Acquisition of a space representation by a naive agent from sensorimotor invariance and proprioceptive compensation

Gurvan Le Clec 'H, Bruno Gas, J. Kevin O'Reagan

Research Article

INTERNATIONAL JOURNAL OF
ADVANCED ROBOTIC SYSTEMS

International Journal of Advanced Robotic Systems
November-December 2016: 1–15
© The Author(s) 2016
DOI: 10.1177/1729881416675134
arx.sagepub.com
SAGE

Acquisition of a space representation by a naive agent from sensorimotor invariance and proprioceptive compensation

Gurvan Le Clec'H[1], Bruno Gas[1], and J. Kevin O'Regan[2]

Abstract

In this article, we present a simple agent which learns an internal representation of space without a priori knowledge of its environment, body, or sensors. The learned environment is seen as an internal space representation. This representation is isomorphic to the group of transformations applied to the environment. The model solves certain theoretical and practical issues encountered in previous work in sensorimotor contingency theory. Considering the mathematical description of the internal representation, analysis of its properties and simulations, we prove that this internal representation is equivalent to knowledge of space.

Keywords

Sensorimotor learning, space representation, sensorimotor contingencies

Date received: 1 February 2016; accepted: 5 September 2016

Topic: Bioinspired Robotics
Topic Editor: Mohsen Shahinpoor

Introduction

Sensorimotor theory

Sensorimotor contingency theory argues that the acquisition of space knowledge in the brain is a result of the interaction between perception and body movement. "Passive perception" alone is not sufficient to create a representation of space, instead many authors propose "active percetion" in which action is a necessary component of perception.[1,2] The connections between sensory inputs and motor outputs are defined by a general set of rules whose properties depend on the characteristics of the surrounding space.[2–8] The agent is able to use its body to compensate for sensory changes. In response to sensory changes, which are a result of changes in the environment or body movements, the agent will move to counter the effect of the initial changes. Poincaré[4,5] described the compensation algorithm as the capability of the body to compensate for a transformation of the environment. Nicod[6,9] applied the concept of compensation to auditory signals and stated that a space representation can emerge when body movements are used in interaction with the auditory system. More recently, O'Regan and Noë[2] used psychological arguments obtained from experiments on humans and animals to clearly define the sensorimotor contingency approach and outline its expectations. Philipona et al.[7,8] performed physical simulations, modeling, and analyses of the tangent spaces of the manifold of sensorimotor interactions. These studies showed that it is possible to retrieve the dimension of the

[1] Institut des Systemes Intelligents et de Robotique, Paris, France
[2] Laboratoire Psychologie de la Perception - CNRS Université Paris Descartes, Paris, France

Corresponding author:
Gurvan Le Clec'H, ISIR-UMR 7222, Univ. Paris 06, 4 place Jussieu, 75005 Paris, France.
Email: gurvan.leclech@gmail.com

learned space for any sensory and motor dimensions. The algorithm of Philipona retrieved the dimension of the group of transformations in which their agent moved without any prior knowledge of it and without knowledge of the sensory and motor dimensions. Laflaquiere et al.[10,11] implemented this type of algorithm in order to retrieve the dimensionality of the environment in which a robot moved its arm.

In a psychological study, Aytekin et al.[12] demonstrated that humans use a sensorimotor approach to learn space from auditory stimulation. The group properties and the metric of the auditory space were shown to be captured by the human brain, thanks to sensorimotor interactions. Terekhov and O'Regan[13,14] showed that an agent using internal compensation movements could acquire external movements and the metric without knowledge of the environment.

The present article extends the field of sensorimotor research by demonstrating a naive agent that learns the group properties of space and provides new insights into theoretical results. The naive agent creates a usable representation of space and retrieves its fundamental properties such as its group operation. The term "naive" indicates that before learning the agent had no awareness of space or its properties. Using compensation, this agent is able to learn an internal representation of space. We extend the Laflaquière and Terekhov models, merging their algorithms. Our results remove all ambiguity in the mathematical representation of space and extend the usability of the representation. This representation is itself a group and is isomorphic to the group of compensable transformations. The agent measures sensory signals from the environment, which correspond to changes in its perception. The change in perception can be due to movements generated by the agent (internal movement, external movement, or both), a change in the environment (a transformation applied to the sources of signal), or to a combination of these. The agent will act to compensate for these changes in order to retrieve the same signal it was experiencing before the change.

In this article, the agent's proprioceptive capabilities alone are used to create a full representation of its environment without any a priori knowledge of it. This representation has the same properties as its embedding space (R^2). The learning algorithm is based on the notion of compensation defined by Poincaré (Poincaré, 1898, 1902) for the visual perception of space.

In order to prove that our agent has learned a representation of space, we use the following points:

- The agent captures invariant proprioceptive domains in a stationary environment, which provide an internal calibration for its own movements.
- The agent learns the transformations of the environment by compensable movements (movements of the foot, retina, or both), which link environment transformation to body transformation.
- The agent can distinguish between learned transformation and non-learned transformation (compensable and non-compensable).
- The agent captures the properties of learned transformations.
- The group properties for combinations of external movement are reproduced by the agent's internal representation. When the agent uses its internal representation to predict or reproduce a combination of movements, the combinatory effect is preserved.
- The internal representation is isomorphic to the group of transformations.
- A handicapped agent that cannot apply the algorithm (no sensory matching) cannot learn the compensable movements and therefore cannot learn the space representation.

Using all of the above, we show that the internal representation is equivalent to a representation of space.

In the next section, we describe the agent in detail. We describe its sensory system and body. We explain the algorithm applied during learning and the theoretical requirements for proof that the agent learned a full representation of space. Next, we present the computational logic and calculations that have been applied to the agent. We present the solution used to validate our theoretical claims and discuss the effects studied. We then present our theoretical results and simulations. We compare the learned compensable transformations to noise and non-compensable transformations. We also show the proof of the group properties learned by the agent and its internal representation of space. Taking the example of an agent with a particular form of handicap, we show how this is reflected in terms of the representation of space. Finally, we discuss the results of the model and present future work.

Theoretical presentation

The environment

The environment is composed of a number of light sources. Signal propagation obeys physical laws and is generated by a simple ray tracing algorithm for each source. The environment and its state is defined by ϵ. The different states of the environment are noted with the q subscript giving ϵ_q to describe the environment in state q. All the possible states belong to the set $\varepsilon : \epsilon_q \in \varepsilon$.

The environment can be subject to transformation T and the environment will then be in a different state.

$$T(\epsilon_q) = \epsilon_{q'}, T : \varepsilon \rightarrow \varepsilon \tag{1}$$

The transformation of the environment ϵ can be rigid transformation (translation and rotation), geometrical transformation (scaling), or any type of transformation (noise, intensity change, and random movement of the sources of light). When subject to a rigid transformation,

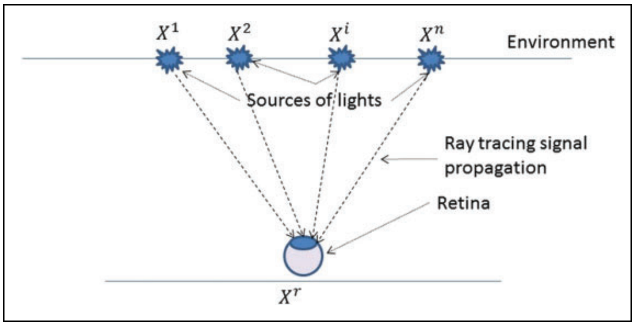

Figure 1. The signal measured by each visual cell depends on the distance between the sources *i* located at *Xi* and the retina located at X^r. The set of signals defines the visual signal vector, which depends on the number of cells in the retina. The signal is given by the function $\sigma(X^r, \epsilon) = \sum_{i=0}^{N} \sigma(X^i - X^r, \epsilon)$, where ϵ is the effect of the global position of the agent in the environment.

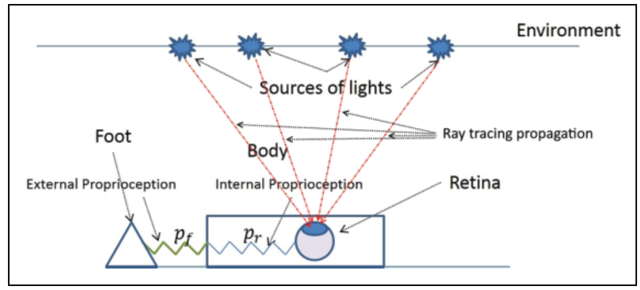

Figure 2. The agent is in a physical space and can move in any direction (left or right along X in this figure) by controlling a foot. The foot displacement is associated with the external proprioception $\boldsymbol{p}^f$. The agent has a retina which can move freely inside the body. The retina position depends on the internal proprioception $\boldsymbol{p}^r$. The environment is composed of a number of lights which can move together (not individually) in any direction.

the effect applies on all the sources of signal with a displacement of their locations in the physical space.

Sensations of the change in the environment

The agent has a retina, which is sensitive to the source of lights. The retina is a detector composed of visual cells, which are sensitive to the sources depending on the source position relative to the sensor. Let $s \in \mathcal{S}$ describe the sensorial activity of the agent. The dimension of $\mathcal{S}$ is the dimension of the retinal signal given by the number of visual cells. An example of this type of sensory detector is shown in Figure 1 . The signals are measured by the agent's detector. The detectors are sensitive to a signal emitted by the sources and the signal variation on a detector is directly linked to the source's localization in the space surrounding the agent. The signals captured by the detector depend on location with a function

$$s = \sigma(X^r, \epsilon) \tag{2}$$

where X^r is the relative position of the retina in the agent and ϵ is the environment's effect on the agent (which depends on the absolute position of the agent in the environment). The σ function is a binary relation between the environment and R^m where m is the dimension of the signal-measuring sensor $\sigma(X^r, \epsilon) = \sum_{i=0}^{N} \sigma(X^i - X^r, \epsilon)$ where the X_i is the lights' location in the environment relative to the retina. It is important to note that the signal function must be continuous and invertible.

Invertibility can be ensured by considering a sufficiently complex detector, that is, one composed of sufficient number of retina cells. We do not suppose we have a full domain where the function is invertible, however, in our simulation, we only got limitations with very few cells in the retina and this showed interesting defects in the space representation. Defects in the space representation obtained using an insufficient number of cells will be discussed later in the article. (The change in the environment can generate sensory variation in the agent.)

Agent description

The agent is a simple body composed of two moving parts: a foot which generates body displacements and a retina which can be moved inside the body (Figure 2). For simplicity, we present, in the Figure, a one-dimensional environment, however, our agent was validated and tested in a two-dimensional environment. The proprioceptive state of the agent $\boldsymbol{p}$ ($\boldsymbol{p} \in R^n$) is defined by combining the retina proprioceptive state $\boldsymbol{p}^r$ and the foot motor activity $\boldsymbol{p}^f$ as $\boldsymbol{p} = (\boldsymbol{p}^r, \boldsymbol{p}^f)^\top \in \mathcal{P} = \mathcal{P}_R \times \mathcal{P}_F$. Where $\mathcal{P}_R$ and $\mathcal{P}_F$ are the sets of all proprioceptive values. The foot proprioceptive state $\boldsymbol{p}^f \in \mathcal{P}_F$ can allow the agent to move both forward and backward.

In the following part of the article, we are using the subscripts q and q' in order to distinguish the state of the studied objects. $\boldsymbol{p}_q$ and $\boldsymbol{p}_{q'}$ are the proprioceptive states of the agent corresponding to two states q and q'. We are also using this notation to distinguish the states of the environment ϵ_q and $\epsilon_{q'}$. Using such a simple agent allows us to use precise mathematical tools to clarify the properties of space perception. Space and compensable transformations, as proposed by Poincaré,[3] are central to our study and are discussed in the two next sections.

The movements of the agent are of two kinds: internal when the agent moves its retina and external when the agent moves its foot. The effect of the movement of the retina is related to the proprioception of the retina $\boldsymbol{p}^r$ and the relative position of the retina in the agent is given by a function Π

$$X^r = \Pi(\boldsymbol{p}^r) \tag{3}$$

The proprioception of the foot $\boldsymbol{p}^f$ generates external movement. The body of the agent is fully displaced when the agent moves its foot. The movement of the foot is given by the function μ

$$\Delta \text{ Agent} = \mu(\boldsymbol{p}^f) \tag{4}$$

Applying the effect of $\mu(\boldsymbol{p}^f)$ to the agent is equivalent to a geometrical transformation to its body location (rotation and translation). When the foot proprioceptive value $\boldsymbol{p}^f$ is equal to 0, the agent is at rest.

Π and μ are bijective and thus are invertible functions. This bijectivity is a particular case that can be generalized.[15]

The absolute position of the retina in the evironment can be given by the position of the agent in the environment and the relative position of the retina in the agent X^r. While the agent doesn't know its absolute position, it can act by changing its retina position with the movements of its body. Only foot movements generate any displacement of the body. Foot movements are not bounded, while the retina movement is solely within the body of the agent. While this is a strong hypothesis, it is necessary for purely mathematical reasons. However, the limits of this hypothesis are not tested as we typically consider small movements centered on the agent's body and only small foot movements are needed to compensate.

Compensable transformation

In 1895, H. Poincaré wrote his work on space and geometry.[3] He defined geometry in relation to a totally naive brain, which has access to its sensorimotor flow only. The geometry can be inferred by considering certain types of sensory changes. Of all the possible sensory variations, some occur without motor commands and, therefore, must be related to external changes. Some changes that are related to external rigid displacements can be compensated by the agent's motor commands. In this case, the sensory variations due to the external changes and the motor commands are opposite, so that the initial and final sensory states are identical. This is what is meant by compensable. Because the function linking the positions of light sources and the agent sensor signal is invertible, for a stationary agent, we can state that for every change of the environment, the sensory perception of the agent also changes. When considering different positions of the agent in the environment with identical values of the proprioceptive state vector $\boldsymbol{p}$ (identical retina position and foot activity), the signals measured by the sensor are necessarily different. In the general case, the agent is unable to infer its sensorial variations from variations in its proprioceptive state. However, compensable transformations can be thought of as a type of sensorial variation due to external transformations that are compensated by a specific movement of the agent. These compensable transformations can be detected by the agent. According to Poincaré's proposition, the set of compensable transformations is a group and is equivalent to the group characterizing the external geometric space. By experiencing the compensable transformations, the agent can capture the most important properties of the (Euclidean) space in which it moves.

Capturing the set of compensable transformations

In this section, we introduce the formalism for compensable transformations by defining Φ as a set of binary relations φ_T that will allow us to build a representation of the group of compensable transformations T. We consider two cases: auto-compensable transformations in a stationary environment and compensable transformations more generally. Making this distinction allows the set of φ_T to map specific transformations T unambiguously. Initially, the φ_T functions were proposed by Terekhov et al.[13,14] They are built from a catalog of all compensable transformations the agent detects. By matching identical sensory inputs before a change and after compensation, φ_T functions map proprioception observed before a change to proprioception observed after the compensation.

For a given displacement T of the environment, the agent will recover the same function φ_T for both different initial positions and different environments. As shown by Terekhov and O'Regan,[14] the functions φ_T provide the agent with the notion of space.

(In this article, the φ_T are not functions but binary relation as they are not unique mappings from $\mathcal{P}$ to $\mathcal{P}$. However they are real functions in the article of Terekhov.)

Compensable transformations

Let $(s, \boldsymbol{p}_q) \in \mathcal{S} \times \mathcal{P}$ describe the state of the agent in the environment. After a given rigid displacement of the environment (lights moving in our case), the agent's new state is $(s', \boldsymbol{p}_q)$. Such a transformation is compensable if there exists a new proprioceptive state $\boldsymbol{p}_{q'}$ corresponding to a displacement of the agent in the environment and the resulting state is $(s'', \boldsymbol{p}_{q'})$ with $s'' = s$. In other words, the agent has moved in order to recover the initial perceptual state. Given a compensable transformation of the environment T, the function φ_T is the mapping between $\mathcal{P}$ and itself such that for all proprioceptive states $\boldsymbol{p}_q \in \mathcal{P}$, one has $\boldsymbol{p}_{q'} = \varphi_T(\boldsymbol{p}_q)$, where $\boldsymbol{p}_{q'}$ is the new proprioceptive state which compensates for T

$$\varphi_T : \mathcal{P} \to \mathcal{P}, \boldsymbol{p}_q \to \boldsymbol{p}_{q'} = \varphi_T(\boldsymbol{p}_q) \tag{5}$$

This definition does not depend on the perceptual state s. This is a consequence of the definition of compensable transformations. Cases where $s'' = s$ can be found as long as the agent can evaluate its sensory flow. Difficulties arise when discussing how an agent can detect such coincidences from the sensorimotor flow alone, without a priori knowledge. How can it infer for a set of $(\boldsymbol{p}_q, \boldsymbol{p}_{q'})$ that they correspond to the same φ_T? That is, how does it determine that they compensate for the same unknown external compensable transformation T? Terekhov and O'Regan proposed a solution[14] whereby they collected all the agent compensations corresponding to the same external transformation T and repeated for all other values of T. However, as the agent itself does not have access to T, this solution is somewhat artificial. In this article we propose a new solution to this problem by introducing the notion of auto-compensable transformation. In this case, the environment is stationary. Since this is not known by the agent, we must also introduce the identity transformation of the agent's body.

(As mentioned earlier, while the φ_{T} are functions in case of the Terekhov's article, in our case there are not. However, when we keep the foot proprioceptive state constant before and after the compensation (respectively the retina proprioceptive value before and after the transformation of the environment), the compensation φ_{T} becomes a function as there is a unique retina state compensating the applied transformation (respectively for the foot). When the environment is stationary, φ_0 is a function (see next section).

In the article, we will then use the term of function even if it can be seen as an inaccurate term.

Auto-compensable transformation

An auto-compensable transformation is the displacement of a part of the agent's body (e.g. retina), which compensates for sensorial variations induced by the movement of another part of the agent's body (e.g. foot displacement). The distinction between auto-compensation and compensation more generally is that the transformation T is fully determined by the agent in the case of auto-compensation. The agent can then build a set of functions φ_T, which can be used as internal references to represent external compensable transformations. In order to avoid any confusion between the agent's transformations T and external transformations, the environment must remain stationary when determining the function φ_T. However, the agent cannot know whether the environment is stationary or not. This problem has been discussed by Roschin and Frolov,[16] Laflaquière,[17] and more recently by Marcel et al.[15] without finding a fully satisfactory solution. We will address this by introducing the identity transformation. The identity transformation can be considered a simple saccadic movement. After a displacement of the agent's foot, the sensorial state of the retina changes. The agent then recovers the initial proprioceptive state of the foot by performing the exact inverse displacement of the foot. The initial sensorial state of the retina will be recovered. If it is not, then the agent can state that the environment changed during the identity transformation. Therefore, by retaining only those transformations which do not involve movements of the external environment, the agent can build unambiguous φ_T functions.

Φ the set of sensorimotor functions

As previously defined, Φ is the set of functions φ_T, which maps the proprioceptive states onto themselves. We will show that our agent can capture the set of compensable transformations T and that it learns the group properties of this set. Using equation (5), we show (in Appendix 1) that if the agent's state modification from $\boldsymbol{p}_q$ to $\boldsymbol{p}_{q'}$ compensates for the environment transformation T, we have

$$\boldsymbol{p}_{q'} = \varphi_T(\boldsymbol{p}_q) = \begin{pmatrix} \Pi \circ \mu(\boldsymbol{p}^f_q) \circ T \circ \mu(\boldsymbol{p}^f_{q'})^{-1} \circ \Pi^{-1}(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q'} \end{pmatrix} \tag{6}$$

$\Pi(p^r_q) = X^r_q$ is the function that describes the sensor position relative to the agent's body in the physical space given the agent's proprioceptive state $\boldsymbol{p}^r_q$. $\mu(\boldsymbol{p}_q)$ is the function that describes the movement of the agent's body given the foot proprioceptive state $\boldsymbol{p}^f_q$. We distinguish between the reverse function of μ, $\mu^{-1}(X_q) = \boldsymbol{p}^f_q$, and the opposite function of μ, $\mu(\boldsymbol{p}^f_q)^{-1}$, which is the opposite displacement of $\mu(\boldsymbol{p}^f_q)$. The function Π^{-1} is the inverse function of Π.

In this article, we present the φ_T functions and their properties and demonstrate that there is a calculable solution. It is important to note that the agent has no knowledge of Π and μ. Only the mapping of $\boldsymbol{p}_q$ to $\boldsymbol{p}_{q'}$ is known. We can describe this mapping by a mathematical function. In order to do so, we first consider transformations that are auto-compensable in a stationary environment before moving to the general case of a nonstationary environment.

Auto-compensable transformation. When considering the auto-compensable transformation, the environment is stationary. We calculate φ_0, where 0 indicates no movement of the environment. When the agent moves its foot, it compensates for the movement with its retina. The absolute position of the retina does not change. It is the same before the foot movement and after the compensating movement of the retina. After the foot movement, the environment perceived by the agent is changed in accordance with the physical laws of signal propagation and perception. Compensation is the act of retrieving the initial signal by a movement different to that which generated the difference. For every foot movement and initial retina position, the final retina position corresponding to the compensated movement is unique in a stationary environment (bijectivity of Π and μ). The signal measured after the compensating retina movement is compared to that before the foot movement. The algorithm applied to the agent is as follows

1. The agent is in an environment at a given location X_q, with the retina at a given position $X^r_q = \Pi(\boldsymbol{p}^r_q)$. The agent can move its foot and generate a body displacement $\mu(\boldsymbol{p}^f_q)$, where $\boldsymbol{p}^f_q$ is the foot state.

2. The agent moves to a new location $X_{q'}$ in response to a foot movement induced by the foot motor activity $\boldsymbol{p}^f_{q'}$. The transformation generating the displacement is given by $X_q \xrightarrow{\mu(\boldsymbol{p}^f_{q'})} X_{q'}$.
3. The agent compensates for the foot movement by moving its retina in such a way that the retina visual perception is the same before and after the initial movement. The practical implementation is given in Appendix 1. What the agent sees after compensation is exactly what it saw before the foot movement. The final relative retina position is measured only by its internal state $\boldsymbol{p}^r_{q'}$.
4. The agent creates an internal representation of auto-compensation by mapping the initial internal state $\boldsymbol{p}_q$ to the final state $\boldsymbol{p}_{q'}$, $\varphi_0(\boldsymbol{p}_q) = \boldsymbol{p}_{q'}$.
5. We show in the Appendix 1 that

$$\varphi_0(\boldsymbol{p}_q) = \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_q) \circ \mu(\boldsymbol{p}^f_{q'})^{-1} \circ \Pi(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q'} \end{pmatrix} \tag{7}$$

Compensable transformation. In this case, the environment is not stationary. When the environment is moved, the agent compensates by movement of the foot, retina, or both. The signal on the sensor is the same when measured before the movement of the environment and after the compensating movement of the agent. The algorithm is as follows

1. The agent is in an environment at a given location X_q, with the retina at a given position $X^r_q = \Pi(\boldsymbol{p}^r_q)$. The agent can move its foot and generate a body displacement based on the function $\mu(\boldsymbol{p}^f_q)$ where $\boldsymbol{p}^f_q$ is the foot state.
2. The environment is moved to $\epsilon_{q'}$ with $\epsilon_q \xrightarrow{T} \epsilon_{q'}$.
3. The agent compensates the external movement by either:

   a. A foot movement only. The body is moved in such a way that its retina visual perception is the same before and after the initial movement. The retina proprioceptive value remains unchanged. The foot proprioceptive value is $\boldsymbol{p}^f_{q'}$.
   b. A retina movement only. The body is not moved but the retina is moved in such a way that its retina visual perception is the same before and after the initial movement. The retina proprioceptive value is $\boldsymbol{p}^r_{q'}$. The final foot proprioceptive value is 0.
   c. Both foot and retina movement. Both the body and retina are moved in such a way that the retina visual perception is the same before and after the initial movement. The foot proprioceptive value is $\boldsymbol{p}^f_{q'}$. The retina proprioceptive value is $\boldsymbol{p}^r_{q'}$.

4. The agent creates an internal representation of the transformation T by mapping the initial internal state of the agent $\boldsymbol{p}_q = \begin{pmatrix} \boldsymbol{p}^r_q \\ \boldsymbol{p}^f_q \end{pmatrix}$ to the final state $\boldsymbol{p}_{q'} = \varphi_T(\boldsymbol{p}_q) = \begin{pmatrix} \boldsymbol{p}^r_{q'} \\ \boldsymbol{p}^f_{q'} \end{pmatrix}$.
5. We show in the Appendix 1 that

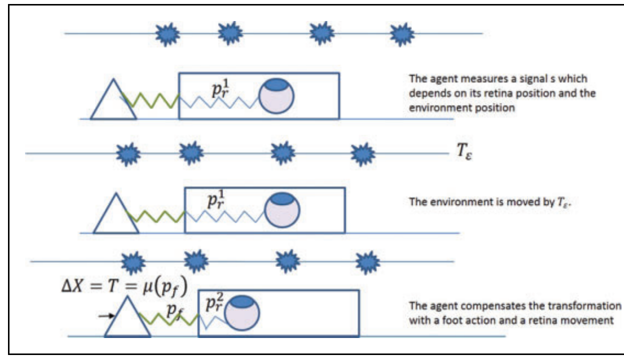

Figure 3. The matching coincidence algorithm is based on the comparison of the signal measured by each of the retina's cells before an external movement is applied to the agent and after the compensating retina internal movement.

$$\varphi_T(\boldsymbol{p}_q) = \varphi_T\begin{pmatrix} \boldsymbol{p}^r_q \\ \boldsymbol{p}^f_q \end{pmatrix} = \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_q) \circ T \circ \mu(\boldsymbol{p}^f_{q'})^{-1} \circ \Pi(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q'} \end{pmatrix} \tag{8}$$

The set of $\varphi_T(\boldsymbol{p})$ defines a manifold in internal state space and can be used to show that the agent captures the geometrical space as defined by Poincaré.[3] The acquired knowledge is a space representation given that its properties mathematically correspond to those of space (including its group properties and the isomorphism with the compensable transformation set). This is discussed in the section "Φ is a representation of the geometrical space." Figure 3 illustrates the steps of the coincidence algorithm outlined above.

Φ *is a representation of the geometrical space*

As outlined previously, we will show that the set Φ of functions φ_T that corresponds to the mapping of proprioceptive states to compensable transformations has all the properties of the external space in which the agent is situated. The agent is capable of capturing the relevant properties even though it is not aware that such a set of transformations is a group. Thus, the set Φ of φ_T is a group and is also isomorphic to the set of T.

Combinatory property. The combinatory operation $\circ$ on the set of φ_T functions is defined such that for any T_q, $T_{q'}$ there exists $T_{q''}$ such that $\varphi_{T_{q'}} \circ \varphi_{T_q} = \varphi_{T_{q''}}$. In other words, the combination of two compensable transformations is itself a compensable transformation.

Demonstration: From the construction of the function φ_T, for any compensable transformation T, there is an associated function φ_T. Using the definition of a compensable transformation, there exists an action, which compensates for an external transformation. The compensation can be either an internal or external movement. As the displacement caused by the movement of the foot can be any length $(\mu(\boldsymbol{p}^f))$, all combinations of compensable transformations can be compensated by foot movement. Thus, any combination of compensable transformations is compensable.

It is also the case that $\forall T_q, T_{q'}$ and $T_{q''}$ where $T_{q'} \circ T_q = T_{q''}$ then $\varphi_{T_{q'}} \circ \varphi_{T_q} = \varphi_{T_{q''}} = \varphi_{T_{q'} \circ T_q}$.

Demonstration

We begin with

$$\varphi_{T_q}(\boldsymbol{p}_q) = \boldsymbol{p}_{q'} \tag{9}$$

Applying equation (9) twice gives

$$\varphi_{T_{q'}} \circ \varphi_{T_q}(\boldsymbol{p}_q) = \varphi_{T_{q'}}(\boldsymbol{p}_{q'}) = \boldsymbol{p}_{q''}$$

Combining these functions with equation (8), we obtain

$$\begin{aligned}
\boldsymbol{p}_{q''} = \varphi_{T_{q'}}(\boldsymbol{p}_{q'}) &= \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_{q''}) \circ T_{q'} \circ \mu(\boldsymbol{p}^f_{q'})^{-1} \circ \Pi(\boldsymbol{p}^r_{q'}) \\ \boldsymbol{p}^f_{q''} \end{pmatrix} \\
&= \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_{q''}) \circ T_{q'} \circ \mu(\boldsymbol{p}^f_{q'})^{-1} \circ \Pi \circ \Pi^{-1} \circ \mu(\boldsymbol{p}^f_{q'}) \circ T_q \circ \mu(\boldsymbol{p}^f_q)^{-1} \circ \Pi(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q''} \end{pmatrix} \\
&= \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_{q''}) \circ T_{q'} \circ T_q \circ \mu(\boldsymbol{p}^f_q)^{-1} \circ \Pi(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q''} \end{pmatrix} \\
&= \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}^f_{q''}) \circ T_{q''} \circ \mu(\boldsymbol{p}^f_q)^{-1} \circ \Pi(\boldsymbol{p}^r_q) \\ \boldsymbol{p}^f_{q''} \end{pmatrix} \\
&= \varphi_{T_{q''}}(\boldsymbol{p}_q) = \varphi_{T_{q'} \circ T_q}(\boldsymbol{p}_q) = \varphi_{T_{q'}} \circ \varphi_{T_q}(\boldsymbol{p}_q)
\end{aligned}$$

The combination of φ_T functions does not depend on the intermediary steps selected by the agent.

Group property. The axioms necessary to validate a group are closure, associativity, the existence of an identity, and the existence of an inverse for every element of the group.

- Closure

$$\forall \boldsymbol{p} \in \mathcal{P}, \forall \varphi_{T_q}, \varphi_{T_{q'}} \text{ then } \varphi_{T_{q'}} \circ \varphi_{T_q}(p) \in \mathcal{P} \tag{10}$$

where $\mathcal{P}$ is the set of possible proprioceptive measures of the retina and foot.

This first property is obvious. It comes from the learning algorithm where the agent changes its proprioceptive state with a retina or foot movement to compensate for an external agent movement. The result of applying φ_{T_q} to any $\boldsymbol{p} = \begin{pmatrix} \boldsymbol{p}^r \\ \boldsymbol{p}^f \end{pmatrix}$ belongs to $\mathcal{P}$. We must demonstrate that $\forall \boldsymbol{p} \in \mathcal{P}, \forall T_q$ and $T_{q'}$ then $\varphi_{T_{q'}} \circ \varphi_{T_{q'}}(\boldsymbol{p}) \in \mathcal{P}$.

Demonstration: From the algorithm, we have $\forall \boldsymbol{p}_q \in \mathcal{P}, \forall T_q$ then $\varphi_{T_q}(\boldsymbol{p}_q) \in \mathcal{P}$. As $\varphi_{T_q}(\boldsymbol{p}_q) = \boldsymbol{p}_{q'}$, we apply the same logic: $\forall \boldsymbol{p}_{q'} \in \mathcal{P}, \forall T_{q'}$ then $\varphi_{T_{q'}}(\boldsymbol{p}_{q'}) = \varphi_{T_{q'}} \circ \varphi_{T_q}(\boldsymbol{p}) \in \mathcal{P}$

- Associativity

$$\forall \varphi_{T_q}, \varphi_{T_{q'}}, \varphi_{T_{q''}} \text{ then } \varphi_{T_q} \circ (\varphi_{T_{q'}} \circ \varphi_{T_{q''}}) = (\varphi_{T_q} \circ \varphi_{T_{q'}}) \circ \varphi_{T_{q''}} \tag{11}$$

Demonstration

$$\begin{aligned}
\varphi_{T_q} \circ (\varphi_{T_{q'}} \circ \varphi_{T_{q''}}) &= \varphi_{T_q} \circ \varphi_{T_{q'} \circ T_{q''}} = \varphi_{T_q \circ T_{q'} \circ T_{q''}} \\
&= \varphi_{T_q \circ T_{q'}} \circ \varphi_{T_{q''}} = (\varphi_{T_q} \circ \varphi_{T_{q'}}) \circ \varphi_{T_{q''}}
\end{aligned}$$

- Identity

$$\forall \varphi_T, \varphi_0 \text{ then } \varphi_T \circ \varphi_0 = \varphi_0 \circ \varphi_T = \varphi_T \tag{12}$$

where φ_0 is the identity

- Inverse

$$\forall \varphi_T \text{ then } \varphi_T \circ \varphi_T^{-1} = \varphi_T^{-1} \circ \varphi_T = \varphi_0 \tag{13}$$

Demonstration for both identity and inverse:

In the group of transformations, $\forall T$ we have $T \circ T^{-1} = Id$ which implies $\forall T,\ T^{-1}$ then $\varphi_T \circ \varphi_{T^{-1}} = \varphi_{T \circ T^{-1}} = \varphi_0$.

This implies that $\varphi_T^{-1} = \varphi_{T^{-1}}$

which proves the existence of an identity element (φ_0) and the existence of a inverse function $\varphi_T^{-1} = \varphi_T^{-1}$ for any φ_T.

These points demonstrate that the set Φ of the functions φ_T is a group.

There is an isomorphism between the set of compensable transformations T and the set of functions φ_T. In order to prove there is an isomorphism, we have to prove that the two sets have the same dimension and that the φ_T is linear and injective.

For the dimension of both sets, we are using previous work from Philipona et al.[7,8] and Laflaquiere et al.,[11] which show that the internal representation of a compensatory agent has the same dimension as the geometrical space of transformation.

For linearity, we have shown that $\forall T_q, T_{q'}$ then $\varphi_{T_q} \circ \varphi_{T_q} = \varphi_{T_q \circ T_{q'}}$. In the limit of compensability, we can easily extend it to $\forall T_q,\ T_{q'},\ \alpha,\ \beta$ then $\varphi_{\alpha T_q} \circ \varphi_{\beta T_{q'}} = \varphi_{\alpha T_q \circ \beta T_{q'}}$ which is the expected linearity.

To show that the φ_T are injective, we must prove that $\forall T_q, T_{q'}$ if $\varphi_{T_q} = \varphi_{T_{q'}}$ then $T_q = T_{q'}$. Since the set of φ_T is a group, for every item in the set, there is a reverse function. Thus, $\varphi_{T_q} = \varphi_{T_{q'}}$ then $\varphi_{T_q} \circ \varphi_{T_{q'}}^{-1} = Id = \varphi_{T_q \circ T_q'^{-1}}$. Previously, we determined that if $\varphi_T = Id$, then $T = Id$. Therefore $T_q \circ T_{q'}^{-1} = Id$ and $T_q = T_{q'}$.

Theses points demonstrate that the set of functions φ_T is isomorphic to the set of compensable transformations and is thus a space representation.

Computation

Learning phase: Algorithm for computation of φ_T

In this section, we present the algorithm used to calculate φ_T where T is the external transformations applied to the agent. The set of transformations is a set of translations in two dimensions.

For the simulation, we considered a limited set of transformations of the environment T and a limited set of possible compensations (both foot and retina; see Appendix 1 for details). Applying this algorithm allowed us to calculate the complete set of functions $\varphi_T(\boldsymbol{p}_q) = \boldsymbol{p}_{q'}$.

Using φ_T functions to retrieve movement

By referencing the memorized tuples of three elements $(T, \boldsymbol{p}_q, \boldsymbol{p}_{q'})$, the combination of any two can be used to retrieve the third. There are multiple possible solutions for $\boldsymbol{p}_{q'} = \begin{pmatrix} \boldsymbol{p}_q^r \\ \boldsymbol{p}_q^f \end{pmatrix}$ given $\boldsymbol{p}_q$ and T.

φ_T functions are only sensitive to compensable transformations

In order to test the algorithm, we applied transformations other than those the agent learned. We first applied a continuous transformation, where the length of translation is not a multiple of the basic step size used in the computation. We then applied a scaling transformation, where the source objects are deformed by homothetic deformation. We also added random noise to the source light signals, with an amplitude ranging from 10% to 500% of the initial signal. Starting with a random initial retina proprioceptive state $\boldsymbol{p}_q^r$ and a random foot movement $\boldsymbol{p}_q^f$, we applied a random transformation T and then compensated for it, giving the final retina proprioceptive state $\boldsymbol{p}_{q'}^r$ and foot action proprioceptive state $\boldsymbol{p}_{q'}^f$. We then retrieved φ_{T_q} from the best proprioceptive set $(\boldsymbol{p}_q, \boldsymbol{p}_{q'})$ and extracted the associated T_q. We compared the applied T_a and extracted T_e (note that the T_q found from the algorithm is noted as the extracted T_e. When applying unknown (or non-compensable) transformations, coincidence matching could not be done. That is, it was not possible to match the states before and after movement. Exact matching was not possible and the error on the comparison of sensory measures increased with the difference between the initial image and the final image. The error between the applied T_a and the extracted T_e also became more significant as we measured $\| T_a, T_e \|$. This test was repeated for 1000 transformations, starting with random retina positions and we calculated the error between the applied movement and the compensated transformation for the best tuple $(\boldsymbol{p}_q, \boldsymbol{p}_{q'})$.

Internal space as a group of transformations

A calculation was performed for the full set of points in the combined φ_T in order to verify $\varphi_{T_1} \circ \varphi_{T_2} = \varphi_{T_3}$ then $T_3 = T_1 \circ T_2$. For any random transformations T_1 and T_2, we calculated $T_3 = T_1 \circ T_2$. We then compared φ_{T_3} acquired directly from the set of φ_T and the calculated $\varphi_{T_1 \circ T_2} = \varphi_{T_1} \circ \varphi_{T_2}$. As shown in the paragraph on the validation of the group axioms, this property proves the associativity, identity, and inverse axioms.

Testing parameters and mathematical functions. Our initial selection of functions and parameters did not affect our results. In order to demonstrate this, we repeated the simulations with multiple sets of proprioceptive functions $\left(\Pi(\boldsymbol{p}^r) = X^r\right)$ and varied the ratio of the applied movements to the retina or foot movements.

Figure 4 illustrates the simple and complex proprioceptive models used in this article.

In all simulations, we used a grid for the environment and a grid for the body displacement. The relevant values for the simulation are as follows:

- The number of steps the agent moves on a grid within the environment, # Steps.
- The ratio of environment displacement step size to body movement step size, $\frac{\Delta e}{\Delta b}$.
- The proprioceptive function measure $\Pi(p)$, which is one of:
  - Affine function: $\Pi(p) = X = a \times p + b$
  - Logarithmic function: $\Pi(p) = X = \sqrt{\ln(1 - p)}$

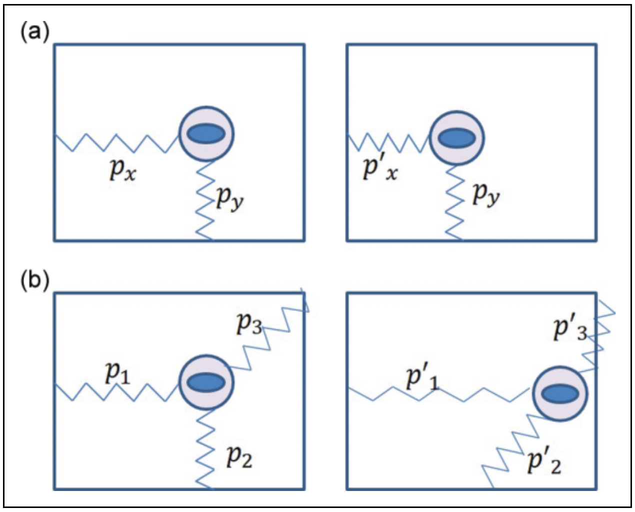

Figure 4. Comparison between simple and complex agents. (a) Simple agent. When the retina moves in the direction X, only the proprioceptive p_x is transformed into p'_x. p_y is not affected by the movement and (b) complex agent. When the retina moves in any direction, all proprioceptive detectors are affected (p_1, p_2, and p_3) then (p'_1, p'_2, and p'_3).

- Proprioceptive coupling
  - Uncoupled proprioceptive function (simple model). Movement in any direction is associated with a single proprioceptive measure. $\Pi(p_x^r) = x$ and $\Pi(p_y^r) = y$. This generates two independent φ_T functions, φ_x and φ_y, where $T(x,y)$ is the applied transformation.
  - Coupled proprioceptive function (complex model). Movement in any direction affects all proprioceptive measures. $\Pi(p) = \Pi(p_1^r, p_2^r,, p_i^r,, p_n^r) = (x, y)$. This generates a single multidimensional φ_T, where T is the applied transformation.

The retina of the agent is composed of randomly located cells in the retina. In the next section, we present the results of varying sensory parameters such as the number of cells and retina size in the simulations.

Results

Compensable movement versus non-compensable movement

The estimated displacement when applying either compensable movement or non-compensable movement is listed in Table 2. Since we know the applied transformation (even though the agent does not), we know the expected position after the transformation. We compare this to the agent's estimated position after transformation and calculate the difference between them. When the difference is zero, the expected and estimated positions are identical. For this test, we selected 1000 random movements (for each type of transformation) and looked for the transformations T, which could be retrieved from the known φ_T. The transformations were as follows:

- Compensable transformations: Translations that are multiples of the agent step length and land on the agent grid positions.
- Continuous transformations: Translations of continuous length which are not multiples of the agent step length.
- Scale transformations: Include both a regular translation and a homothetic transformation applied to the source object.
- Noise transformations: Random noise is applied to the signal of the source object. The noise factor 100%. The source signal varied from 0 to 2.

Results are presented in Tables 1, 2, and 3.

Results in Tables 1 and 2 show that only compensable movements were learned by the agent. We measured the distance $\| T_a, T_e \|$ between the applied transformation T_a and the extracted transformation T_e. The estimated movement is retrieved from the φ_T, which maps the priopriceptive state before the transformation to the proprioceptive state after compensation. We retrieved the learned transformation using the tuple p_b and p_a where p_b is the proprioceptive measure of the agent (retina location before the movement and agent in rest position) and p_a is the proprioceptive measure of the retina and the foot movement after compensation. The applied transformation T_a is directly linked to (p_b, p_a). The Tables 1, 2, and 3 compare the distances $\| T_a, T_e \|$ (T_a applied transformation and T_e estimated transformation).

For compensable transformations, the agent always retrieved the correct transformation. That is, the difference between T_a and T_e was always zero. This capacity to retrieve the applied movement T for any tuple $\boldsymbol{p}_a = \varphi_T(\boldsymbol{p}_b)$ was expected from the model. Non-compensated transformations were not properly retrieved. The error on the estimated movements was typically on the order of the agent size but was on the order of the size of the environment for scale and noise transformations. The continuous transformations gave better results than the scale and noise transformations. Contrary to the results for scale and noise transformations, the average distance for continuous transformations was dependent on the size of the agent step length. Continuous movements will always yield a better match than non-regular movements (scale or noise transformations). This can be interpreted by considering the grid of possible agent positions. As the φ_T were sampled on the grid of possible agent positions, any estimation of movement was always on a node of this grid. Thus, the estimations for continuous movements (indeed any tested transformations) always fell on the grid. In this case, the continuous movements were translations and the error depended on the agent step size and whether the continuous movement landed between two grid positions. The nearest

Table 1. Simple model (uncoupled proprioceptive measure).[a]

Parameters			Error on $\| T_a, T_e \|$			
# Steps	$\frac{\Delta e}{\Delta b}$	Π	Compensable T	Continuous T	Scaling T	Noise 100%
10 × 10	1	Affine	0	0.447	1.684	1.659
20 × 20	1	Affine	0	0.293	1.588	1.564
10 × 10	1	Logarithmic	0	0.445	1.716	1.684
20 × 20	1	Logarithmic	0	0.258	1.627	1.596
10 × 10	2	Affine	0	0.478	1.666	1.614
20 × 20	2	Affine	0	0.306	1.632	1.594
10 × 10	2	Logarithmic	0	0.478	1.720	1.676
20 × 20	2	Exponential	0	0.289	1.635	1.606

[a]Proprioceptive fields are decoupled with $p = (p_x, p_y)$. Movements in X direction affect only p_x. Movements in Y direction affect only p_y. φ_x, and φ_y functions were calculated for different step lengths and different retina to agent step length ratios. We measured the distance $\| T_a, T_e \|$ between the applied transformation T_a and the extracted transformation T_e

Table 2. Complex model (coupled proprioceptive measure).[a]

Parameters			Error on $\| T_a, T_e \|$			
# Steps	$\frac{\Delta e}{\Delta b}$	Π	Compensable T	Continuous T	Scaling T	Noise 100%
10 × 10	1	Affine	0	0.482	1.702	1.658
20 × 20	1	Affine	0	0.302	1.639	1.600
10 × 10	1	Logarithmic	0	0.478	1.662	1.623
20 × 20	1	Logarithmic	0	0.287	1.684	1.650
10 × 10	2	Affine	0	0.507	1.680	1.658
20 × 20	2	Affine	0	0.285	1.615	1.584
10 × 10	2	Logarithmic	0	0.466	1.689	1.650
20 × 20	2	Logarithmic	0	0.294	1.601	1.579

[a]Results for linked proprioceptive sensors. Moving the retina in any direction affects all proprioceptive measures. We measured the distance $\| T_a, T_e \|$ between the applied transformation T_a and the extracted transformation T_e.

Table 3. Varying sensory parameters and number of signal sources for compensable and non-compensable movement.[a]

Parameters		Error on $\| T_a, T_e \|$			
# Visual cells	# Sources	Compensable T	Continuous T	Scaling T	Noise 100%
1	1	0.657	0.946	1.687	1.654
2	1	0.04	0.753	1.669	1.649
3	1	0	0.473	1.645	1.645
1	10	0	0.473	1.714	1.707
2	10	0	0.476	1.679	1.680
3	10	0	0.479	1.712	1.712

[a]Since the system was able to properly compensate with only three retina cells for one source and with any number of retina cells for 10 sources, further results are not included. The # Steps are 10 × 10 and the retina to agent step length ratio is 1. The distance $\| T_a, T_e \|$ between the applied transformation T_a and the extracted transformation T_e has been measured.

grid position was used as the estimated position. Although the agent did not learn continuous movements, continuous movements are similar to compensated movements within the precision of the agent step length.

For scale and noise transformations, the deformations of the source signal could not be matched by the agent space representation. As the deformation of the image is not a rigid transformation, the perceived sensory signal before and after the transformation cannot correspond exactly. This gives rise to a significant error on the difference in sensory signals. The error was on the order of the agent size and did not depend on any of the variable parameters (agent step length, retina step length, and proprioceptive function). For an agent with one or two retina cells, the compensation algorithm did not work properly and the agent could not retrieve the movement. The results show that the compensable transformations were fully learned while non-compensable transformations were not mapped very well.

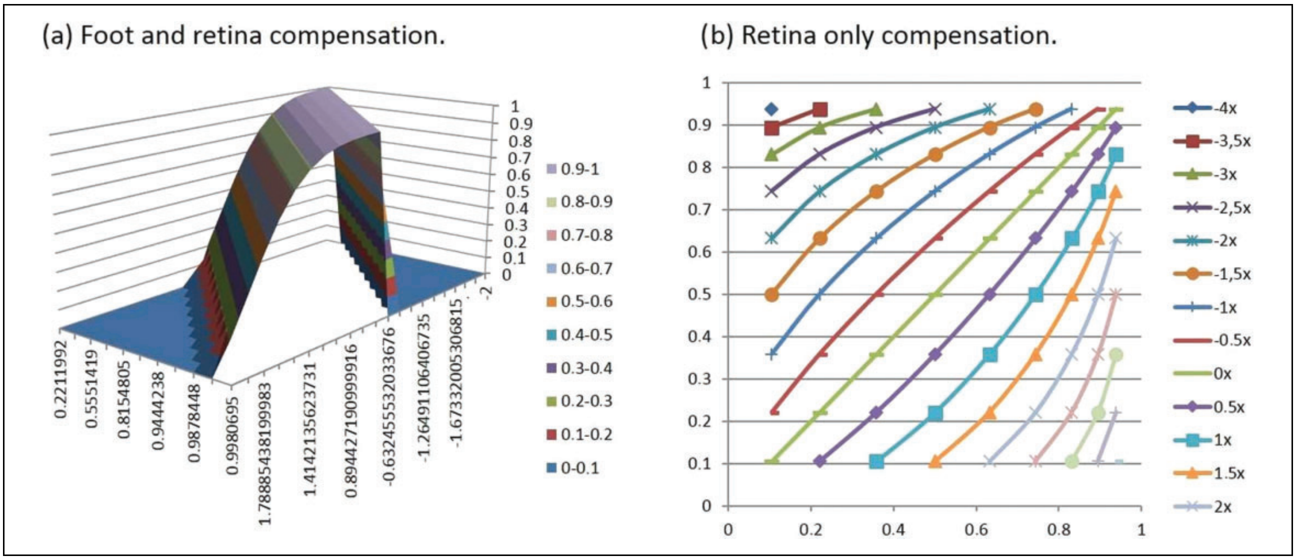

Figure 5. Set of curves φ_x where x is the amplitude of the translation of the agent. The step size of the translation is 0.5 with an agent length of 4. For this simple agent (x and y movements are decoupled), there are eight proprioceptive steps. We noted the external transformation with the x and y following the length of the displacement. (a) φ_x is expressed in terms of foot motor proprioceptive values (from -2 to 2) and the retina proprioceptive measure (from 0 to 1). The ordinate is the retina value after compensation. This surface corresponds to all possible couples p_f, p_r for a single movement ($-1{,}2$ x in our case) and (b) projection of φ_x on the retina proprioceptive coordinates without foot movement. Each curve corresponds to a unique external displacement. We see that the curves are symmetric $\varphi_x \circ \varphi_{-x} = \varphi_0$ then $\varphi_x = \varphi_{-x}^{-1}$. φ_0 is the invariant curve (no external transformation).

The agent learned the compensable transformations and was able to distinguish between compensable and non-compensable transformations.

Group properties of the function φ_T

The results of combining the φ_T functions in order to validate the group properties are presented in Table 4. In order to observe any effect on the quality of the space representation, we varied the numbers of cells and sources used in the simulations.

Discussion

Sensorimotor contingency theory

The agent is a sensorimotor contingency model. The knowledge acquired by the agent is not the result of direct sensory analysis but of the creation of an abstract representation built on the interaction between the sensory inputs and motor control via proprioceptive signals. This result is predicted by the sensorimotor contingency theory, where abstract notions do not reflect regularities in the sensory inputs per se but reflect robust laws describing the possible changes of sensory inputs following actions on the part of the agent. The set of φ_T is the space of internal representation where T is the transformation in the environment. By applying the compensation algorithm, we showed that the set of φ_T is indeed a space representation of the agent's environment.

Space knowledge without a priori knowledge of body or environment. It is important to note that the agent acquires its space representation without any a priori knowledge of either the structure of space or the group of transformations that describe it. There is no initial hypothesis that the agent is in space. The agent only has its sensory data, motor action, and the sensorimotor association. Furthermore there is no need for a strong hypothesis on the sensory information the agent needs to use. Only visual coincidence matching is used; no preprocessing of images or knowledge of the metric is necessary. It is important to note that no model for the environment is given to the agent and no assumptions are made about its body (proprioceptive organization or sensory capabilities).

Distinguishing external movement from internal action. Using this model, the agent is able to distinguish between movements of the environment and its own movements. During the learning phase with both stationary and non-stationary environments, the agent acquires the set of φ_T and can compare changes in the external environment to its proprioceptive changes. The auto-compensable transformations act as a reference for any external transformations of the environment (see Figure 5).

The set Φ of φ_T is a representation of space

Poincaré makes the distinction between sensible space and geometrical space. Sensible space is explicitly related to raw measures from different sensory systems. Poincaré argues that despite the major differences between these spaces, an agent can retrieve the properties of geometrical space from the sensible spaces by considering the effects of actions on the sensible space. As the geometrical space can be defined by the group of rigid transformations, if an agent can capture these rigid transformations and their group property, the agent acquires a representation of the geometrical space. As in our study, the agent does not have any a priori knowledge of the rigid transformations or their properties. However, using its sensorimotor system alone, the agent can acquire a subset of these transformations (the compensable transformations). Poincaré argued that the agent will acquire not only the set of compensable transformations but also learn that they behave as a mathematical group.

In the present article, the agent fully learned the set of compensable transformations and that this set was a group. More importantly, because the internal representation is isomorphic to the group of compensable transformations, the agent can also learn other information about it, for example, the metric or topology. Terekov has used a similar model to retrieve the metric.[14]

While the agent did not have any a priori knowledge of its body or environment, we made some assumptions about the mathematical functions of the model such as the bijectivity and invertibility of the functions Π and μ for the retina and foot proprioceptives states. Another assumption that we made was that the foot displacement was unbounded. Further generalization could use Marcel's formalism[15] of a surjective model for the proprioceptive states and include boundaries for the agent displacement.

While our theoretical framework does not limit or specify the type of compensable transformation, all the simulations were done using translation transformations. Other type of transformations may also be simulated and we are currently working on rotations. Furthermore, this work used the visual sensory system for the coincidence matching, but other sensory systems, such as auditory, tactile, or vestibular systems, could be tested. In this study, we applied the sensorimotor compensation theory to a two-dimensional geometrical space. However, we believe that a more general compensation theory could be formalized on any type of physical space (not necessarily a geometrical space) with its own specific types of compensable transformations.

Defects in φ_T give defects in the space representation

Our algorithm requires the signal function to be continuous and invertible. The number of retina cells is an important parameter as it affects the continuous and invertible properties of the sensorial signal. If the number of cells is too low, for different absolute retina positions, the sensory measure will not be unique. There calculated compensating position will be ambiguous. For translations where the compensation is exact (i.e. the agent's retina can find the exact same position the agent had before a transformation), there is no effect of retina size, number of cells, or proprioceptive signal as long as the conditions of a continuous signal and reversibility are met. When we analyzed the structure of the φ_T function for extreme parameters (one and two visual cells), we found that the agent was not able to properly recognize compensable movements. In the case of one single retina cell, the recognition was extremely limited. The effects of these extreme parameters on both transformation retrieval and group properties are shown in Tables 3 and 4.

Table 4. Testing the group propery of the function φ_T by calculating the combination of φ_{T_1} and φ_{T_2} in order to generate a φ_{T_3} where $T_3 = T_1 \circ T_2$.[a]

Parameters		Results
# Visual cells	# Sources	Correct group combination (%)
1	1	5
2	1	75.7
3	1	100
1	10	100
2	10	100
3	10	100

[a]Since the system combined properly with only three retina cells for one source and with any number of retina cells for 10 sources, further results are not included. The # Steps are 10×10 and the retina to agent step length ratio is 1.

The curves plotted in Figure 6 illustrate this effect. When the space representation is fully learned, the φ_T is well separated as in (a). However, this is not the case when the space representation is invalid or incomplete. For two retina cells (b), the φ_T curves occasionally overlap. They are totally invalid for one retina cell (c).

With such kind of handicap, an agent is not able to properly develop a representation of space.

Problem of the rotation

We presented in this article a general framework and exact mathematical proofs that are related to rigid transformation as rotation and translation. However, during the simulation, we have shown only results on the translation. It is important to note that we are currently working on rotation simulation. But rotations have two noticeable effects.

First, the simulation grid is not invariant by rotation but it is by translation. Thus, the rotation transformation shows similar defects on the representation of space as the handicapped agent. We have been able to find a solution to resolve this but giving the full explanation in this article would have been problematic.

Second, the rotation is periodic. Rotating by 2π is equivalent to no rotation in terms of sensory perception. This very interesting property creats complexity that will be presented in a future article.

These reasons forced us to not include rotation in the present article. But we consider that this point does not limit the results as the mathematical results are general for both translation and rotation and only the simulation was restricted.

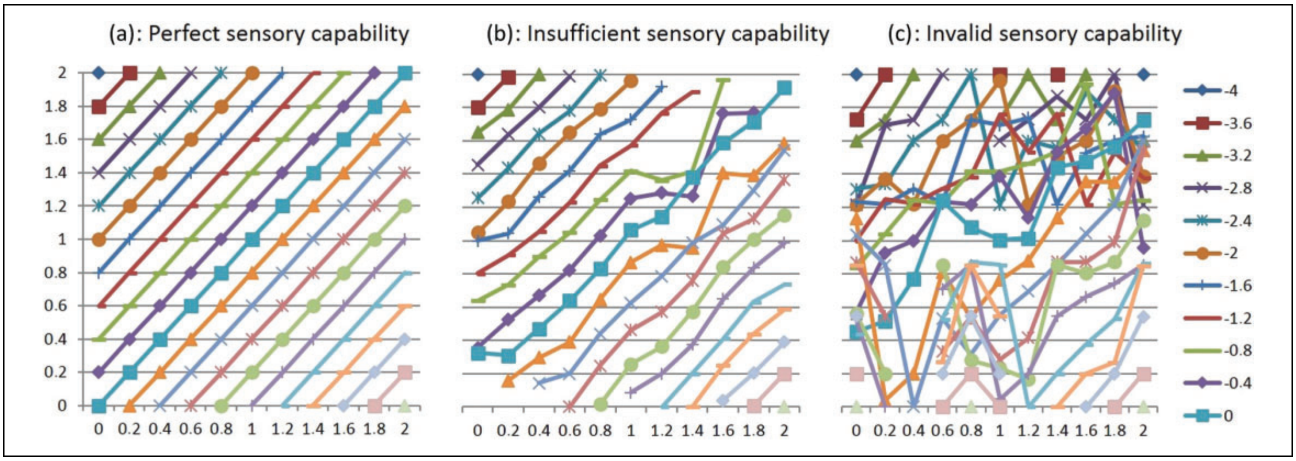

Figure 6. The set of φ_T curves for an affine uncoupled proprioceptive retina measure. Each curve corresponds to a learned movement T. (a) When the sensory system is large enough to compensate properly (continuity and reversibility conditions are met), the agent exactly compensates the movements and the curves are well separated and linear (for affine proprioceptive only); (b) when the sensory system is not sufficiently large (two retina cells), the compensatory algorithm yields ambiguous results and cannot retrieve a well-separated manifold of φ_T for different transformations T, resulting in an incorrect representation of space; and (c) when the sensory system is totally invalid (one retina cell), the algorithm cannot find a match and the curves are invalid. The algorithm does not generate a representation of space in this case.

Conclusion

General statistical learning algorithms for perceptual capabilities require a prerequisite model of the environment and body in order to acquire the ability to behave and generate actions. Space knowledge is predefined in the model and is thus restricted by any assumptions of the model. By contrast, our agent, whose learning is based on proprioceptive compensation (or more generally, algebraic learning), learns the properties of the surrounding space without any prior assumptions. The learned representation is a group, as proven in this article. Furthermore, the agent's internal representation can be used to distinguish the agent's own movements from those of the environment. Our algorithms and implementation allow the usage of φ_T for space computation. Future work includes a further theoretical formalization and simulations using other transformations such as rotations. We will also consider the acquisition of other types of knowledge, for example, object knowledge and arithmetic knowledge.

Appendix 1

Formal expression of the φ_T function

Figure 7 illustrates the compensation process. Starting in a initial state $\boldsymbol{p}_q$, the agent has a retina position $\left(X_q^r = \Pi(\boldsymbol{p}_q^r)\right)$ in its body and moves the body by a foot movement $\mu(\boldsymbol{p}_q^f)$. The environment is moved homogeneously by a transformation T. The absolute position of the retina in the environment changes by the combination of $\mu(\boldsymbol{p}_q^f) \circ T$. The agent compensates for the transformation by moving its body by a new foot movement $\mu(\boldsymbol{p}_{q'}^f)$ and retina movement τ positioning it at $X_{q'}^r = \Pi(\boldsymbol{p}_{q'}^r)$.

The compensation is equal to the initial transformation applied to the agent

$$\mu(\boldsymbol{p}_q^f) \circ T = \tau \circ \mu(\boldsymbol{p}_{q'}^f)$$
$$\mu(\boldsymbol{p}_q^f) \circ T \circ \mu(\boldsymbol{p}_{q'}^f)^{-1} = \tau$$

The final retina position in the body is

$$X_{q'}^r = \Pi(\boldsymbol{p}_{q'}^r) = \tau \circ X_q^r = \tau \circ \Pi(\boldsymbol{p}_q^r)$$
$$\boldsymbol{p}_{q'}^r = \Pi^{-1} \circ \tau \circ \Pi(\boldsymbol{p}_q^r)$$
$$\boldsymbol{p}_{q'}^r = \Pi^{-1} \circ \mu(\boldsymbol{p}_q^f) \circ T \circ \mu(\boldsymbol{p}_{q'}^f)^{-1} \circ \Pi(\boldsymbol{p}_q^r)$$

This gives us the φ_T function.

$$\varphi_T(\boldsymbol{p}_q) = \boldsymbol{p}_{q'}$$
$$\varphi_T(\boldsymbol{p}_q) = \boldsymbol{p}_{q'} = \begin{pmatrix} \boldsymbol{p}_{q'}^r \\ \boldsymbol{p}_{q'}^f \end{pmatrix}$$
$$\varphi_T(\boldsymbol{p}_q) = \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}_q^f) \circ T \circ \mu(\boldsymbol{p}_{q'}^f)^{-1} \circ \Pi(\boldsymbol{p}_q^r) \\ \boldsymbol{p}_{q'}^f \end{pmatrix}$$

In the case of a stationary environment, the previous formula simplifies to

$$\varphi_0(\boldsymbol{p}_q) = \boldsymbol{p}_{q'} = \begin{pmatrix} \Pi^{-1} \circ \mu(\boldsymbol{p}_q^f) \circ \mu(\boldsymbol{p}_{q'}^f)^{-1} \circ \Pi(\boldsymbol{p}_q^r) \\ \boldsymbol{p}_{q'}^f \end{pmatrix}$$

Implementation of the algorithm

The environment movements can be done on a grid. The length of the grid steps is Δe. The agent movements (foot and retina) are done on another grid of steps length Δb.

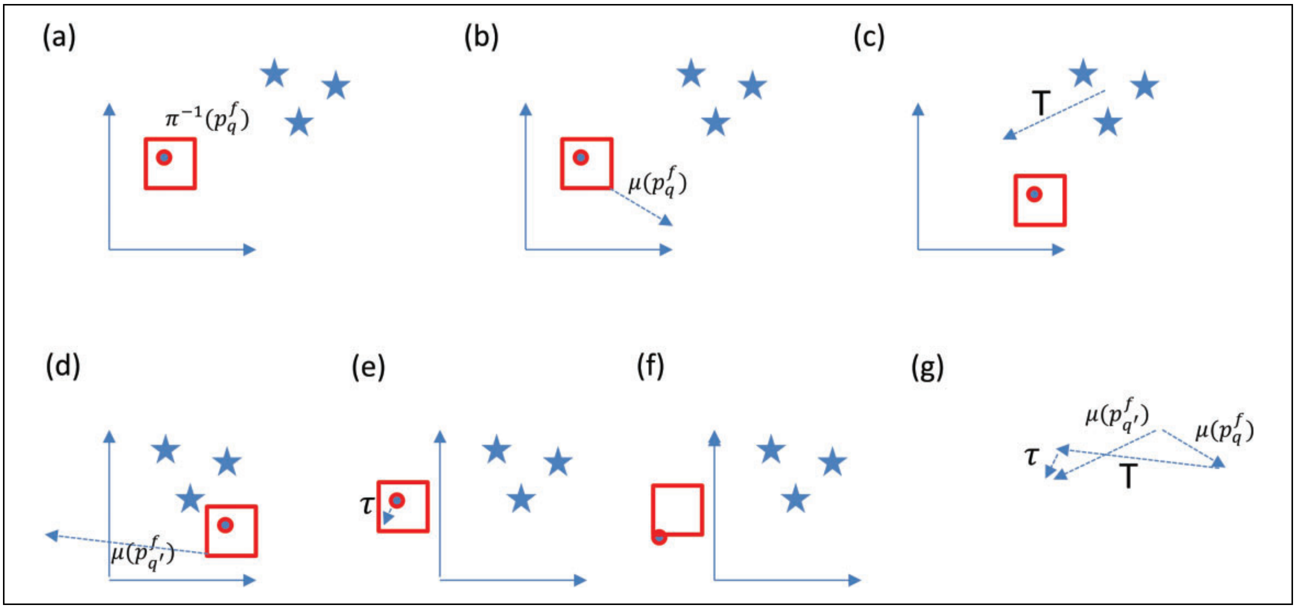

Figure 7. (a) The agent is in a physical space. The position of the retina in its body is determined by $\Pi(\boldsymbol{p}_q^r)$, (b) the agent moves its body by a foot movement $\mu(\boldsymbol{p}_q^f)$, (c) the environment is moved (all points in the environment are moved) following a transformation T, (d) the agent compensates with a foot movement $\mu(bfp_{q'}^f)$ and a new retina position $X^{r'} = \Pi(\boldsymbol{p}_{q'}^r)$, (e) The local movement of the retina is associated with a local transformation τ, (f) the final position of the retina relative to the environment is identical to its initial position, and (g) the applied transformations are equivalent $\mu(\boldsymbol{p}_q^f) \circ T = \tau \circ \mu(\boldsymbol{p}_{q'}^f)$.

Here is the algorithm in pseudo code: Loop over all agent position in the environment X^a. Loop over all prorpioceptives states $\boldsymbol{p} = \begin{pmatrix} \boldsymbol{p}^r \\ \boldsymbol{p}^f \end{pmatrix}$. Loop over all transformation of the environment T.

Compensate the transformation T knowing the initial state $\boldsymbol{p}$ by calculating the new proprioceptive state $bf\ p^{'}$ corresponding to a movement of the body and a new retina position.

The compensation is done by applying a coincidence visual signal matching before applying T and after applying T. The foot and the new retina position minimizing the distance between the initial signal and the compensated signal: $\boldsymbol{p} = \mathrm{Min}\left(\| \sigma(X^a, X^r), \sigma(T \circ X^{a'}, X^{r'}) \| \right)$. Calculate all proprioceptive states $\boldsymbol{p}$ for solving this equation.

Save the tuples $(T, \boldsymbol{p}, \boldsymbol{p}')$ as $\varphi_T(\boldsymbol{p}) = \boldsymbol{p}'$ with $\boldsymbol{p} = \begin{pmatrix} \boldsymbol{p}^r \\ \boldsymbol{p}^f \end{pmatrix}$ and $\boldsymbol{p}' = \begin{pmatrix} \boldsymbol{p}^{r'} \\ \boldsymbol{p}^{f'} \end{pmatrix}$

Continue looping for all combinations of $T, \boldsymbol{p}, \boldsymbol{p}'$ in order to generate all φ_T functions.

Acknowledgements

Gurvan Le Clec'H and Bruno Gas are now affiliated to Sorbonne Universités, UPMC Univ. Paris 06, UMR 7222, ISIR, F-75005 Paris, France and CNRS, UMR 7222, ISIR, F-75005 Paris, France.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: J. Kevin O'Regan acknowledges funding from ERC Advanced Grant 323674 "FEEL".

References

1. Berthoz A. *Le sens du mouvement*. Odile Jacob, Paris, 1997.
2. O'Regan JK and Noë A. A sensorimotor account of vision and visual consciousness. *Behav Brain Sci* 2001; 24(5): 939–1031.
3. Poincaré H. L'espace Et la Géométrie. *Revue de Métaphysique Et de Morale* 1895; 3(6): 631–646.
4. Poincaré H. On the foundations of geometry. *The Monist* 1895; 9: 1–43.
5. Poincaré H. La science et l'Hypothèse. *Flammarion*, Paris, 1902.
6. Nicod J. L'espace et la géométrie. *Revue de Métaphysique Et de Morale* 1921; 28: 537–543.
7. Philipona D, O'Regan JK and Nadal J-P. Is there something out there?: Inferring space from sensorimotor dependencies. *Neural Comput* 2003; 15(9): 2029–2049.
8. Philipona D, Kevin O'Regan J, Nadal J-P, et al. Coenen. Perception of the structure of the physical world using unknown sensors and effectors. *Adv Neural Inf Proces Syst* 2004; 15: 945–952.
9. Nicod J. *La Géométrie dans le monde sensible*. PhD thesis, Université de Paris, France, 1923.
10. Laflaquiere A, Argentieri S, Gas B, et al. Space dimension perception from the multimodal sensorimotor flow of a naive

robotic agent. In: *IEEE/RSJ international conference on intelligent robots and systems, IROS'10*, 2012, pp. 1520–1525.

11. Laflaquiere A, Gas B and Argentieri S. Perception de la dimension de l'espace via une exploration active. In: *Reconnaissance des Formes et Intelligence Artificielle, RFIA'2012*, 2012.
12. Aytekin M, Moss CF and Simon JZ. A sensorimotor approach to sound localization. *Neural Comput* 2008; 20(3): 603–635.
13. Terekhov AV and O'Regan JK. Space as an Invention of Active Agents. *Front Robot* 2016. AI 3:4.doi: 10.3389/frobt.2016.00004.
14. Terekhov AV and O'Regan JK. Learning abstract perceptual notions: the example of space. In: *Development and learning and epigenetic robotic (ICDL-Epirob)*, 2014; pp. 368–373.
15. Marcel V, Garcia B, Argentieri S, et al. Building the representation of an agent body from its sensorimotor invariants. In: *IEEE/RSJ international conference on intelligent robots and systems, IROS'15*, 2015, pp. 39–44.
16. Roschin VY and Frolov AA. A neural network model for the acquisition of a spatial body scheme through sensorimotor interaction. *Neural Comput* 2011; 23: 1821–1834.
17. Laflaquière A. *Approche sensorimotrice de la perception de l'espace pour la robotique autonome*. PhD thesis, Sorbonne Universités, Université Pierre et Marie Curie, France, 2013.