

Interpersonal Maps: How to Map Affordances for Interaction Behaviour

Verena V. Hafner¹ and Frédéric Kaplan²

¹ Humboldt-Universität zu Berlin, Institut für Informatik, Berlin, Germany
hafner@informatik.hu-berlin.de

² Ecole Polytechnique Federale de Lausanne, CRAFT, Lausanne, Switzerland
frederic.kaplan@epfl.ch

Abstract. In a study of how the concept of affordances could be applied to interaction behaviour, we introduce the notion of “interpersonal maps”, a geometrical representation of the relationships between a set of proprioceptive and heteroceptive information sources, thus creating a common representation space for comparing one’s own behaviour and the behaviour of others. Such maps can be used to detect specific types of interactions between agents such as imitation. Moreover, in cases of strong couplings between agents, such representations permit to map directly an agent’s body structure onto the structure of an observed body, thus addressing the body correspondence problem. These various cases are studied with several robotic experiments using four-legged robots either acting independently or being engaged in delayed imitation. Through a precise study of the effects of the imitation delay on the structure of the interpersonal maps, we show the potential of this “we-centric” space to account for both imitative and non imitative interactions.

1 Introduction

In 1977, perceptual psychologist J. J. Gibson defined an affordance as “a resource or support that the environment offers an agent for action, and that the agent can directly perceive and employ” [1]. Even though people are disagreeing about the possible applications of this theory in the cognitive sciences, affordances can be seen as a useful theory of interaction for many disciplines including robotics. Affordances link perception and action depending on the current goals or intentions of an agent. Gibson also stated that affordances are not classifications of objects, but rather a function-centered view, and therefore provide a more intuitive view of oneself in a certain environment or situation. An everyday example for this view are sorting systems in human environments. Following Gibson’s theory of affordances, putting a priority on the affordance of an object (e.g. cutting, connecting) is more intuitive than sorting objects by their name or appearance. Humans can also employ attentional processes to focus on the perception of a particular affordance. This leads to joint attention [2] in the case of interaction behaviour.

Most of the research in affordances has focused on interactions between an agent and its environment [3]. However, we believe that affordances are also a relevant concept in the case of the interaction between two agents. In another research, we have explored

a novel framework permitting unsupervised activity classification [4] based on coordination patterns. Here, we introduce and explore the concept of “interpersonal maps”, the application of the same idea to interaction behaviour. This concept is defined in the framework of information theory and can be applied in the context of interactions between living entities as well as between artifacts. Although information theory has historically been mainly concerned with information transmission between a sender and a receiver [5,6], several lines of research have focused on addressing issues concerning relationships between information sources [7,8,9]. In particular, it has been shown that the space of information can be equipped with a metric [10]. It is therefore possible to adapt some of the vocabulary and tools of geometry to the domain of information theory. Interpersonal maps are geometrical representations of relationships between a set of information sources.

The notion of interpersonal maps is related to several existing concepts in psychology and neuroscience. To account for early imitation, Meltzoff and Moore argue for the existence of an intermodal mapping establishing equivalence relations between different modalities such as vision or motor actions [11,12]. Such a model suggests that both perceived (self) and observed (others) behaviour could be represented in a shared neural format. Similarly, Gallese has argued that since the beginning of our life we inhabit a shared multidimensional interpersonal space. When we observe other individuals, “a meaningful embodied interpersonal link is established”. Gallese refers to this form of intersubjectivity as the *shared manifold space*. Furthermore, his theory predicts the existence of “somatosensory mirror neurons” giving the capacity to map different body locations during the observation of the bodies of others [13]. However, few models try to give a precise account on how such interpersonal or intermodal mappings could be developed.

The approach presented in this article is directly inspired by several methods concerning unsupervised map building recently described in the field of artificial intelligence and autonomous robotics. Pierce and Kuipers present a method for building maps of a sensory apparatus out of raw uninterpreted sensory data [14,15]. This so-called sensory reconstruction method is based on various distances between sensors such as a normalised Hamming distance metric and a frequency metric. Sensors are clustered into subgroups based on their relative distance. The dimensionality of each subgroup can then be computed, related sensors can be projected to form a sensor map. Building on this sensory reconstruction method, Olsson, Nehaniv and Polani [16] have suggested to use the information metric defined by Crutchfield [10] as a more interesting measure of the distance between two information sources. They have conducted experiments with various sensor sets including visual and proprioceptive sensors on an AIBO robot. Related approaches were also investigated by Kuniyoshi’s research team [17]. Most of these approaches interpret such sensory reconstruction methods as a way of building maps of sensors in an unsupervised manner. Some of these works make the comparison with somatosensory maps discovered in the brain.

We extend and, more importantly, reinterpret the sensory reconstruction method. The sensory reconstruction method is well-adapted to address processes underlying the emergence of behavioural complexity, but it may be misleading to interpret it only as a formation of a body map. A particular set of distances captures not only aspects of an

agent's embodiment, but can also reflect the agent's current activities and the situated nature of its interaction with the environment. In particular, a specific configuration may appear in the case of couplings with other agents or in cases of remarkable coordination patterns, thus allowing the system to be interpreted as an application of the theory of affordances. We will now present the approach in a more formal manner and give results of preliminary experiments showing how this framework can be used in the context of robot-robot imitation.

2 Maps Based on Information Distances

This section reviews and illustrates the basic principles of map construction based on information distances as it is used in several other articles (e.g. [14,16,4]). This will provide the basic elements for introducing in the next section the notion of interpersonal maps.

2.1 Definition

Distance Between Information Sources. Let us assume that the robot R_X is equipped with n sensors (proprioceptive and distance sensors). At any time t its sensory state can be captured by the vector $X(t)$

$$X(t) = (X_1(t), X_2(t), \dots, X_n(t)) \quad (1)$$

For any sensor X_i the entropy $H(X_i)$ can be calculated as

$$H(X_i) = - \sum_{x_i} p(x_i) \log_2 p(x_i)$$

where $p(x_i)$ is the probability mass function over all possible discretised values x_i . To calculate it, the distribution of the values of X_i has to be computed with a careful choice of the number of bins (see [18]). A good solution to avoid this problem is to introduce adaptive binning [19]. In such a case, the size of the bins is variable and chosen in a way that maximises the entropy for each sensor.

The conditional entropy for two sensors X_i and X_j can be calculated as

$$H(X_j|X_i) = - \sum_{x_i} \sum_{x_j} p(x_i, x_j) \log_2 p(x_j|x_i)$$

where $p(x_j|x_i) = p(x_j, x_i)/p(x_i)$.

Crutchfield defines the information distance between two information sources as:

$$d(X_j, X_i) = H(X_i|X_j) + H(X_j|X_i) \quad (2)$$

and the normalised information distance as

$$d_N(X_j, X_i) = \frac{H(X_i|X_j) + H(X_j|X_i)}{H(X_i, X_j)} \quad (3)$$

set of constraints optimally, but in order to get a lower dimension representation an approximation must be taken. Pierce and Kuipers describe a method used by statisticians to determine a good dimensionality for projecting a given set of data [14]. In the rest of the article, two-dimensional projections are used for illustrative purposes although they may not be the optimal ones.

In order to create a two-dimensional map we can apply a relaxation algorithm. The algorithm is an iterative procedure of positioning the sensors in a two-dimensional space in such a way that the metric distance between two sensors in this map is as close as possible to the distance in the n -dimensional information space. Different algorithms exist in the literature [22,23,24]. Here, the algorithm of Pierce is used since it does not require any information about the relative orientation of connections between sensor nodes [24].

More precisely, the algorithm used in this paper consists of an iteration of two simple steps. Before these two steps, each sensor X_i is randomly assigned to a point \mathbf{p}_i on a two-dimensional plane.

1. The force f_i on each point \mathbf{p}_i is computed as:

$$f_i = \sum f_{ij}$$

where

$$f_{ij} = (\|\mathbf{p}_i - \mathbf{p}_j\| - d(X_i, X_j))(\mathbf{p}_j - \mathbf{p}_i) / \|\mathbf{p}_j - \mathbf{p}_i\|$$

2. Each point \mathbf{p}_i is moved according to the force f_i :

$$\mathbf{p}_i = \mathbf{p}_i + \eta f_i$$

where $\eta = 1/n$.

The energy E of the map can be calculated using the difference of the information distances d and the Euclidean distances l of sensor points in the map.

$$E = \sum_{ij} (d_{ij} - l_{ij}).$$

2.2 Example

Sensory data have been collected from an AIBO robot (Sony AIBO ERS-7, dimensions: 180 (W) x 278 (H) x 319 (D) mm) performing a slow walk while moving its head continuously from side to side. The walk was a straight movement performed in an open space (no obstacle). For this first experiment, we tried to limit the influence of the environment on the behavior. Each leg has 3 degrees of freedom, as well as the head. Infrared distance sensors are mounted on the head and on the main body¹. The recorded sensors were:

¹ The robot has a colour camera mounted above its mouth, electro-static touch sensors, paw sensors, LED lights, all of which are not used in the present experiment but have been exploited in other research conducted with this robot (e.g. [25,26]).

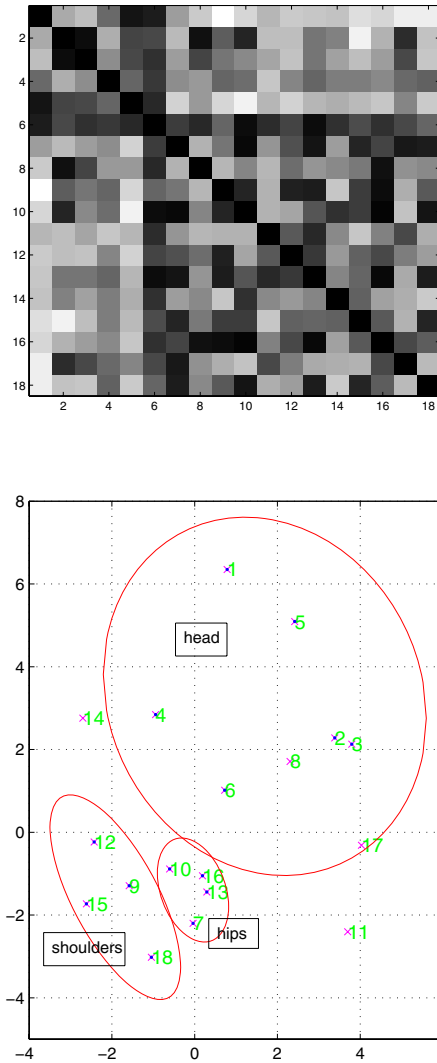


Fig. 1. Information distance matrix and bodymap. The values in the matrices range from zero (dark) to high (light). The mapping from the sensors to the position of the sensors on the robot's body is clearly visible.

- 1-3 distance sensors
- 4-6 head (proprioceptive sensors)
- 7-9 right front leg
- 10-12 right hind leg
- 13-15 left front leg
- 16-18 left hind leg

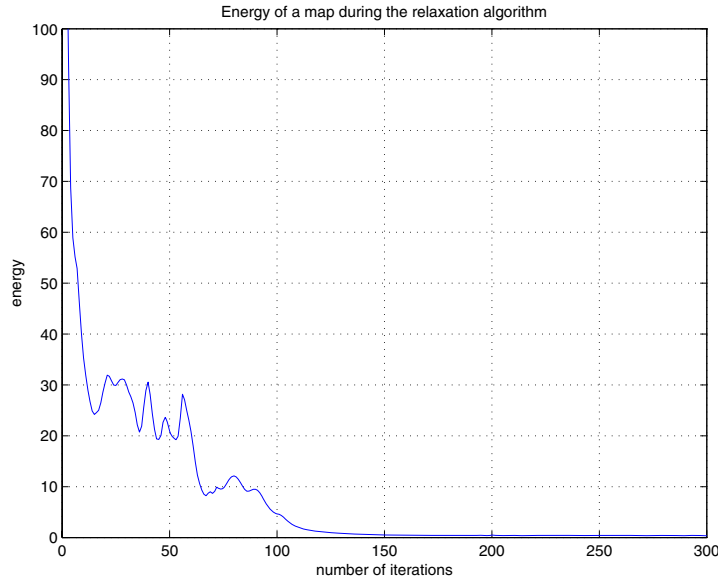


Fig. 2. Energy decrease during the relaxation algorithm

During the walk, 1000 sensor values have been collected for each of these 18 sensors. Figure 1 shows an example of the distance matrix and the maps resulting from the relaxation algorithm using the sensor measurements of the AIBO robot. In this experiment, hip and shoulder joints show remarkable coordination patterns (and also distance sensors and head joints to some extent). Figure 1 bottom shows the two-dimensional map of the robot sensors after applying the relaxation algorithm until the position of the sensor points converged. The decrease in energy of the map can be seen in figure 2. In the map of figure 1, the arrangement of the sensors in the body map already corresponds roughly to the sensor distribution on the body of the robot. Distance and head sensors are arranged in the upper right half of the map, the knee joints of all four legs on the lower right of the map and all other leg sensors on the left side. The exact map depends on the random initial conditions which are different for each run of the relaxation algorithm, but the maps have comparable structures.

The particular emergent organisation of the map results from the body structure of the robot as well as from the behavioural patterns it conducts in a particular environment. In this particular setting, embodiment constraints linking sensor information are probably the most significantly captured (e.g. spatially close similar sensors). In that sense, such maps can be interpreted as a body image. However, for other coordination patterns emergent configurations may differ greatly [4].

3 Interpersonal Maps

In this section, the maps from the previous section that represented both the body and the behaviour of one robot, will be extended to maps including the interaction with

another robot. This is applied to the scenario of an imitation behaviour between robots, in which a strong coupling exists. Different interaction cases are independent behaviour, perfect imitation behaviour, and imitation behaviour with a delay in imitation.

3.1 Definition

The concept of a map can be extended to include not only internal proprioceptive sensors but also external sensors such as visual information. This permits to relate in the same format information about the robot's own body with information about other robots perceived through sensors. Let us define the state of the robot R_Y by a vector of size m :

$$Y(t) = (Y_1(t), Y_2(t), \dots, Y_m(t)) \quad (6)$$

A possible formalisation of this situation can be obtained by supposing that the behaviour of the other robot R_Y is perceived through k new sensors in addition to the ones dedicated to proprioception. The new vector $X(t)$ of size $n + k$ can be expressed as below, where g is a potentially complex function linking the state of R_Y (dimension m) to the perceived state of R_X (dimension k).

$$X(t) = (X_1(t), \dots, X_n(t), g_1(Y(t)), \dots, g_k(Y(t))) \quad (7)$$

In such conditions, a map can be built using the same method as the one described in the previous section. In general, the sensors corresponding to the perceived state of R_Y will not be correlated with the activity of R_X , but they should show separated intracorrelated patterns. In such a case, the body schemas of R_X and R_Y should appear as two distinct clusters in the maps. However in some cases, some intercorrelations could be found between the two sets of sensors. This could be in particular the case when the two robots interact in a closely coupled manner, for instance during a direct imitation task. Such maps can be seen as conceptual signatures for the body correspondence problem. We will now show examples of these two situations.

For the sake of simplicity, we assume in the following examples that g offers a linear mapping linking the sensory states of the observed robot to the states perceived by the observing robot. We will discuss this assumption in the next section.

3.2 Example 1: No Intercorrelation

In this example, we used the sensors recorded from the walking robot together with the sensors of another robot it could have observed. The other robot was sitting and stretching its legs and neck. Altogether, this results in a recording of 36 sensors during 1000 time steps.

Since there is no interaction between the two robots, the two sensor groups are not directly correlated. This results in a smaller information distance on average between two sensors of the same robot than between two sensors of different robots. The interpersonal body map in figure 3 therefore shows two clusters. The first cluster has sensor indices from 1 to 18, the second cluster has sensor indices from 19 to 36. The clusters are indicated with an ellipse each that corresponds to the confidence region of the cluster assuming a Gaussian distribution. Since there are only 18 data points per cluster, this is only a rough approximation and serves the understanding of the graph. The

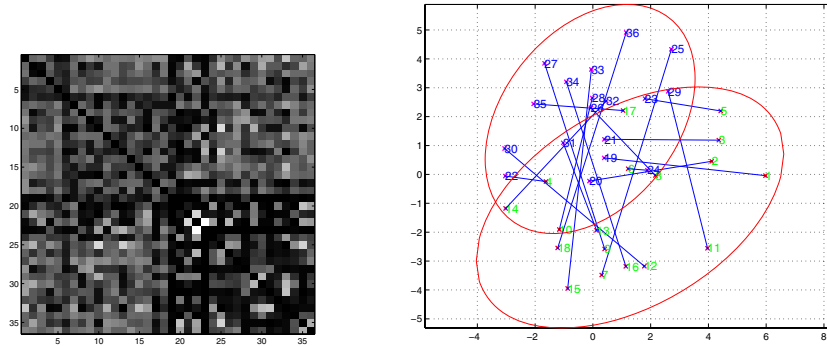


Fig. 3. Information distance matrix and interpersonal body map for a robot observing another robot behaving independently

body schemas within the two clusters are more distorted than the one in figure 1 bottom due to the interplay of the sensors, but a concentration of the head and distance sensors towards the centre of the map is still visible.

3.3 Example 2: Intercorrelation

This example studies the sensory information of one robot imitating the behaviour of the other. In this case, the robots were performing the same programmed walking pattern with a time delay of 10 recordings which corresponds to about half a second (figure 4). In this case, the interpersonal body map does not show two clusters anymore but shows a mapping between sensors of a similar type. Sensors with indices i and $i + 18$ are very close to each other on the body map. In the graphs, they are connected by lines. These lines are much shorter than in the previous graph showing that the information distances between corresponding sensors are small (e.g. X_1 and X_{19}).

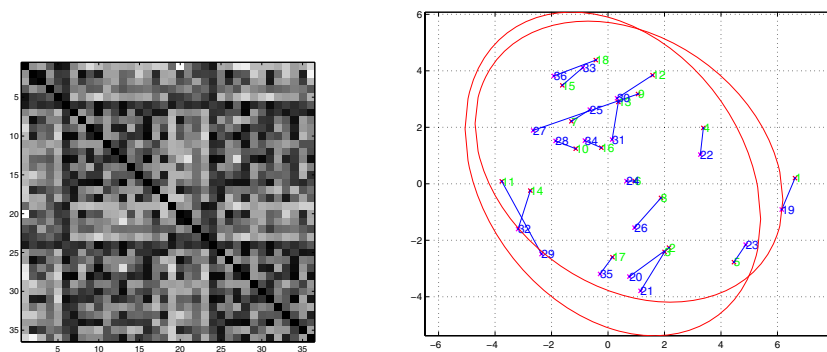


Fig. 4. Information distance matrix and interpersonal body map for a robot being imitated by another robot with a time delay of half a second

3.4 Influence of Delay on the Map Organization

In a further series of experiments, the delay in imitation of the behaviour of one robot by another robot has been varied. The experiments were performed on a robot walking behaviour of 40 seconds imitated with a delay Δ between 0 and 10 seconds. The temporal resolution was again 20 Hz.

To quantify the amount of clustering given the information distances between sensor measurements, we introduced a measurement for clustering, the *clustering factor* c .

$c = (A_1 + A_4)/(A_2 + A_3)$ where A_1, A_2, A_3, A_4 are the sums of distances in the quadrants of the distance matrix \mathbf{D} .

A_1	A_2
A_3	A_4

In the case of a single cluster c should be 1, in the case of two separated clusters c will be smaller than 1.

In figure 5, the clustering factor c over the imitation delay is shown. From a delay of about 1s, a shift occurs and c decreases strongly. It can also be seen that several oscillations occur with a length of about 40 data points (2s). These correspond to the oscillation pattern in the walking behaviour. However, the shift in the clustering factor and therefore in the interpersonal map around the imitation delay of 1s is more prominent than the changes due to the oscillatory behaviour.

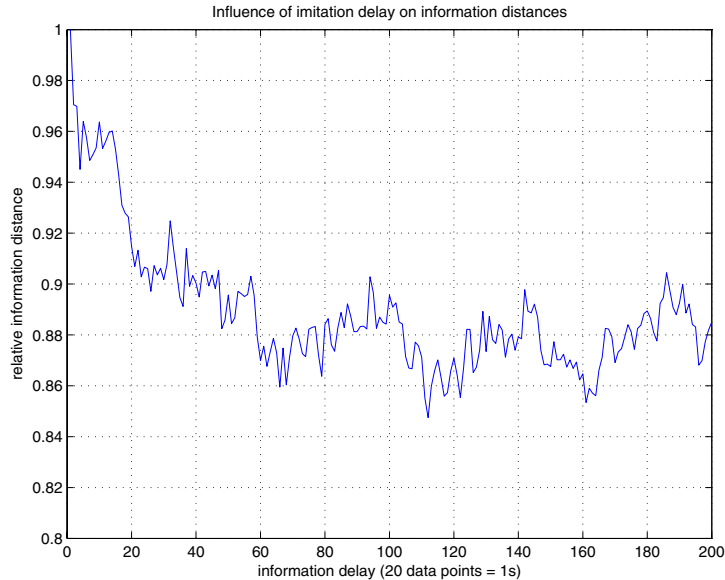


Fig. 5. Effect of the delay in imitation on the clustering of the interpersonal map

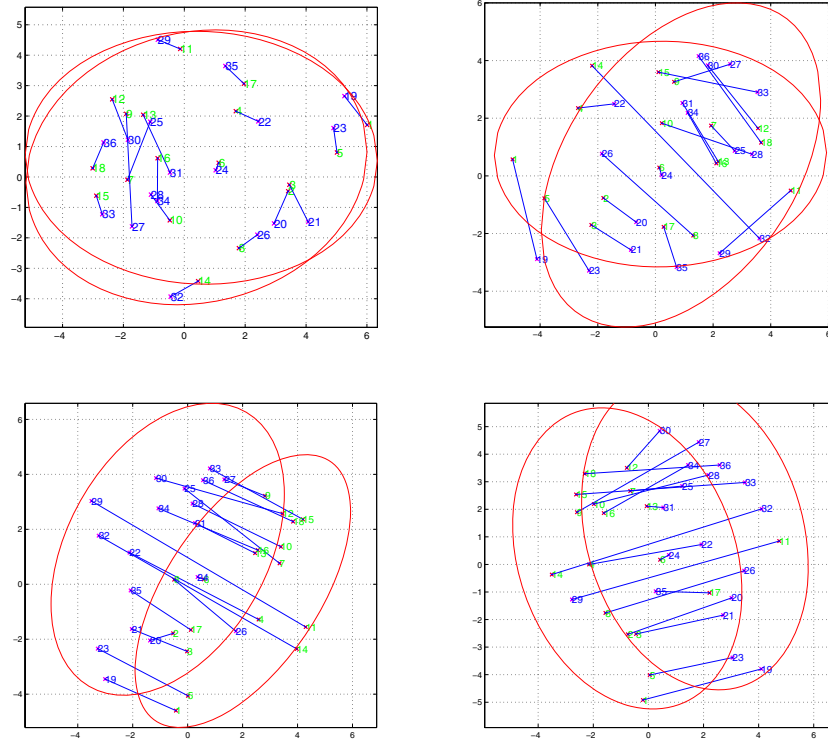


Fig. 6. Interpersonal maps for different delays in imitation: 10 (upper left), 30 (upper right), 60 (bottom left), and 120 (bottom right) data points (20dp=1s)

In figure 6, the interpersonal maps with different delays in imitation are displayed together with the two ellipses indicating the cluster of each robot's sensors as well as lines connecting corresponding sensors between the two robots. The chosen delays in imitation are $0.5s$, $1.5s$, $3s$ and $6s$. The clustering becomes more clear when increasing the imitation delay up to $3s$, but is similar for $3s$ and $6s$. Please note that these are examples of relaxation maps for the given imitation time delays, but they will look different for each run due to the random initialisation parameters.

As the delay increases, the configuration of the interpersonal map progressively shifts. This evolution can be represented by performing principal component analysis (PCA) in the configuration space (the space of the distance matrices) and projecting the data onto its first few principal components. The points corresponding to each configuration for an imitation delay between 0 and 100 data points (corresponding to $5s$) are plotted in figure 7 using the first three principal components. The rapid evolution for the initial high coupling imitation configurations to a cluster corresponding to low coupling situations can be clearly seen. This gives yet another view of this transition for coupled to non-coupled situations.

These different measures and representations support the idea that interpersonal maps can act as signatures of the types of coupling between interacting agents. For

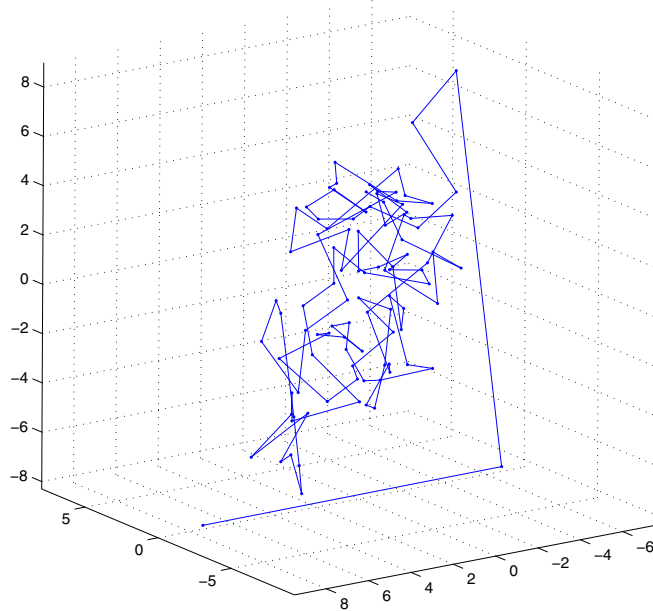


Fig. 7. Projection of the inter-robot information distances onto their first principal components

the present research, we have not implemented a way of classifying these patterns in a fully autonomous manner but we are confident that this is possible although the way to implement such an automatic process would certainly depend on the application considered.

4 Discussion

Our model makes a series of assumptions that need to be discussed. The first one is to separate sensors related to proprioception with sensors related to external perception. In practice, such a clear distinction cannot be obtained. Our embodied perception merges both internal and external stimuli without a priori discrimination. However, presenting the model this way helps clarifying the mechanism we describe.

More importantly, we assume that robot R_X 's perception of the behaviour of robot R_Y can be modelled using a function g mapping the state of R_Y to R_X 's perceptual state. This is a reasonable assumption in the sense that in some way or another the observation of the behaviour of R_Y can be related to R_Y 's internal state. The fact that relevant information about R_Y 's state can be reconstructed after this function has been applied is potentially more questionable. In our context, what counts is that some inter-correlation between Y and X can still be discovered.

We must admit that it is likely that g is a rather complex function. Even in that case intercorrelations could potentially be discovered in several circumstances. One possibility is that R_Y scaffolds the interaction to make its perceived behaviour more

tuned to its own internal state. It has been well studied that adults adapt to children in order to make their overt behaviour more easily analysed [27,28].

Another possibility is that the biases of g are evaluated by a separated mechanism. More generally, the progressive awareness of self and others is likely to be linked with several other developmental processes. Other embodied developmental models suggest for instance that discrimination based on levels of predictability could play a key role in development of the animate/inanimate distinction and the self/other discrimination [29].

Here, we have deliberately chosen not to focus on these important issues in order to investigate first what could be captured by approaches based on information distance matrices. Interpersonal maps may offer a possible unified framework accounting for the structure of the agent's body schema as well as a representation of the observed behaviour of another agent. In cases of strong couplings between agents, a "we-centric" cluster can emerge in which the agent's body structure can be directly mapped onto the structure of an observed body. The system takes different features into account, such as the body, the environment, as well as the agent's behaviour and actions, thus being a perfect architecture for mapping affordances.

We strongly believe that the dynamics responsible for self-other distinction are tightly related with the ones accounting for the construction of the body schema and that both processes must be studied together. It is also clear that information available in such kinds of maps can have a direct influence on the coupling behaviour itself. Our future research will therefore investigate further the consequences of the structuring of this interpersonal space and the possible usage of this type of relational information in the larger context of robotic control architecture. In other words, we must now "close the loop" and show how such interpersonal maps can be used to structure interactions in return. The present research has explored some forms of bottom-up building of information maps, computation performed using such maps should now result in top-down influences.

Acknowledgements

Research funded by Sony CSL Paris with additional support from the ECAGENTS project founded by the Future and Emerging Technologies programme (IST-FET) of the European Community under EU R&D contract IST-2003-1940.

References

1. Gibson, J.J.: The theory of affordances. In: Shaw, R., Bransford, J. (eds.) *Perceiving, Acting, and Knowing*, Lawrence Erlbaum Associates, Hillsdale, NJ (1977)
2. Kaplan, F., Hafner, V.V.: The challenges of joint attention. *Interaction Studies* 7(2), 129–134 (2006)
3. Gibson, J.J.: *The Ecological Approach to Visual Perception*. Houghton-Mifflin, Boston (1979)
4. Kaplan, F., Hafner, V.V.: Information-theoretic framework for unsupervised activity classification. *Advanced Robotics* 20(10), 1087–1103 (2006)
5. Shannon, C., Weaver, W.: *The mathematical theory of communication*. University of Illinois Press (1962)

6. Cover, T., Thomas, J.: Elements of information theory. John Wiley and Sons, Inc., Chichester (1991)
7. Kullback, S.: Information theory and statistics. Dover, Mineola, NY (1968)
8. Tononi, G., Edelman, G., Sporns, O.: Complexity and coherency: integrating information in the brain. *Trends in cognitive sciences* 2(12), 474–484 (1998)
9. Sporns, O., Pegors, T.: Information-theoretical aspects of embodied artificial intelligence. In: Iida, F., Pfeifer, R., Steels, L., Kuniyoshi, Y. (eds.) *Embodied Artificial Intelligence*. LNCS (LNAI), vol. 3139, pp. 74–85. Springer, Heidelberg (2004)
10. Crutchfield, J.P.: Information and its metric. In: Lam, L., Morris, H.C. (eds.) *Nonlinear Structures in Physical Systems – Pattern Formation, Chaos, and Waves*, pp. 119–130. Springer, Heidelberg (1990)
11. Meltzoff, A., Gopnick, A.: The role of imitation in understanding persons and developing a theory of mind. In: Baron-Cohen, S., Tager-Flusberg, H., Cohen, D. (eds.) *Understanding other minds*, pp. 335–366. Oxford University Press, Oxford (1993)
12. Moore, C., Corkum, V.: Social understanding at the end of the first year of life. *Developmental Review* 14, 349–372 (1994)
13. Gallese, V.: The manifold nature of interpersonal relations: the quest for a common mechanism. In: Frith, C., Wolpert, D. (eds.) *The Neuroscience of Social Interaction*, pp. 159–182. Oxford University Press, Oxford, UK (2004)
14. Pierce, D., Kuipers, B.: Map learning with uninterpreted sensors and effectors. *Artificial Intelligence* 92, 169–229 (1997)
15. Kuipers, B., Besson, P., Modayil, J., Provost, J.: Bootstrap learning of foundational representations. *Connection Science* 18(2) (2006)
16. Olsson, L., Nehaniv, C., Polani, D.: From unknown sensors and actuators to actions grounded in sensorimotor perceptions. *Connection Science* 18(2) (2006)
17. Kuniyoshi, Y., Yorozu, Y., Ohmura, Y., Terada, K., Otani, T., Nagakubo, A., Yamamoto, T.: From humanoid embodiment to theory of mind. In: *Embodied Artificial Intelligence*, pp. 202–218. Springer, Heidelberg (2004)
18. Schreiber, T.: Measuring information transfer. *Physical Review Letters* 85(2), 461–464 (2000)
19. Olsson, L., Nehaniv, C., Polani, D.: Sensor adaptation and development in robots by entropy. In: *CIRA 2005. Proceedings of the 6th IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Espoo, Finland, IEEE Computer Society Press, Los Alamitos (2005)
20. Amari, S., Nagaoka, H.: *Methods of information geometry*, vol. 191. Oxford University Press, Oxford, UK (2000) *Translations of mathematical monographs*
21. Lafferty, J., Lebanon, G.: Diffusion kernels on statistical manifolds. *Journal of Machine Learning Research* 6, 129–163 (2005)
22. Hafner, V.V.: Cognitive maps for navigation in open environments. In: *IAS-6. Proceedings of the 6th International Conference on Intelligent Autonomous Systems*, Venice, Italy, pp. 801–808 (2000)
23. Duckett, T., Marsland, S., Shapiro, J.: Fast, on-line learning of globally consistent maps. *Autonomous Robots* 12, 297–300 (2002)
24. Pierce, D.M.: *Map Learning with Uninterpreted Sensors and Effectors*. PhD thesis, The University of Texas at Austin (1995)
25. Steels, L., Kaplan, F.: Aibo’s first words: The social learning of language and meaning. *Evolution of Communication* 4(1), 3–32 (2000)
26. Hafner, V.V., Kaplan, F.: Learning to interpret pointing gestures: experiments with four-legged autonomous robots. In: Wermter, S., Palm, G., Elshaw, M. (eds.) *Biomimetic Neural Learning for Intelligent Robots*. LNCS (LNAI), vol. 3575, pp. 225–234. Springer, Heidelberg (2005)

27. Schaffer, H.: Early interactive development in studies of mother-infant interaction. In: Proceedings of Loch Lomonds Symposium, pp. 3–18. Academic Press, New York (1977)
28. Kaye, K.: The mental and social life of babies. University of Chicago Press, Chicago (1982)
29. Kaplan, F., Oudeyer, P.Y.: The progress-drive hypothesis: an interpretation of early imitation. In: Dautenhahn, K., Nehaniv, C. (eds.) Models and Mechanisms of Imitation and Social Learning: Behavioural, Social and Communication Dimensions, Cambridge University Press (to appear, 2007)