Interpersonal Maps and the Body Correspondence Problem

Verena V. Hafner and Frédéric Kaplan
Sony Computer Science Laboratory, 6 rue Amyot, Paris, France
hafner@csl.sony.fr, kaplan@csl.sony.fr
www.csl.sony.fr

Abstract

In this paper, we introduce the concept of "interpersonal maps". They realize a representation of one's own body to include the body of one's peers. In cases of strong couplings between agents, a "we-centric" space can emerge in which the agent's body structure can be directly mapped onto the structure of an observed body. Based on a set of robotic experiments, we argue that this unified representation can help to elucidate both the formation of a body schema and the body correspondence problem.

1 Introduction

The establishment of the self-other identity is a crucial milestone towards the development of more sophisticated forms of social interaction. It serves as a basis for developing intentional understanding, joint attention and imitative capabilities. Matching and discriminating between oneself and others results certainly from the interplay of several developmental dynamics. In this paper, we focus on a subset of this complex issue by considering the links between the formation of the body schema and the body correspondence problem (Nehaniv and Dautenhahn (2002)). We introduce the concept of "interpersonal maps", realizing a representation of one's own body as well as the body of peers.

This idea is related to several existing concepts. 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 (Meltzoff and Gopnick (1993); Moore and Corkum (1994)). Such a model suggests that both perceived and observed 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 (Gallese (2004)).

However, few models try to give a precise account on how such interpersonal or intermodal mappings could be developed. We believe that research in developmental robotics can play a relevant role to progress in understanding the development of such mappings. Designing algorithms addressing the body correspondence problem and the constitution of the body schema is one of the major challenges of this domain (Kaplan and Hafner (2004)). These issues have been investigated in separate manners (e.g. Yoshikawa et al. (2002, 2004) for the body scheme and Nehaniv and Dautenhahn (2002) for approaches of the correspondence problem). Our model results in a preliminary investigation in trying to address both problems in a unified framework.

2 Maps Based on Information Distances

2.1 Definition

Our approach takes inspiration from research carried out by Olsson et al. (2004) concerning the use of information distances between sensors. This research shows that maps can be built as metric projections showing informational relationships between sensors. It is based on the methods by Pierce and Kuipers (1997) on map learning. In such maps, sensors that are informationally related are close to each other. A related approach was investigated by Kuniyoshi et al. (2004). They argued that such information maps could appropriately be related to "somatosensory maps" such as the ones known to exist in the cortex (Penfield and Rasmussen (1950)).

Such a map can be built in the following way:

Computation of the information distance matrix

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 $X(t)$

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

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 histogram of X_i has to be calculated with a careful choice of the number of bins (see Schreiber (2000)).

The conditional entropy for two sensors X_i and X_j can be calculate 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 $p(x_j|x_i) = p(x_j, x_i)/p(x_i)$.

We chose to use

$$d(X_j, X_i) = H(X_i|X_j) + H(X_j|X_i)$$

as the distance used in the distance matrix since it has several advantages compared to the mutual information (Crutchfield (1990)). d is a metric for the space of information sources. This means that it has the three properties of symmetry, equivalence and triangle inequality.

- $d(X, Y) = d(Y, X)$ follows directly from the symmetry of the definition
- $d(X, Y) = 0$ if and only if X and Y are recoding-equivalent (in the sense defined by Crutchfield Crutchfield (1990)).
- $d(X, Z) \leq d(X, Y) + d(Y, Z)$

Two-dimensional metric projection

A two-dimensional projection is ideal for visualisation of the data. In order to create a two-dimensional body map from the sensor data, we 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[1] to the distance in the n-dimensional information space.

[1] a perfect mapping given the $n \times n$ information distance matrix is possible in an $(n-1)$-dimensional space.

Different algorithms have been suggested (Hafner (2000); Duckett et al. (2002); Pierce (1995)) which convert an n-dimensional input into an m-dimensional map ($m < n$). Here, the algorithm of Pierce (1995) is used since it does not require any information about the relative orientation of connections between sensor nodes.

The algorithm used in this paper consists of an iteration of two simple steps:

First, 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 advantage of using the relaxation algorithm is that it only requires the distances, and not the actual positions, which are not available in our case. A Kohonen self-organising map would therefore not be applicable on this data (Kohonen (2001)).

2.2 Example

Sensory data have been collected from an AIBO robot performing a slow walk while moving its head continuously from side to side. The recorded sensors are:

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

During the walk, 1000 sensor values have been collected for each of these 18 sensors. Figure 1 shows an example of the development of the distance matrices and the maps using the sensor measurements

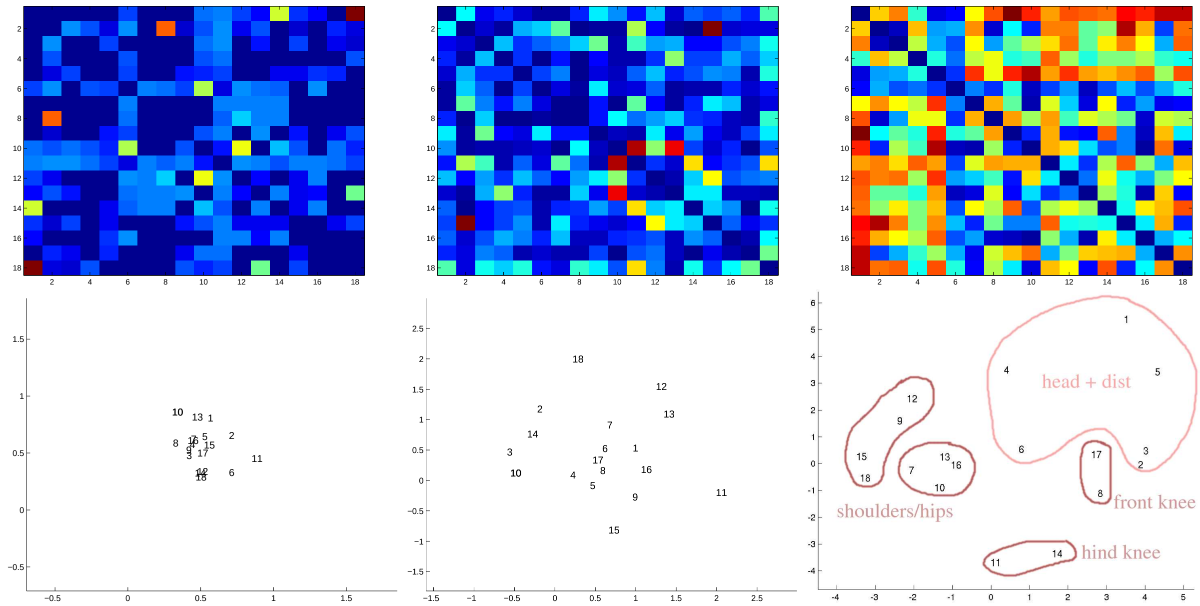

Figure 1: Development of distance matrices and corresponding body maps over time. Left: 10 measurements, centre: 100 measurements, right: 1000 measurements. The values in the matrices range from zero (dark blue) to high (red). In the body map on the right, the mapping from the sensors to the position of the sensors on the robot's body is already clearly visible.

of the AIBO robot after $10, 100$ and 1000 steps. The 18×18 information distance matrix D is symmetrical with zeros in the diagonal, since $d(X_i, X_i) = 0$ and $d(X_i, X_j) = d(X_j, X_i)$.

In the map of figure 1 right, 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 that sense, such maps can be interpreted as a body image.

3 Interpersonal Maps

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:

$$Y(t) = (Y_1(t), Y_2(t), \ldots, Y_m(t)) \tag{2}$$

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).

$$X(t) = (X_1(t), \ldots, X_n(t), g_1(Y(t)), \ldots, g_k(Y(t))) \tag{3}$$

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 intra-correlated 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 higher information distance on average between two sensors of the same robot than between two sensors of different robots. The interpersonal body map in figure 2 therefore shows two clusters. The first cluster can be seen on the lower part of the body map with sensor indices from 1 to 18 printed in black, the second cluster can be seen above the first one with sensor indices from 19 to 36 printed in red. The body schemas within the two clusters are more distorted than the one in figure 1 right 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 walking. The experiment has been performed with imitation with a time delay of 10 recordings which corresponds to about half a second (figure 3). 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 and are plotted in the same colour (e.g. X_1 and X_{19} on the upper right side).

4 Discussion

Our model makes a series of assumptions that can 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 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 its 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 intercorrelation between Y and X can still be discovered. For instance if g is a linear transformation, such kind of information will be entirely conserved.

But it is likely that g is a much more 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 (Schaffer (1977); Kaye (1982)).

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 (Kaplan and Oudeyer (2005)).

5 Conclusion

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" space can emerge in which the agent's body structure can be directly mapped onto the structure of an observed body. We strongly believe that the dynamics responsible for self-other distinction are tightly related with

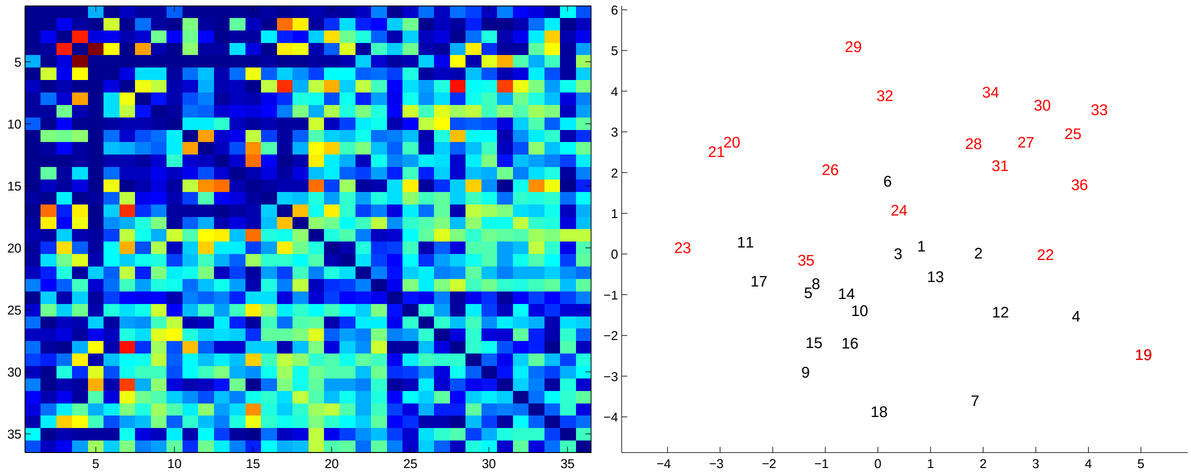

Figure 2: Information distance matrix and interpersonal body map for a robot observing another robot behaving independently.

the ones accounting for the construction of the body schema and that both processes must be studied together. Our future research in developmental robotics will investigate further the conditions for the emergence of this interpersonal space and the possible usage of this information representation in the larger context of robotic control architecture. We also wish to address more precisely the relevance of this mechanism for the development of the self-other matching and discrimination as observed during children's early development.

6 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

J. P. Crutchfield. Information and its metric. In L. Lam and H. C. Morris, editors, *Nonlinear Structures in Physical Systems – Pattern Formation, Chaos, and Waves*, pages 119–130. Springer Verlag, 1990.

T. Duckett, S. Marsland, and J. Shapiro. Fast, on-line learning of globally consistent maps. *Autonomous Robots*, 12:297–300, 2002.

V. Gallese. The manifold nature of interpersonal relations: the quest for a common mechanism. In C. Frith and D. Wolpert, editors, *The Neuroscience of Social Interaction*, chapter 7, pages 159–182. Oxford University Press, 2004.

V. V. Hafner. Cognitive maps for navigation in open environments. In *Proceedings of the 6th International Conference on Intelligent Autonomous Systems (IAS-6)*, pages 801–808, 2000.

F. Kaplan and V.V. Hafner. The challenges of joint attention. In L. Berthouze, H. Kozima, C. Prince, G. Sandini, G. Stojanov, G. Metta, and C. Balkenius, editors, *Proceedings of the 4th International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic System*, pages 67–74. Lund University Cognitive Studies 117, 2004.

F. Kaplan and P-Y. Oudeyer. The progress-drive hypothesis: an interpretation of early imitation. In K. Dautenhahn and C. Nehaniv, editors, *Models and mechanisms of imitation and social learning: Behavioural, social and communication dimensions*. Cambridge University Press, 2005. to appear.

K. Kaye. *The mental and social life of babies*. University of Chicago Press, Chicago, 1982.

T. Kohonen. *Self-Organizing Maps. Extended edition*. Springer, Berlin, 2001.

Y. Kuniyoshi, Y. Yorozu, Y. Ohmura, K. Terada, T. Otani, A. Nagakubo, and T. Yamamoto. From

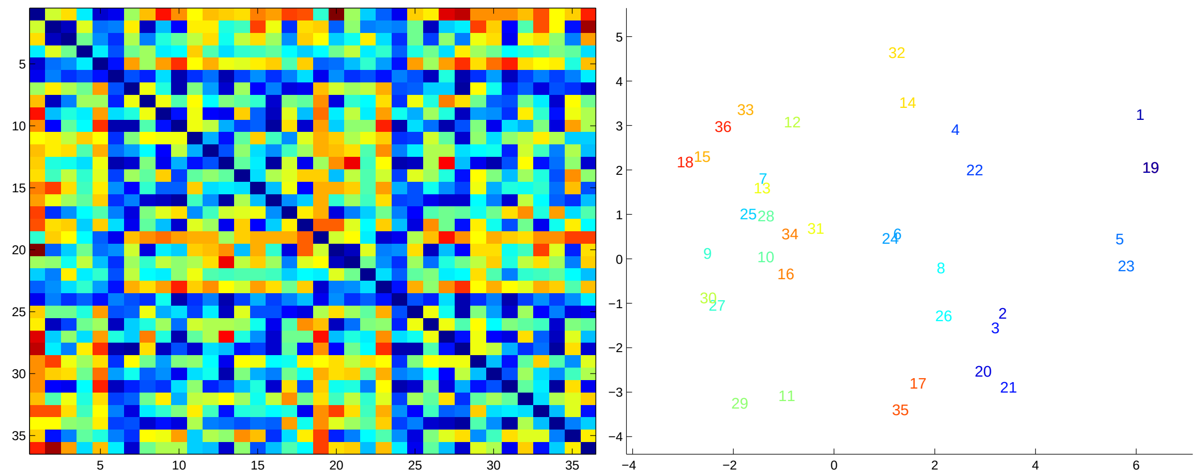

Figure 3: Information distance matrix and interpersonal body map for a robot being imitated by another robot with a time delay of half a second.

humanoid embodiment to theory of mind. In *Embodied Artificial Intelligence*, pages 202–218. Springer Verlag, 2004.

A. Meltzoff and A. Gopnick. The role of imitation in understanding persons and developing a theory of mind. In H. Tager-Flusberg S. Baron-Cohen and D.Cohen, editors, *Understanding other minds*, pages 335–366. Oxford University Press, Oxford, England, 1993.

C. Moore and V. Corkum. Social understanding at the end of the first year of life. *Developmental Review*, 14:349–372, 1994.

C.L. Nehaniv and K. Dautenhahn. The correspondence problem. In K. Dautenhahn and C.L. Nehaniv, editors, *Imitation in animals and artifacts*, pages 41–61. MIT Press, 2002.

L. Olsson, C.L. Nehaniv, and D. Polani. The effects on visual information in a robot in environments with oriented contours. In *Proceedings of the Fourth International Workshop on Epigenetic Robotics*, pages 83–88, 2004.

W. Penfield and T. Rasmussen. *The Cerebral Cortex of Man*. Macmillan, New York, 1950.

D. Pierce and B. Kuipers. Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92:169–227, 1997.

D. M. Pierce. *Map Learning with Uninterpreted Sensors and Effectors*. PhD thesis, The University of Texas at Austin, 1995.

H. Schaffer. Early interactive development in studies of mother-infant interaction. In *Proceedings of Loch Lomonds Symposium*, pages 3–18, New York, 1977. Academic Press.

T. Schreiber. Measuring information transfer. *Physical Review Letters*, 85(2):461–464, 2000.

Y. Yoshikawa, K. Hosoda, and M. Asada. Binding tactile and visual sensations via unique association by cross-anchoring between double-touching and self-occlusion. In *Proceedings of the Fourth International Workshop on Epigenetic Robotics*, pages 135–138, 2004.

Y. Yoshikawa, H. Kawanishi, M. Asada, and K. Hosoda. Body scheme acquisition by cross map learning among tactile, image, and proprioceptive spaces. In *Proceedings of the Second International Workshop on Epigenetic Robotics*, 2002.