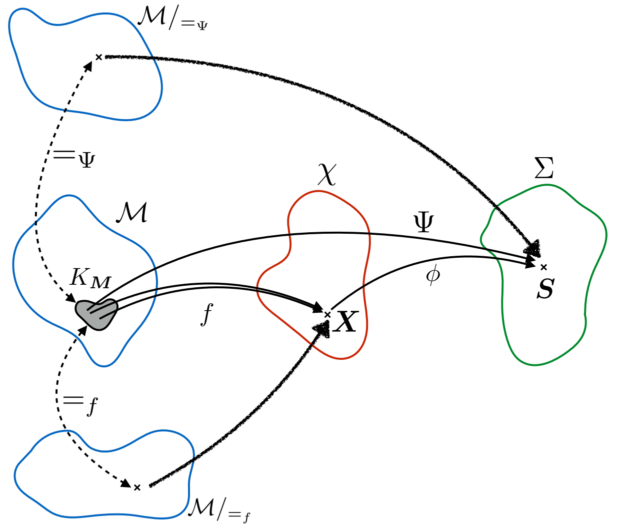

Fig. 2: Schematic representation of the different spaces involved in the paper. $\mathcal{M}$ is the motor space, χ is the working space, Σ is the sensory space, $\mathcal{M}/_{=_f}$ and $\mathcal{M}/_{=_\Psi}$ both represent the internal representation of χ and Σ respectively.

this mathematical formalization to the case of an agent endowed with a tactile-like sensor, and proposes a simulation setup and results which illustrate the building of an internal representation of the agent body. Then, a short discussion is proposed in §IV. Finally, a conclusion ends the paper.

II. Representation of an agent's working space

Laflaquière showed in [13] through simulations how a partition of the motor space, i.e. the aforementioned kernel manifolds, can be used to represent the position of a robotic arm end-effector in the 3D space. In this section, a careful mathematical formalization of these kernel manifolds is provided through the introduction of equivalence classes, on which it becomes possible to apply the isomorphism theorems. This formalization will then serve as the main support throughout this article to show how to obtain an internal representation of the body without a priori information.

A. Mathematical foundations

As a first step, this subsection is devoted to the formalization of the building of an internal representation of an agent's working space. For the sake of simplicity, an arm-type robotic agent compelled to move in the 3D space will be used to illustrate the demonstration, while not being limited to.

1) Notations: The proposed robotic system, whose kinematics is inspired by traditional industrial robots, is composed of moving rigid parts, each of them being connected to each other via a joint parameterized by a scalar $m_i \in \mathbb{R}$, $i = \{1, \ldots, M\}$, see Fig. 1. The motor configuration of the agent is entirely described by the vector $\boldsymbol{M} = (m_1, m_2, \ldots, m_M)^T \in \mathcal{M}$, with $\mathcal{M} \in \mathbb{R}^M$ the so-called *motor configuration space*, and $.^T$ the transpose operator. Thanks to these motor commands, the end-effector of the system is able to move in the so-called *working space* $\chi \in \mathbb{R}^3$. In all the following, let's denote $\boldsymbol{X} = (x_1, x_2, x_3)^T \in \chi$ the end-effector position vector. Thanks to the forward kinematics model of the system $f(.)$, the end-effector position $\boldsymbol{X}$ can be computed from the motor configuration $\boldsymbol{M}$, i.e. $\boldsymbol{X} = f(\boldsymbol{M})$. Due to the possible agent's redundancy, f can be surjective, as different motor configurations can lead to the same end-effector position.

2) Demonstration: On this basis, one can now define the equivalence relation $=_f$ between any pair of elements $(\boldsymbol{M}_1, \boldsymbol{M}_2)$ of $\mathcal{M}$ with

$$\boldsymbol{M}_1 =_f \boldsymbol{M}_2 \Leftrightarrow \exists \boldsymbol{X} \in \mathcal{X}, f(\boldsymbol{M}_1) = f(\boldsymbol{M}_2) = \boldsymbol{X}. \quad (1)$$

Proposition 1. *The quotient set of $\mathcal{M}$ with respect to $=_f$ is isomorphic to χ.*

Proof. It follows that $=_f$ is by definition the equivalence kernel of f. For any $\boldsymbol{M} \in \mathcal{M}$ such that $f(\boldsymbol{M}) = \boldsymbol{X}$, let's now denote by $K_{\boldsymbol{M}}$ the equivalence class of $\boldsymbol{M}$, with

$$K_{\boldsymbol{M}} = \{\boldsymbol{R} \in \mathcal{M} \mid \boldsymbol{R} =_f \boldsymbol{M}, \forall \boldsymbol{R} \in \mathcal{M}\}. \quad (2)$$

$K_{\boldsymbol{M}}$ can then be modded out to form the quotient set $\mathcal{M}/_{=_f}$, with

$$\mathcal{M}/_{=_f} = \{K_{\boldsymbol{M}} \mid \boldsymbol{M} \in \mathcal{M}\}. \quad (3)$$

According to the first isomorphism theorem, if f is surjective (i.e. if the agent is redundant), then $\mathcal{M}/_{=_f}$ is isomorphic to χ. □

3) Interpretation: Figure 2 highlights the relationships between $\mathcal{M}$, χ and $\mathcal{M}/_{=_f}$. Regrouping all the motor commands leading to the same end-effector position $\boldsymbol{X}$ in the same set $K_{\boldsymbol{M}}$ defines a so-called *motor kernel manifold* (gray area). Each kernel manifold can then be conceptually represented in $\mathcal{M}/_{=_f}$ by a point thanks to the equivalence relation $=_f$ (dotted line), each of them being also linked to an end-effector position $\boldsymbol{X}$ (brushed lines). Then, according to the previous subsection, it appears that $\mathcal{M}/_{=_f}$ is isomorphic to χ. In other terms, $\mathcal{M}/_{=_f}$ can be used by the agent as an internal representation of its working space χ.

B. Discussion

Proposition 1 states that the quotient set $\mathcal{M}/_{=_f}$ is an excellent candidate for representing the working space $\mathcal{X}$. But $\mathcal{M}/_{=_f}$ can only be built by regrouping all the motor commands leading to same end-effector position. How can the agent know that its end-effector position is not actually changing when it moves? Indeed, the naive agent has only access to its *sensorimotor flow*, made of its motor commands and exteroceptive data originating from sensors possibly equipping the system. The question is then: is there any other way to obtain an internal representation of $\mathcal{X}$ based on this sensorimotor flow? The authors have shown in a previous contribution that this can be achieved thanks to the introduction of a new kernel set which trades "invariant end-effector position" for "invariant sensation" [10]. This idea has been successfully assessed in simulations with a visually-inspired modality, while not being mathematically formalized. This will be extended in the next section, which will be illustrated with a tactile-inspired sensory modality. It will then make the agent able to build –without any a priori– an internal representation of a subset of its working space: its own body.

III. Tactile representation of an agent's body

The previous section was devoted to the building of the internal representation of an agent's working space. As already argued, it is postulated that this internal representation can be discovered by an agent endowed with adequate sensing capabilities. Among others, the focus is put on the tactile modality in the following, while not being necessarily limited to. It will be proved in the first subsection that this will allow an agent endowed with an arm and a body covered with tactile receptive fields to build an internal representation of its own body. Then, the agent setup used in simulation to illustrate the building of this internal representation is detailed in a second subsection. Finally, simulations results are provided in the last subsection.

A. Dealing with sensory inputs

The agent described in the previous section is deprived from sensory inputs, and is thus unable to form any representation of its working space by itself. Consequently, the existence of $\mathcal{M}/_{=_f}$ is not relevant to the agent, since it will not be able to build it without any a priori knowledge. Suppose now that the system is endowed with sensing capabilities, and is thus able to have access to a sensation vector $\boldsymbol{S} = (s_1, s_2, \dots, s_S)^T \in \mathcal{S}$, with $\mathcal{S} \in \mathbb{R}^S$ the *sensory space*. For any vector $\boldsymbol{X} \in \chi$, one can obtain the sensation vector $\boldsymbol{S}$ through the *forward sensory function* $\phi(.) : \mathcal{X} \to \mathcal{S}$ so that $\boldsymbol{S} = \phi(\boldsymbol{X})$. Importantly, by further restricting $\mathcal{S}$ to $\Sigma = \mathrm{Im}\,(\phi)$, ϕ can be rendered surjective from χ to Σ. In other words, Σ is the set of all "physically" reachable sensory states, and will thus also be called the *sensory space*. As illustrated in Fig. 2, a given motor configuration $\boldsymbol{M}$ can now be associated to a sensation vector $\boldsymbol{S}$ thanks to the application $\Psi = \phi \circ f$. On this basis, and with the same reasoning as in §II-A, the equivalence kernel of Ψ can be formed along

$$\boldsymbol{M}_1 =_\Psi \boldsymbol{M}_2 \Leftrightarrow \Psi(\boldsymbol{M}_1) = \Psi(\boldsymbol{M}_2). \quad (4)$$

Proposition 2. *The quotient set of $\mathcal{M}$ with respect to $=_\Psi$ is isomorphic to Σ.*

Proof. Lets denote $\widetilde{K}_{\boldsymbol{M}}$ the equivalent class of $\boldsymbol{M}$, i.e.

$$\widetilde{K}_{\boldsymbol{M}} = \{\boldsymbol{R} \in \mathcal{M} \mid \boldsymbol{R} =_\Psi \boldsymbol{M}, \forall \boldsymbol{R} \in \mathcal{M}\}, \quad (5)$$

which can be modded out to form the quotient set $\mathcal{M}/_{=_\Psi}$

$$\mathcal{M}/_{=_\Psi} = \left\{\widetilde{K}_{\boldsymbol{M}} \mid \boldsymbol{M} \in \mathcal{M}\right\}. \quad (6)$$

Again, thanks to the first isomorphism theorem, if Ψ is surjective, then Σ is isomorphic to $\mathcal{M}/_{=_\Psi}$. □

Importantly, one can show that if ϕ is also injective, then $\mathcal{M}/_{=_f}$ and $\mathcal{M}/_{=_\Psi}$ also become isomorphic. In other words, since $\mathcal{M}/_{=_f}$ is isomorphic to the working space χ, then $\mathcal{M}/_{=_\Psi}$ is also isomorphic to χ. The authors have already exploited this property in a previous work dealing with the learning of the agent spatial configuration [10]. In this past work, the forward sensory function $\phi(.)$ was representing a retina-like sensor whose outputs were sensitive to lights placed in the vicinity of the robots, thus making the agent able to build an internal representation of $\mathcal{X}$. This present work is more concerned with a tactile-like modality, which will be used by the agent to build an internal representation of a subset $\mathcal{X}_b$ of $\mathcal{X}$: its own body. Considering the physics of this modality, it is clear that $\phi(.)$ will be defined as a surjective function. Indeed, it is obvious that $\forall \boldsymbol{X} \notin \mathcal{X}_b, \phi(\boldsymbol{X}) = \boldsymbol{0}$. Working with the restriction $\phi_{|\mathcal{X}_b} : \mathcal{X}_b \to \Sigma^*$ allows then to form a bijective sensory function, since it seems reasonable to hypothesize that two different end-effector positions on the agent body lead to two different sensory vectors. Consequently, with $\psi = \phi_{|\mathcal{X}_b} \circ f$, the quotient set $\mathcal{M}/_{=_\Psi}$, –which is built by considering all the motor commands leading to the same sensory vector– can be used to represent any vector $\boldsymbol{X} \in \chi_b$, thus making the agent able to build a representation of its own body.

B. Simulation setup and algorithm

A simulated robotic system (i.e. the agent) is used in all the following to illustrate 1/how $\mathcal{M}/_{=_\Psi}$ can be built from the sensorimotor flow of the simulated agent, and 2/how $\mathcal{M}/_{=_\Psi}$ can be used to represent its body. The agent is made of a spherical body covered with S tactile receptive fields and a multi-DoF arm identical to the one used in §II (see Fig. 1). The arm basis and the spherical body are sticked together, so that only the arm is able to move according its $M = 5$ degrees of freedom. Importantly, the entire approach does not require any *a priori*, so that the agent does not have access to its forward kinematics model $f(.)$. Consequently, the only naive operation it can conduct consists in generating P random motor commands $\boldsymbol{M}^{(p)} = (m_1^{(p)}, \dots, m_M^{(p)})^T$, each of them being associated with a perception $\boldsymbol{S}^{(p)} = (s_1^{(p)}, \dots, s_S^{(p))})^T$, $p = 1, \dots, P$.

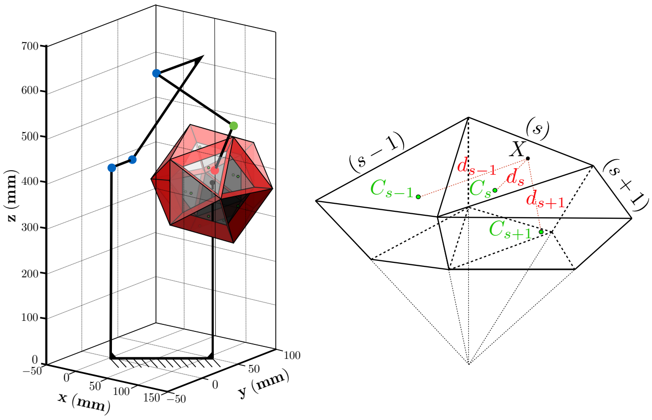

Fig. 3: Agent setup. (Left) The arm touches at $\boldsymbol{X} \in \chi_b$ the spherical body endowed with S tactile receptive fields, thus producing the sensory vector $\boldsymbol{S} \in \Sigma^*$. (Right) Surface of the body; each sensory vector component s_s depends on the distance between the contact point $\boldsymbol{X}$ and the s^{th} sensitive field center $\boldsymbol{C}_s$, $s = 1, \ldots, S$. Note that the simulated receptive field depth is not infinitesimal, each of them having the same non-zero depth.

1) Rough random motor sampling: In all the following, the agent explores its motor space by using a random walk model applied to each component of the p^{th} motor vector $\boldsymbol{M}^{(p)}$, i.e.

$$m_i^{(p)} = \mod\left(m_i^{(p-1)} + \mu^{(p)} + \pi, 2\pi\right) - \pi, i = 1, \ldots, M, \tag{7}$$

with $\mu^{(p)}$ a realization of the random variable $\mu \sim \mathcal{N}(0, \sigma)$. A total of L trajectories, sampled with V values, are computed for each motor component. Eq. (7) is reset every V iterations to start a new random motor exploration, see §III-B.6.

2) Sensory vector generation: Each of the $P = V \times L$ motor configuration $\boldsymbol{M}^{(p)}$ of the agent is associated to a position $\boldsymbol{X}^{(p)}$ of the end-effector, together with a perception $\boldsymbol{S}^{(p)}$ along $\boldsymbol{S}^{(p)} = \Psi(\boldsymbol{M}^{(p)})$, whose i^{th} sensory vector component is computed along

$$s_i^{(p)} = \phi_i(\boldsymbol{X}^{(p)}) = \begin{cases} 0, \text{ if } \boldsymbol{X}^{(p)} \text{is not on the body,} \\ \exp\left(-K \dfrac{\|\boldsymbol{X}^{(p)} - \boldsymbol{C}_i\|_2}{d_{\text{body}}}\right) \text{otherwise,} \end{cases} \tag{8}$$

with $i = 1, \ldots, S$. $\boldsymbol{C}_i$ denotes the center of the i^{th} sensitive field, K and d_{body} being a normalization constant and the spherical body diameter respectively. Of course, only a subset of the P sensation vectors are different from $\boldsymbol{0}$. Additionally, multiple motor configurations are related to the same non-null sensory vectors because of the agent redundancy. But the random motor configuration sampling prevents the agent to *exactly* obtain these sets of motor commands without any a priori. Consequently, *very close* sensory vectors will be regrouped together up to a certain threshold, so as to estimate the motor quotient set.

3) Kernel space sampling: Let's now select N so-called *target* sensory vectors $\boldsymbol{S}_i^*$, $i = 1, \ldots, N$ among all the R non-zero vectors $\boldsymbol{S}^{*(r)} \in \{S^{(1)}, \ldots, S^{(P)}\} \subset \Sigma^*$, $r = 1, \ldots, R$. The N most distant vectors from each others are selected here in order to efficiently represent the sensory space Σ^*, see III-B.6. One can then form the N sets S_i regrouping all the $\boldsymbol{S}_i^*$ neighbors

$$S_i = \left\{\boldsymbol{S}^{*(r)}, r = 1, \ldots, R \mid d(\boldsymbol{S}^{*(r)}, \boldsymbol{S}_i^*) < \delta/2\right\}, \tag{9}$$

with $d(.)$ the Euclidean distance between two vectors and δ a threshold. Importantly, δ must be selected so as to avoid any intersection between all S_i, i.e.

$$\delta \leq \min_{i,j \in [1,N]^2} d(\boldsymbol{S}_i^*, \boldsymbol{S}_j^*). \tag{10}$$

On this basis, one can then also form the N kernel sets M_i regrouping all the motor commands leading to a sensory vector in S_i, with

$$M_i = \left\{\boldsymbol{M}^{*(r)}, r = 1, \ldots, R \mid \boldsymbol{S}^{*(r)} = \Psi(\boldsymbol{M}^{*(r)}) \in S_i\right\}. \tag{11}$$

M_i allows then to define an approximation $\widehat{\mathcal{M}}/_{=_\Psi}$ of the quotient set $\mathcal{M}/_{=_\Psi}$ along

$$\widehat{\mathcal{M}}/_{=_\Psi} = \{M_i, i = 1, \ldots, N\}. \tag{12}$$

With $\widehat{\mathcal{M}}/_{=_\Psi}$, the agent has now built an internal representation of its own body. While not being mandatory for the agent, $\widehat{\mathcal{M}}/_{=_\Psi}$ can also be projected to a lower dimension so as to be able to interpret this representation. This is performed thanks to a final CCA step.

4) Metric computations: The internal representation space $\widehat{\mathcal{M}}/_{=_\Psi}$ contains the topological information about the agent's body. A metric can be defined in this internal space so that data can be projected in a low dimensional space with CCA. Let's take two points U_i and U_j in this internal representation, both of them being related to the kernel sets M_i and M_j respectively. The distance ρ between these 2 points in $\widehat{\mathcal{M}}/_{=_\Psi}$ can be computed with

$$\rho(U_i, U_j) = \min_{\substack{\boldsymbol{M}^{*(r)} \in M_i \\ \boldsymbol{M}^{*(s)} \in M_j}} \mathrm{d}(\boldsymbol{M}^{*(r)}, \boldsymbol{M}^{*(s)}). \tag{13}$$

This distance is not a proper metric, as the triangle inequality does not hold. But experiments show that this simple distance definition is sufficient to capture kernel sets distances in practice.

5) CCA projection: CCA is a non-linear projection technique which will be used to represent in a low dimensional space every points in $\widehat{\mathcal{M}}/_{=_\Psi}$ by preserving the topology of the underlying manifold. In this paper, the projection will be performed in a $3D$-space, which is the smallest dimension required to preserve the representation of the body. Again, this projection is only performed here to visualize and interpret the internal representation built by the agent. Consequently, considering a $3D$ projection does not bring any a priori in the system. For more details on the CCA method, a complete description of the algorithm can be found in [14] and an example application by the authors in [8].

6) Algorithm: All the previous steps are summarized in the Algorithm 1. This algorithm will be now used to obtain the internal representation of the agent body for different cases in the next subsection.

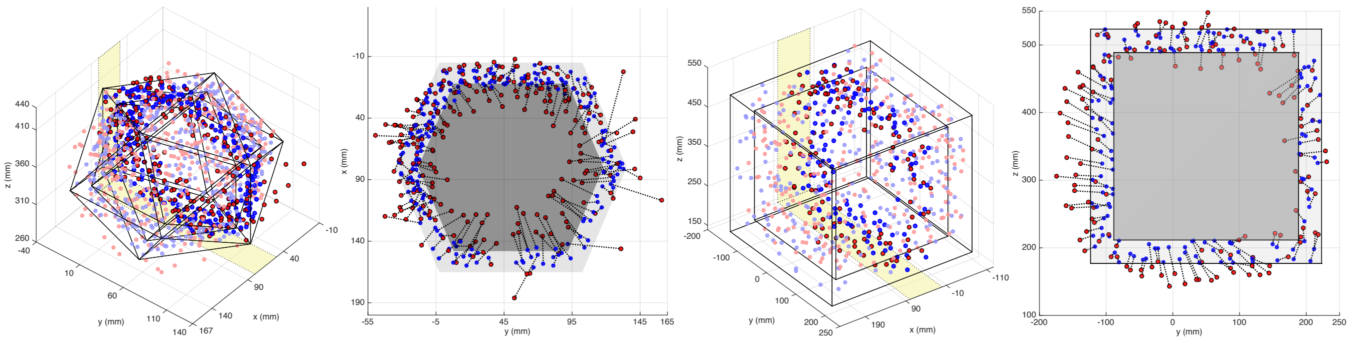

Fig. 4: Illustration of the agent internal representation for the N target sensations in the robotic $3D$ frame. (Red) transformed internal representation of N target sensations by the agent, (blue) original end-effector points that led to the target sensations. From the left to the right: (i) spherical body with an highlighted slice of data, (ii) links between transformed internal representation and real positions of target sensations on the body in a (xy) plan cut of the spherical body, (iii) cubic body with an highlighted slice of data, (iv) (yz) plan cut of the cubic body.

C. Simulations and results

1) Spherical body: In this first scenario, the agent body is made of $S = 20$ tactile receptive fields, thus forming an icosahedral body with a diameter $d_{\text{body}} = 100$mm. The normalization constant in Eq. (8) has been set to $K = 20$. The body center is placed at position $C = [80, 50, 350]$ (to avoid any effect of symmetrization) in the frame $B_0 = (O_0, x_0, y_0, z_0)$ centered at the root basis of the robotic arm, see Fig. 3. The rough motor sampling step is performed with $L = 10$ trajectories of length $V = 10^6$ samples with a standard deviation parameter $\sigma = 0.1$, while $N = 1000$ target sensory vectors $\boldsymbol{S}_i^*$ distant from a minimal euclidean distance of $\delta = 40.10^{-3}$ are selected.

With all these parameters, Algorithm 1 is run and then produces a $3D$ representation of the agent body. While visualizing this CCA projection could be sufficient to interpret the consequent representation, an additional isometric transformation is applied to the CCA output so as to represent this projection in the set $\mathcal{X}$. This will allow to directly compare the N target positions $\boldsymbol{X}_i^*$ such that $\phi(\boldsymbol{X}_i^*) = \boldsymbol{S}_i^*$, with the N transformed CCA outputs. In practice, the isometric transformation relies on the Horn's quaternion-based method [15], which aims at minimizing the least square error between the target positions and the transformed CCA outputs. Importantly, it is postulated that the resulting RMSE between these two sets of representation can be used to measure the quality of the obtained internal representation.

Figure 4 exhibits the simulation results. Lets focus on the first two left subfigures for now, in which blue and red points represent the N target end-effector positions $\boldsymbol{X}_i^*$ and the N isometrically-transformed CCA outputs respectively. The first subfigure exhibits the entire icosahedral agent body, while the second one focuses on the yellow slice of it so that information about thickness and depth of the body can be seen. The blue and red points can be linked by pair as they both represent the same target position: one by its real position on the body (blue), and the other one by its (projected and transformed) internal representation (red). These links are represented by dotted lines in the second subfigure. As shown in Figure 4, the two sets of point are relatively closed to each other, which means that the agent is able to capture the body geometry by using only sensorimotor information. This is confirmed via the small RMSE which is as low as 17.6mm, representing a relative error of 17.6% w.r.t. the body size. This error can be easily lowered by using smallest target neighborhoods, but at the cost of a higher computational cost. Indeed, a small δ value will lead to a lower collection of sensations and motor commands, which can be countered by a finest –but computationally longer– random exploration by the agent.

2) Cubic body: In order to illustrate the genericity of the proposed approach, a second form of the agent body is used. This time, $S = 6$ tactile receptive fields are used together to form a body with a cube shape of length $d_{\text{body}} = 300$mm. The normalization constant in Eq. (8) is now set to $K = 0.2$. The body center position, together with the number of trajectories L, the samples number V and standard deviation σ remain unchanged, while δ has been now set to $\delta = 6.10^{-3}$. All the remaining steps in Algorithms 1 are the same, together with the CCA isometric transformation mentioned it the first simulation. The results are shown in the two right subfigures in Figure 4, which again exhibits the target points on the body (blue) and those obtained by the construction of the internal representation (red).It can be seen that the cubic shape has been relatively well approximated, with a RMSE of 32.68 mm representing 10.9% of the cubic body size.

IV. Discussions

The proposed approach allows to obtain an internal map of the agent body from the sensorimotor flow. But why would the agent build such a map? From the authors point of view, this map is the first step towards the ability to plan a movement of the agent arm on the body. Indeed, first very preliminary results indicate that this can be achieved through very naive interpolation approaches, i.e. one can infer the motor commands to be applied so as to reach a target sensation never felt by the agent before. In other terms, the proposed internal representation can be generative. But one have to keep in mind that the isomorphism theorems exploited all along the paper do not provide any

Algorithm 1 Generation of the internal representation

L: random walk trajectories number
V: trajectories length
N: number of target sensations on the body
$\mathbf{S}_i^*$: i^{th} target sensation
δ: minimum distance separating target sensations
M_i: kernel set for target sensation $\mathbf{S}_i^*$
σ: max step size for random exploration
d($\cdot$,$\cdot$): Euclidean distance
$\rho(\cdot,\cdot)$: distance between kernel sets
$\Psi(\cdot)$: sensorimotor function

Require: N, δ, L, σ, V

{Rough random motor sampling}
r=1;
for $l = 1 : L$ **do**
  % Initialization with a random motor configuration
  $\mathbf{M}^{(0)} = (m_1^{(0)}, \cdots, m_M^{(0)}) = 2\pi\,(\mathrm{rand}(M,1) - 0.5)$;
  for $v = 1 : V$ **do**
    % Random walk, along Eq. (7)
    $\mathbf{M}^{(v)} = \mathrm{mod}\left(\mathbf{M}^{(v-1)} + \sigma\,\mathrm{randn}(M,1) + \pi, 2\pi\right) - \pi$;
    % Collect the corresponding sensation
    $\mathbf{S}^{(v)} = \Psi(\mathbf{M}^{(v)})$;
    if $\mathbf{S}^{(v)} \neq \mathbf{0}$ **then**
      $\mathbf{M}^{*(r)} = \mathbf{M}^{(v)}$;
      $\mathbf{S}^{*(r)} = \mathbf{S}^{(v)}$;
      $r = r + 1$;
    end if
  end for
end for
$R = r$; *% Total number of non-null sensations*

{Kernel set sampling}
$\boldsymbol{S}_1^* = \boldsymbol{S}^{*(\mathrm{randi}(R))}$; *% Pick randomly a sensation among R*;
for $i = 2 : N$ **do**
  Find first $r \in [1, R]$ **such that** $\mathrm{d}(\mathbf{S}^{*(r)}, \mathbf{S}_j^*) \geq \delta, \quad \forall j < i$;
  $\mathbf{S}_i^* = \boldsymbol{S}^{*(r)}$;
end for
% Append the satisfying motor configurations to the kernel set
for $i = 1 : N$ **do**
  for $r = 1 : R$ **do**
    if $\mathrm{d}(\mathbf{S}_i^*, \mathbf{S}^{*(r)}) \leq \delta/2$ **then**
      Append $\mathbf{M}^{*(r)}$ to M_i
    end if
  end for
end for

{Computation of the metric}
for $i = 1 : N$ **do**
  for $k = 1 : N$ **do**
    $D(i,k) = \rho(M_i, M_k)$;
  end for
end for

{Computation of a low-dimensional projection through CCA}
$C = \mathrm{CCA}(D)$;
return C

continuity proof. In practice, the interpolations performed on the internal map are consistent, thus suggesting that a richest mathematical structure –possibly involving topological considerations– might be introduced. These aspects are being investigated, together with the extension of the proposed formalism to the building of an internal representation of the peripersonal space outside the body, and the evaluation of the approach genericity to various kind of agent body structure. Importantly, the provided mathematical proofs show that if the agent is able to build motor kernel manifolds by touching its own body, then an internal representation of it can be build. This might imply kinematics constraints (like the need of motor redundancies) which must be carefully identified.

V. Conclusion

Is a system able to build an internal representation of its own body, without any other information than its motor commands and the subsequent sensory information? This work has proposed a first positive answer to this question. For that purpose, basic mathematical proofs have been written, showing that motor quotient sets are ideal candidates for the representation of the end effector position within the agent working space. Then, the paper has focused on the use of a tactile-like modality. Together with the adequate motor quotient set, we have shown in simulation how a naive agent can build its own internal body representation. Results show that the global topology of the body is preserved, exhibiting a small error between target points on the body and their corresponding representative in the internal representation, the comparison being possible only after a CCA and an isometric transformation.

References

[1] D. Knill and W. Richards, *Perception as Bayesian Inference*. Cambridge University Press, Oct. 1996.

[2] H. Poincaré, "Paris: La science et lhypotheèse," *Flammarion*, 1902.

[3] H. Poincaré, "Science and hypothesis," *London: Walter Scott Publishing*, 1905.

[4] H. Poincaré, "L'espace Et la Géométrie," *Revue de Métaphysique Et de Morale*, vol. 3, no. 6, pp. 631–646, 1895.

[5] J. K. O'Regan and A. Noë, "A sensorimotor account of vision and visual consciousness.," *The Behavioral and brain sciences*, vol. 24, pp. 939–73; discussion 973–1031, Oct. 2001.

[6] D. Philipona, J. K. O'Regan, and J.-P. Nadal, "Is there something out there?: Inferring space from sensorimotor dependencies," *Neural Comput.*, vol. 15, no. 9, pp. 2029–2049, 2003.

[7] D. Philipona, J. K. O'Regan, J.-P. Nadal, and O. J.-M. Coenen, "Perception of the structure of the physical world using unknown sensors and effectors," *Advances in Neural Information Processing Systems*, vol. 15, 2004.

[8] A. Laflaquière, S. Argentieri, O. Breysse, S. Genet, and B. Gas, "A non-linear approach to space dimension perception by a naive agent," in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*, pp. 3253–3259, 2012.

[9] V. Y. Roschin and A. Frolov, "A neural network model for the acquisition of a spatial body scheme through sensorimotor interaction," *Neural Computation*, vol. 23, pp. 1821–1834, 2011.

[10] A. Laflaquière, J. K. O'Regan, S. Argentieri, B. Gas, and A. Terekhov, "Learning agents spatial configuration from sensorimotor invariants," *Robotics and Autonomous Systems*, vol. 71, pp. 49–59, September 2015.

[11] M. Lopes and B. Damas, "A learning framework for generic sensory-motor maps," in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pp. 1533–1538, Oct 2007.

[12] J. Stober, R. Miikkulainen, and B. Kuipers, "Learning geometry from sensorimotor experience," in *Proceedings of the First International Conference on Development and Learning and Epigenetic Robotics*, (Frankfurt am Main, Germany), August 2011.

[13] A. Laflaquière, "Approche sensorimotrice de la perception de l'espace pour la robotique autonome," 2013.

[14] P. Demartines and J. Herault, "Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets.," *IEEE Transactions on Neural Networks*, vol. 8, pp. 148–54, Jan. 1997.

[15] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America A*, vol. 4, no. 4, pp. 629–642, 1987.