A Self-organizing Network with Varying Density Structure for Characterizing Sensorimotor Transformations in Robotic Systems

Omar Zahra[(✉)] and David Navarro-Alarcon

Department of Mechanical Engineering,
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
`omar.zahra@connect.polyu.hk`, `dnavar@polyu.edu.hk`

Abstract. In this work, we present the development of a neuro-inspired approach for characterizing sensorimotor relations in robotic systems. The proposed method has self-organizing and associative properties that enable it to autonomously obtain these relations without any prior knowledge of either the motor (e.g. mechanical structure) or perceptual (e.g. sensor calibration) models. Self-organizing topographic properties are used to build both sensory and motor maps, then the associative properties rule the stability and accuracy of the emerging connections between these maps. Compared to previous works, our method introduces a new varying density self-organizing map (VDSOM) that controls the concentration of nodes in regions with large transformation errors without affecting much the computational time. A distortion metric is measured to achieve a self-tuning sensorimotor model that adapts to changes in either motor or sensory models. The obtained sensorimotor maps prove to have less error than conventional self-organizing methods and potential for further development.

Keywords: Self-organizing maps · Sensorimotor models ·
Associative learning · Adaptive systems · Robot manipulators ·
Motor babbling

1 Introduction

Among the many interesting cognitive abilities of animals and humans is the motor babbling process that leads to the formation of the sensorimotor map. Many theories have been introduced about how these behaviors develop since prenatal stages [1]. This sensorimotor adaptation paradigm has proven to be useful in robotics for relating motor commands with sensory outputs when

prior/exact knowledge of its model is unavailable (which is common in practice). A robot with changes to its mechanical structure (e.g. due to damage) or perceptual system (e.g. due to sensor decalibration) is generally not able to properly coordinate its motions without updating such sensorimotor relations. Drawing inspiration from the adaptive properties of living organisms, artificial neural systems can be developed to cope with these uncertainties. The development of computational sensorimotor models with adaptation properties can lead to the emergence of valuable self-calibrating behaviors. Additionally, these could help to (safely) verify theories about the internal workings of the human brain, but with machines.

Previous studies with primates have concluded a topographic arrangement in areas dedicated to motor and sensory processing, where adjacent body parts tend to have an adjacent representation in the brain cortex [2,3].

Thus, to represent perceptual computing units in a biologically inspired manner, such topologically preserving property was considered.

Topographic models are useful for characterizing sensory and motor spaces in robots. Yet, to co-relate how a particular motion/configuration produces a sensory stimuli, additional associative properties must be considered. One common model for linking different brain areas based on shared activity patterns is the so-called Hebbian rule [4]. It states that if two neuronal regions are persistently activated together, the connection between them is strengthened; the connection is weakened if no simultaneous activity is present. Topographic and associative properties are the basis for the sensorimotor adaptive method that we propose in this paper.

In the literature, many efforts have been placed to model human sensorimotor abilities with methods based on self-organizing maps (SOM) [5–7]. Most of these works use SOMs as a topography-preserving and dimension-reducing tool, to map several sensor readings with motor actions.

In [6], an SOM is used to form a sensory map with visual feed. However, the learning process to form a sensorimotor map takes place mainly through gradient-descent rule which makes it less biologically plausible. In [5], two dynamic SOMs (DSOM) [8] representing head and arm of a humanoid robot were used to achieve visuo-motor coordination. Yet, that model suffered from a degradation in performance when perturbations were added to motor commands. In [7], the sensorimotor coordination is achieved by utilizing bi-directional neural modularity such that motor output can be predicted from sensory input and vice versa. For the proposed method in this paper, the learning paradigm allows for the development of reciprocal correlations inherently while maintaining high accuracy.

In this study, we propose a new method for representing sensorimotor transformations of robotic systems. The neuro-inspired method combines self-organizing and associative properties to model continuously adapting relations between sensory and motor spaces. Compared to previous works, our new method proposes a varying density SOM (VDSOM) that reduces the transformation error that is typically present at the periphery of standard SOMs. This is done by automatically adjusting a parameter that controls the density of

neighboring nodes at regions with large transformation errors. In case of changes in either motor or sensory models, a distortion metric is measured to readjust the formed sensorimotor map to suit these changes. The resulting computational model can effectively reduce the mean error over the whole map, while coping with changes in the original sensorimotor model. Several cases of study (such as transformation accuracy, amputation, limb extension) are presented to thoroughly evaluate the proposed method.

The rest of this paper is organized as follows: Sect. 2 describes the computational model; Sect. 3 presents its quantitative evaluation; Sect. 4 gives final conclusions.

2 Methods

2.1 A Biologically-Inspired Sensorimotor Model

Human bodies have different morphologies which develop over years (from birth to death) and even subject to drastic changes as in the case of amputations. However, the brain somehow always manages to find or re-adapt such mappings between sensory feedback and motor actions. In infants, for example, motor babbling helps to adaptively obtain these sensorimotor relations, where by performing motions covering the workspace, the brain is able to correlate bodily configurations with its corresponding motor actions [9,10].

It is also clear from recent studies that in both sensory and motor areas in the brain, adjacent body parts have also contiguous representations [11]. Moreover, many of these areas are connected together by some synapses which develop connections based on their joint activity. Among these rules is the well-known Hebbian learning rule [4].

To represent such learning paradigm, a model for human sensorimotor mapping is constructed using SOM (modeling topographically arranged brain areas) and Hebbian learning rules (modeling connections among these areas). Both of these models have clear biologically-inspired properties as they can represent topographic organization of neurons and modulation of strength of synaptic connections, respectively.

SOM are built upon the underlying rules of development of cognitive functions, as they encode competition, cooperation and adaptation [12]. The nodes (neurons) of the SOM compete against each other such that only one becomes the best matching neuron (BMU) for a given input. However not only the BMU contribute to the output, but also the neighboring neurons as well, such that the closer to the BMU the greater would be the contribution to the output. This represents the lateral interaction between neurons in a network. Adaptation by modulation of weights of BMU (and neighborhood nodes) occurs to enhance the chance of the BMU to represent the input vector and act as the BMU again for a similar input.

The Hebbian learning rule wires the SOMs representing the sensory space and motor space together, such that neurons active on both sides at the same time have the strength of the synaptic connection in between increased proportional to the magnitude of activity of both the pre-synaptic and post-synaptic neurons.
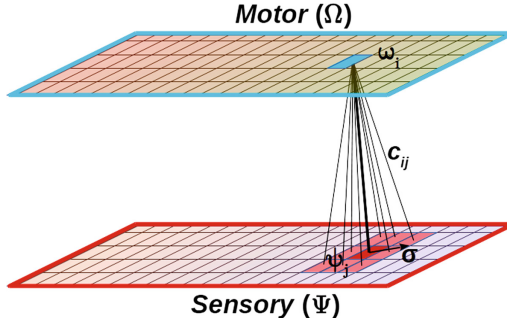
Fig. 1. Motor space (Ω) connected to Sensory space (Ψ) through Hebbian connections ($\boldsymbol{c_{ij}}$). As learning process proceeds, active nodes in Ω (ω_i) have the connections to active ones in Ψ reinforced (ψ_j and neighborhood within a radius of σ).

These connections achieve sensorimotor correlation between the motor actions and the corresponding sensory input that happen to be active at the same time.

2.2   Modeling Sensory and Motor Spaces

The SOM is formed of a 2 dimensional lattice of M neurons (nodes), each of them associated with a weight vector (w_i) of dimension as each vector in the input space (X). These weights are initially set to random values, then, data training points are introduced in a random fashion to the SOM. When a vector of data x is introduced to the SOM, the node with least Euclidean distance between its weights and the input vector is chosen to be the best matching unit (winning neuron) based on:

$$i = \arg \min_{j} \|w_j - x\| \tag{1}$$

where i the denotes the index of the BMU. The weights of all the neurons in the neighborhood around the BMU are updated to give a closer approximation of the input vector x. This node is computed with the following update rule:

$$w_j(t + 1) = w_j(t) + \alpha(t)h_{ji}(t)(x - w_j(t)) \tag{2}$$

where h_{ji} is the neighboring function, which is computed with the following Gaussian function:

$$h_{ji}(t) = \exp\left(\frac{-\|r_j - r_i\|^2}{2\sigma^2(t)}\right) \tag{3}$$

where r_i and r_j are the positions of the BMU and the neighboring jth node within the lattice, respectively. The learning rate α and neighborhood radius σ are set to decrease exponentially with time such that:

$$\sigma(t) = \sigma_{init} \exp\left(\frac{-t}{T}\right), \ \alpha(t) = \alpha_{init} \exp\left(\frac{-t}{T}\right) \tag{4}$$

where t is the time of current iteration, T is the desired time constant of decrease, α_{init} and σ_{init} are the initial values of the learning rate and neighborhood radius, respectively. By tunning the adequate parameters for the learning process, the weights of the nodes are updated to give an adequate mapping for both sensory and motor states within the identified robot workspace.

2.3   Formation of Sensorimotor Mapping

To provide a correlation between activity of each node in motor space a_i to sensory space a_j back and forth as shown in Fig. 1, the Hebbian Oja learning rule [13] is applied by applying the equation:

$$c_{ij}(t+1) \ = \ c_{ij}(t) \ + \ \eta \left(a_i a_j - c_{ij}(t) a_j^2 \right) \tag{5}$$

where c_{ij} represents the strength of the connection between the pre and post synaptic nodes, while η is the learning rate. Nodes from both maps that happen to be active at the same time tend to have high correlations and thus stronger synaptic connection between these nodes. The first term in the parenthesis ensures applying Hebbian learning rule to achieve the correlation. The second term guarantees the stability of the learning process where a forgetting term is included such that in case some nodes are not active for a long time the strength of the connection is attenuated.

The activity a_j of each node is calculated by applying the following Gaussian kernel for the Euclidean distance between the weights of the nodes and the input vector:

$$a_j(t) \ = \ \exp \left(\frac{-\|w_j(t) - x\|^2}{\sigma^2(t)} \right) \tag{6}$$

This expression gives rise of a one-to-one mapping between the nodes of the two SOMs (that respectively model the motor and sensory spaces). The resulting connections are reciprocal (i.e. bidirectional). This means that they can be used to either predict the sensory states based on a given motor action, or to compute the required motor actions to achieve a certain sensory state [14].

2.4   Varying Density Structure

The sensorimotor mappings can be achieved by combining SOM and Oja-Hebbian learning rules, as described above. However, the naive use these method results in regions (e.g. the periphery of the lattice) with large transformation errors. Two initial hypothesis were assumed to cause this problem. The first is that having a small number of training points at these regions may cause that problem. The second one is that having comparatively low number of neurons near the boundaries to represent the sensorimotor correlations may the culprit (e.g. having fewer neurons affect the accuracy of the estimated values). Such problem at the boundaries is one of the drawbacks of the SOM mentioned in the literature [15].

For the former hypothesis, training data with higher density at the lattice boundaries was used, however, it did not improve the mapping accuracy. A viable solution was to increase the density of the neurons near the problematic regions such as the boundaries of both the sensory and motor maps to give a better representation at these points. To achieve this behavior, the SOM update rule was modified by proposing a different neighborhood function that gives the required variable density of nodes. This is done by calculating the summation of the norm of the weights of the BMU to each node in the lattice then applying the Gaussian function. The *node density coefficient* ρ is computed as follows:

$$\rho \,=\, \exp\left(-\sum_{i\in O}\|w_{bmu} \,-\, w_i\|^2\right) \tag{7}$$

for O as the local neighborhood surrounding the neuron. This function aims to give a smooth gradient effect of contribution of proximal nodes.

The coefficient ρ can be used to quantitatively determine neurons with a small number of neighbors. More neurons can be attracted to these nodes to have a denser population and therefore give a better approximation of corresponding values in the sensorimotor map. The resulting map is characterized by having a variable density (even when using uniform training data) that controls the number of nodes in a region based on ρ; we call this network a varying density SOM (VDSOM). The additional term ρ shall have a minimal effect in the formation of the network at the beginning and increase as the learning process proceeds. On the other hand, if it increases at a slow rate the exponential decay term of the neighborhood radius would make the effect of that term minimal.

To achieve this effect, the new neighborhood is defined as follows:

$$h(t) \,=\, \left(\frac{t}{\rho T}\right)^4 \exp\left(\frac{-t}{\sigma^2(t)T}\right) \tag{8}$$

where the new term was chosen to be of the fourth order to have adequate values without disturbing the dynamics of the learning process.

By adding that term, the lattices formed for both the sensory and motor spaces are more dense at the boundaries. This helps to reduce the transformation errors that occurs in these regions without the need to increase the total number of neurons in the network. This density regulation concept may not (yet) have some proof from a neuro-biological perspective, however, varying densities of neurons is certainly present in many different areas of the brain and within each area. For example, in the primary visual cortex, the central region has a higher density of neurons relative to the peripheral regions. In most primates, the central vision area is the main region of interest when observing a scene [16]. Additionally, the proposed mechanism to automatically increase the number of neurons agrees with studies in which high neuronal density is observed for processing hand and face fine motions [17]. Although this study focuses on VDSOM, the same concept can be applied to vary the structure of a Growing Neural Gas(GNG) network [18] to obtain the optimal number of nodes to represent the same sensorimotor model.

2.5   Adaptation to Changes in the Sensorimotor Model

Note that in case of changes in body morphology (e.g. generated by attaching of an external limb or amputation) or changes in the perceptual system (e.g. by wearing vision inverting goggles [19]) the computed sensorimotor model is no longer representative. For this situation, both, sensory and motor maps should be updated accordingly, as well as the inter-connections representing the transformations between these spaces. However, in traditional SOM, once the learning process reaches the specified number of iterations, changes in input data—corresponding to sensory/motor information—will not modify the networks structure. This results in a model that no longer adapts, and therefore is not able to represent the new (and actual) sensory/motor configurations.

To overcome this drawback, a distortion metric ζ is incorporated into the method. If the ζ is found to exceed a give (arbitrary) threshold value after the mapping is established, the neighborhood radius σ is reset to an adequate value to be able to re-adapt the network's structure. Such distortion metric is computed as:

$$\zeta = \frac{1}{n} \sum_{i=1}^{n} \sum_{x \epsilon X} \|x - w_i\|^2 \tag{9}$$

where n is the number of data vectors x available in the data set X. The new neighborhood radius σ_r is set to be initially equal to $\sigma(\tau)$, when the distortion metric after the perturbations occur is equal to that. Then, the value of $\sigma_r(t)$ can be calculated from the equation:

$$\sigma_r(t) = \sigma_{init} exp \left(\frac{-(t + \tau)}{T} \right) \tag{10}$$

If the value of distortion after perturbations is higher than that at the beginning of the learning process, then the radius is set to the maximum value which is the radius of the SOM. On the other hand, a modified version of Oja-Hebbian connections is used to adapt better to these changes.

$$c_{ij}(t + 1) = c_{ij}(t) + \eta(a_i a_j - \beta c_{ij}(t) a_j^2) \tag{11}$$

The additional term allows to control the *forgetting rate* of the already formed connections.

Thus the values of the additional term β and the learning rate η are set to allow for new connections to be formed in a faster manner. These terms

Table 1. Mean and maximum errors for forward and inverse mappings using SOM and VDSOM.

Error	X (mm)	Y (mm)	θ_1 (°)	θ_2 (°)
	Mean (Max)	Mean (Max)	Mean (Max)	Mean (Max)
SOM	2.6 (26.0)	2.9 (29.0)	0.37 (3.45)	0.61 (6.70)
VDSOM	1.15 (11.7)	1.26 (15.0)	0.31 (2.25)	0.44 (5.36)

Table 2. Mean and maximum errors for forward and inverse mappings using VDSOM with different number of nodes.

Error	X (mm)	Y (mm)	θ_1 (°)	θ_2 (°)
	Mean (Max)	Mean (Max)	Mean (Max)	Mean (Max)
VDSOM (30×30)	4.16 (36.3)	3.93 (21.1)	0.91 (4.60)	1.53 (7.62)
VDSOM (50×50)	2.13 (13.5)	2.11 (15.3)	0.52 (2.80)	0.80 (5.05)
VDSOM (70×70)	1.15 (11.7)	1.26 (15.0)	0.31 (2.25)	0.44 (5.36)

are assigned high values that decrease exponentially based on the following expressions:

$$\beta(t) = \beta_{init} \exp\left(\frac{T-t}{T}\right), \quad \eta(t) = \eta_{init} \exp\left(\frac{T-t}{T}\right) \tag{12}$$

3   Results

3.1   Setup

A simulation for the computational model of the sensorimotor mapping was built using Tensorflow library [20] on a PC with i7-6500 16 GB RAM. Both the system without and with the modifications were simulated for 2D lattice SOMs with square grid of 30×30, 50×50 and 70×70 nodes.

A kinematic model of two link robotic arm was used as the prototype system. The end-effector task space is assumed to be measured with an external positions sensor (e.g. a camera). In our sensorimotor model, the joint space is represented with motor SOM, whereas the task space is represented with a sensory SOM. Random joint angles within certain ranges were used to generate end-effector positions. L_1 and L_2 denote the lengths of first and second link, respectively, θ_1 and θ_2 the joint angles of first link relative to the horizontal axis and joint angle of second link relative to the first link. The forward kinematics relation can be simply computed as:

$$\begin{aligned} X &= L_1 \cos(\theta_1) + L_2 \cos(\theta_1 + \theta_2) \\ Y &= L_1 \sin(\theta_1) + L_2 \sin(\theta_1 + \theta_2) \end{aligned} \tag{13}$$

The connections between both joint space and task space SOMs were developed, as described above, based on the Oja-Hebbian learning rule. As can be seen from Table 1 that both the mean and the maximum error values were drastically reduced after applying the proposed solution to the SOM for forward and inverse mappings. It can also be concluded from Table 2 that as the number of nodes increases the error decreases at the expense of increasing the computational time required to build the network and establish the connections.
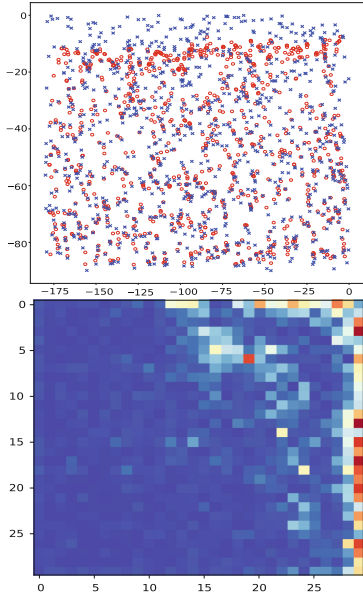
Fig. 2. Motor SOM with heatmap representing error at each node when it was chosen as a BMU. (Color figure online)
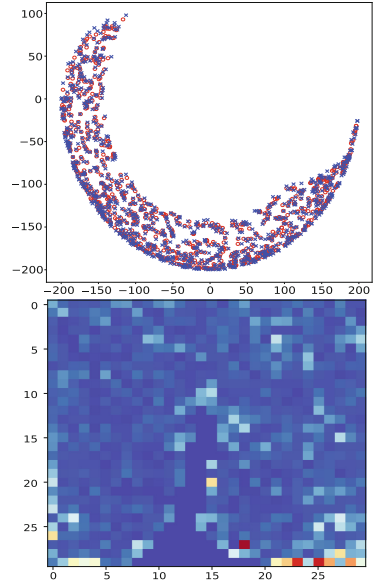

Fig. 3. Sensory SOM with heatmap representing error at each node when it was chosen as a BMU. (Color figure online)
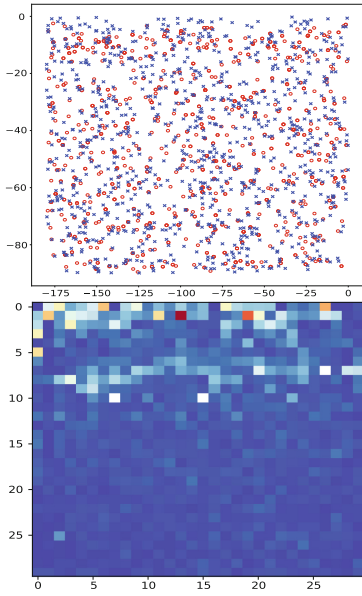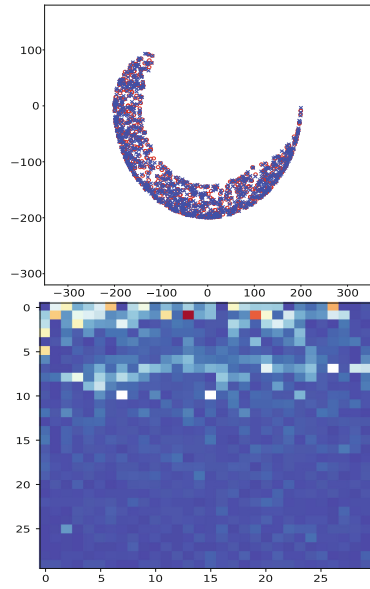

Fig. 4. Motor VDSOM with error heatmap.

Fig. 5. Sensory VDSOM with error heatmap.

3.2   Enhanced Accuracy

Figures 2 and 3 show the final SOMs developed after running several trials to obtain the most adequate parameters for each SOM. As shown in Fig. 3, the original SOM covers the whole workspace uniformly but less dense at the peripherals. It can be observed from the error plot in Figs. 2 and 3 that higher error values occur at these areas, where the dark blue color and the dark red color represent low error and high error, respectively. The effect of the added factor ρ can be noticed in Figs. 4 and 5 where higher density can be observed at the contour of the workspace, and less error in these areas in both forward and inverse mappings. Although the introduced method have an error that is relatively high compared to conventional control methods, it takes one step forward in the formation of biologically-inspired sensorimotor maps.

3.3   Adaptation to Changes in Morphology

The robot morphology was altered to simulate attaching and removing a tool from the end effector. To allow the system to detect such changes, the ζ is calculated based on Eq. (9) and compared with a threshold value. Consequently, when such changes are predicted to occur, the learning process is reset to update the mapping. In case of limb length extension, it can be concluded from Figs. 6 and 7 that the map adapts to re-accommodate for that change and decreases the distortions detected in the computed maps. The connections between the sensory and motor maps are updated to represent the new configuration. Similarly, in the case of limb length reduction, as shown in Figs. 8 and 9, the distortion is measured. The change in distortion value, triggers the adaptation mechanism that allows for the maps to be recomputed.
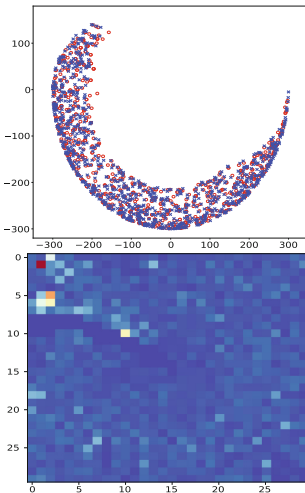

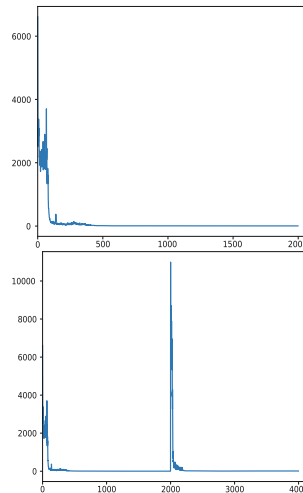**Fig. 6.** Sensory VDSOM after stretching the links with error heatmap.

Fig. 7. Distortion in sensory map before and after stretching the links.
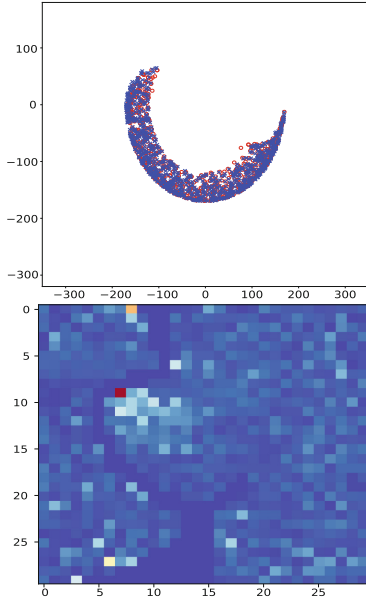
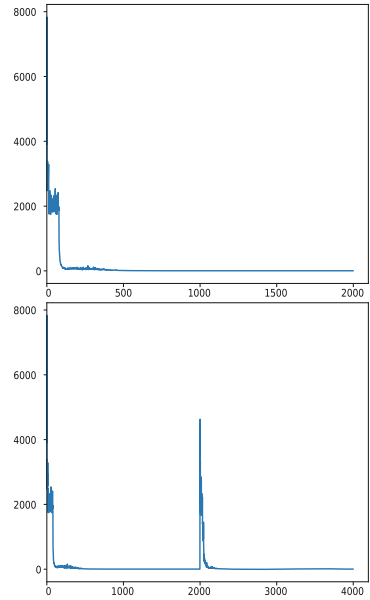**Fig. 8.** Sensory VDSOM after shortening a link with error heatmap.

Fig. 9. Distortion in sensory map before and after shortening the link length.

4   Conclusions

A sensorimotor map was built to correlate sensory and motor spaces in a discretized form with bidirectional connections. This solution relies on collecting data samples by motor babbling, thus it is adequate to be used for various robotic manipulators without any prior information about robot kinematics. Using the SOM introduced by Kohonen with Oja-Hebbian learning rules the mapping was achieved with noticeable error values at the contour of the SOM -and thus the workspace-. A new neighborhood function was proposed to increase the density of nodes at the contour to give better approximation for the corresponding values. The proposed neighborhood increases the density of the nodes wherever the distance between the weights of the BMU and the neighboring nodes has small values. Finally, a perturbation was introduced to simulate a change in either sensory or motor map. A distortion metric was used to assess the state of the robot and reset the learning parameters to adequate values in case of changes in the morphology. Thus adaption process takes place to update the sensorimotor map, by allowing for changes in both the formed VDSOMs and connections.

Concerning the current limitations of this method, these maps can only be used for coarse control. A large number of nodes would be needed for fine discretization of the workspace which is computationally inefficient. Additionally, an extended study is needed to utilize the dimension reduction properties of SOM to be fit for robots with higher degrees of freedom.

References

1. Zoia, S., et al.: Evidence of early development of action planning in the human foetus: a kinematic study. Exp. Brain Res. **176**(2), 217–226 (2007). https://doi.org/10.1007/s00221-006-0607-3

2. Kaas, J.H.: Topographic maps are fundamental to sensory processing. Brain Res. Bull. **44**(2), 107–112 (1997)

3. Silver, M.A., Kastner, S.: Topographic maps in human frontal and parietal cortex. Trends Cogn. Sci. **13**(11), 488–495 (2009)

4. Hebb, D.: The Organization of Behavior, 1949. Wiely, New York (2002)

5. Schillaci, G., Hafner, V.V., Lara, B.: Online learning of visuo-motor coordination in a humanoid robot a biologically inspired model. In: 2014 Joint IEEE International Conferences on Development and Learning and Epigenetic Robotics (ICDL-Epirob), pp. 130–136. IEEE (2014)

6. Kumar, S., Premkumar, P., Dutta, A., Behera, L.: Visual motor control of a 7DOF redundant manipulator using redundancy preserving learning network. Robotica **28**(6), 795–810 (2010)

7. Buessler, J., Kara, R., Wira, P., Kihl, H., Urban, J.: Multiple self-organizing maps to facilitate the learning of visuo-motor correlations. In: IEEE International Conference on Systems Man and Cybernetics, vol. 3, pp. 470–475 (1999)

8. Rougier, N., Boniface, Y.: Dynamic self-organising map. Neurocomputing **74**(11), 1840–1847 (2011)

9. Piaget, J., Cook, M.T.: The origins of intelligence in children. International Universities Press, New York (1952)

10. Von Hofsten, C.: Eye-hand coordination in the newborn. Dev. Psychol. **18**(3), 450 (1982)

11. Penfield, W., Boldrey, E.: Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain **60**(4), 389–443 (1937)

12. Kohonen, T.: Self-organizing Maps. Springer Series in Information Sciences, p. 502. Springer, Heidelberg (2001). https://doi.org/10.1007/978-3-642-56927-2

13. Oja, E.: Simplified neuron model as a principal component analyzer. J. Math. Biol. **15**(3), 267–273 (1982). https://doi.org/10.1007/BF00275687

14. Saegusa, R., Metta, G., Sandini, G., Sakka, S.: Active motor babbling for sensorimotor learning. In: International Conference on Robotics and Biomimetics, pp. 794–799, February 2009

15. Kohonen, T.: Essentials of the self-organizing map. Neural Netw. **37**, 52–65 (2013)

16. Collins, C.E., Airey, D.C., Young, N.A., Leitch, D.B., Kaas, J.H.: Neuron densities vary across and within cortical areas in primates. In: Proceedings of the National Academy of Sciences, vol. 107, no. 36, pp. 15927–15932 (2010)

17. Young, N.A., Collins, C.E., Kaas, J.H.: Cell and neuron densities in the primary motor cortex of primates. Front. Neural Circuit. **7**, 30 (2013)

18. Fritzke, B.: A growing neural gas network learns topologies. In: Advances in Neural Information Processing Systems, pp. 625–632 (1995)

19. Kohler, I.: Experiments with goggles. Sci. Am. **206**(5), 62–73 (1962)

20. Abadi, M., et al.: TensorFlow: large-scale machine learning on heterogeneous systems (2015). http://tensorflow.org/