In Berthouze, L., Kozima, H., Prince, C. G., Sandini, G., Stojanov, G., Metta, G., and Balkenius, C. (Eds.)
Proceedings of the Fourth International Workshop on Epigenetic Robotics
Lund University Cognitive Studies, 117, ISBN 91-974741-3-4

The Effects on Visual Information in a Robot in Environments with Oriented Contours

Lars Olsson, Chrystopher L. Nehaniv, Daniel Polani
Adaptive Systems Research Group
Department of Computer Science
University of Hertfordshire
College Lane, Hatfield Herts AL10 9AB
United Kingdom
{L.A.Olsson, C.L.Nehaniv, D.Polani}@herts.ac.uk

Abstract

For several decades experiments have been performed where animals have been reared in environments with orientationally restricted contours. The aim has been to nd out what e ects the visual eld has on the development of the visual system in the brain. In this paper we describe similar experiments performed with a robot acting in an environment with only vertical contours and compare the results with the same robot in an ordinary of- ce environment. Using metric projections of the informational distances between sensors it is shown that all visual sensors in the same vertical column are clustered together in the environment with only vertical contours. We also show how the informational structure of the sensors unfold when the robot moves from the environment with oriented contours to a normal environment.

1. Introduction

In nature one nds that most animals are highly adapted to their speci c environment. One example of this is the wide variety of sensory organs that are well adapted to the speci c animals and their respective environment (Dusenbery, 1992). It is believed that in many animals the functionality of the sensory organs is almost completely innate while in other animals the functionality can be altered by experiences during the lifetime of the speci c animal. This age-old question of nature versus nurture has been particularly studied in the visual system (Callaway, 1998). Many experiments have been performed with animals where their visual eld somehow has been restricted to contours of a certain orientation, see for example (Wiesel, 1982) and (Callaway, 1998) for an overview. The results of these experiments are not completely conclusive but some experiments show that animals that have been reared in for example an environment with only vertical contours have more neurons selective for vertical contours. It has also been found that ferrets have more neurons selective for vertical and horizontal contours than other angles (Chapman et al., 1996).

Why is it so that animals have more neurons selective for vertical and horizontal contours than contours of other angles? In a very interesting study by Coppola et al. (Coppola et al., 1998) the distribution of contours of di erent orientations in both man-made environments and a natural forest environment was analysed. That they found more vertical and horizontal contours than oblique angled contours in the man-made environments might not be a big surprise. But, interestingly enough, they also found that natural environments contain more vertical and horizontal contours than contours of other angles. This might explain why visual systems innate have more neurons selective for horizontal and vertical contours.

Why, then, is this important when studying and building robots? In contrast with natural systems, sensors of arti cial systems are often, due to practical and historical reasons, seen as something that is “given” and xed. But, given that robots usually are limited by computational resources as well as power consumption, robots need to use their limited resources e ciently. One way to do this is to try and extract relevant information from the environment as early as possible in the processing steps and then focus the computational resources on these relevant pieces of information. But, how can a robot know what information that is relevant to perform its certain task? One notion of relevant information was introduced and formalized in (Nehaniv, 1999) and extended in (Polani et al., 2001) by associating the relevance of information with the utility to an agent to perform a certain action. Given knowledge of the most relevant information from a number of sensors it might then be possible to adapt the sensoric system to discriminate only between events that are of use for the system. The layout of sensors can also be evolved to a certain environment, which for example has been studied in (Olsson et al., 2004a). It is also possible to use the *sensor reconstruction method* first described in (Pierce and Kuipers, 1997) and extended in (Olsson et al., 2004b) to find sensors that produce the same (redundant) information. If several sensors produce more or less the same information the information from all of them but one can be discarded or some of the sensors can be placed at different positions.

In this paper we perform a similar experiment similar to the ones peformed with for example kittens (Wiesel, 1982) with a robot where the robot is acting in an environment with only vertical contours. Using the sensor reconstruction method we show that in this kind of environment most visual sensors produce redundant information and can be discarded without a loss of information. The results also indicate that a rich visual environment is necessary if it is to learn to distinguish between contours of different orientation. We also show an example of *unfolding* of sensors, where the robot moves from the environment with oriented contours to a normal office environment. In this case the sensors move in the metric projections from the clusters of all sensors from a certain column of visual sensors to a layout that reflects the physical layout of the sensors.

The remainder of this paper is organized as follows. Section 2 contains a brief overview of informational distances between sensors and the sensory reconstruction method. Section 3 contains the results of the experiments with a robot in an environment with oriented contours, and also results where the robot moves from the vertical environment to a normal environment. Finally we summarize the paper and discuss some potential applications of the results and possible future directions of the presented work.

2. Information Distances between Sensors

In order to discuss the information distance between sensors a distance metric is needed. To do this a number of different methods can used, e.g., the Hamming distance and frequency distribution distance (Pierce and Kuipers, 1997). In (Olsson et al., 2004b) these distance metrics are compared with the *information metric*, which was defined and proved to be a metric in (Crutchfield, 1990). The distance between two information sources is there defined in the sense of classical information theory (Shannon, 1948) in terms of conditional entropies. To understand what the information metric means we need some definitions from information theory.

Let $\mathcal{X}$ be the alphabet of values of a discrete random variable (information source, in this case a sensor) X with a probability mass function $p(x)$, where $x \in \mathcal{X}$. Then the entropy, or uncertainty associated with X is

$$H(X) = -\sum_{x \in \mathcal{X}} p(x) \log_2 p(x) \tag{1}$$

and the conditional entropy

$$H(Y|X) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) \log_2 p(y|x) \tag{2}$$

is the uncertainty associated with the discrete random variable Y if we know the value of X. In other words, how much more information do we need to fully predict Y once we know X.

The *mutual information* is the information shared between the two random variables X and Y and is defined as

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \tag{3}$$

To measure the dissimilarity of two information sources Crutchfield's information distance (Crutchfield, 1990) can be used. The *information metric* is the sum of two conditional entropies, or formally

$$d(X, Y) = H(X|Y) + H(Y|X). \tag{4}$$

Note that X and Y in our system are information sources whose $H(Y|X)$ and $H(X|Y)$ are estimated from the time series of two sensors using (2).

It is worth noting that two sensors do not need to be identical to have a distance of 0.0 using the information metric. What an information distance of 0.0 means is that the sensors are completely correlated. As an example, consider two sine-curves where one is the additive inverse of the other. Even though they have different values in almost every point the distance is 0.0 since the value of one is completely predictable from the other. In this case, the mutual information, on the other hand, will be equal to the entropy of either one of the sensors.

In the sensory reconstruction method (Pierce and Kuipers, 1997, Olsson et al., 2004b) metric projections(maps) are created that show the informational relationships between sensors, where sensors that are informationally related are close to each other in the metric projections. To create a metric projection the value for each sensor at each time step is saved, where in this paper each sensor is a specific pixel in an image captured by the robot. A number of frames are captured from the camera of the robot and each frame is one time step. The first step in the method is to compute the distances between each sensor. This is computed by considering the time series of sensor values from a particular sensor as an information source X. The distance between two sensors X and Y is then computed using equation 4. From this 2-dimensional distance matrix a 2-dimensional metric projection can be created using a number of different methods like metric-scaling (Krzanowski, 1988), Sammon mapping, and elastic nets (Goodhill et al., 1995), which positions the sensors in the two dimensions of the metric projection. In our experiments we have used the relaxation algorithm described in (Pierce and Kuipers, 1997).

3. Experiments and Results

In our experiments we have used a SONY AIBO[1] robot dog. The robot walked around more or less at random in an ordinary office environment and one other environment where the visual field consists of black vertical lines on a white background. This environment was created with big white sheets of paper with 2cm wide stripes of black paper glued to the white paper, see Figure 1. Since the AIBO did not move its head up and down and the environment was quite small, most of the collected frames of the visual field consisted only of the striped walls but in some frames part of the uniform floor was also visible. Figure 2 shows an example frame captured by the AIBO in the environment with vertical lines.

Figure 1: The SONY AIBO robot in its environment with vertical contours.

To collect data we used the wireless network of the AIBO to download images from the camera of the robot where the image is 88 pixels wide and 72 pixels high. The frame rate was on average 10 frames per second with a minimum of 9.7 frames per second and a maximum of 10.2 frames per second, where the maximum and minimum values were computed

[1]AIBO is a registered trademark of SONY Corporation.

Figure 2: The environment from the robot's perspective.

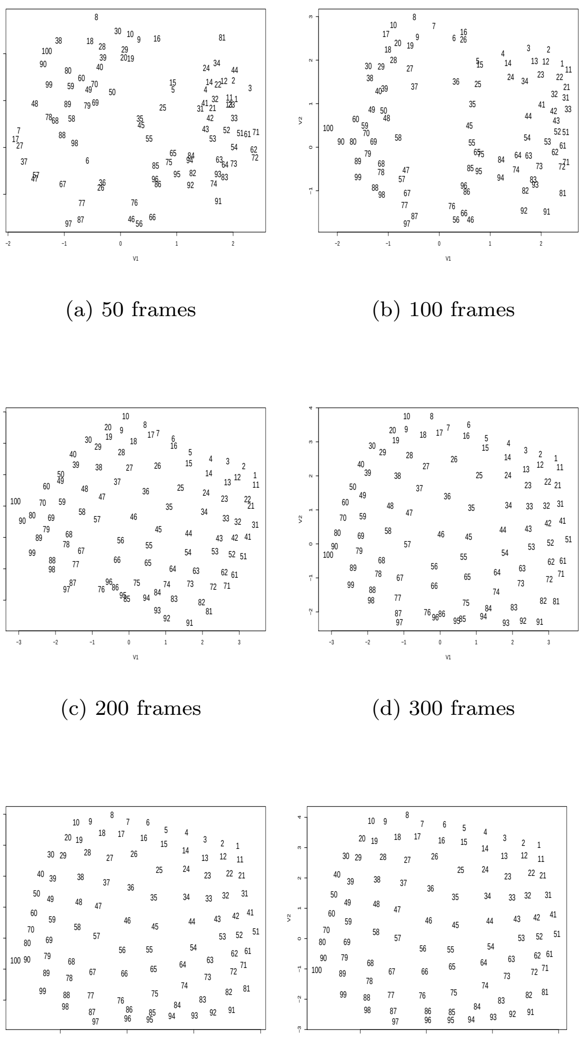
as averages over five seconds. To make the results of the sensor reconstruction method easier to interpret we used a 10 by 10 pixel image from each frame taken from the upper left corner of the image. The pixels of this 10 by 10 image are numbered from 1 to 100, see Figure 3. To verify that the results were not due to the fact that we only used a part of the visual field we performed experiments with the whole image with similar results.

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Figure 3: The layout of the visual sensors (individual pixels).

First consider when the robot is walking in the ordinary office. Contours of all possible angles are visible even though there are probably more vertical and horizontal contours as found in (Coppola et al., 1998). Figure 4 shows the metric projection of the 100 sensors after 50, 100, 200, 300, 400, and 500 frames have been captured. After only 50 frames the sensors are spread out over the metric map but it is hard to distinguish some real order. After 100 frames the metric projection of the sensors has more order. As more frames are processed the structure becomes clearer and clearer and after

500 frames the sensors are ordered in more or less a square with sensor 1 in the upper right corner and sensor 100 in the lower left corner. Thus, the layout of the sensors shown after 500 frames is the mirror image of the real layout found in Figure 3. This orientation of the layout as a mirror image of the real sensors in this case is just a coincidence and in fact all eight possible orientations are equally likely. This is because the data contains no directional information and it is therefore impossible to find the correct physical layout without applying higher-level image analysis. Therefore only the relative positions can be computed, see (Olsson et al., 2004b).



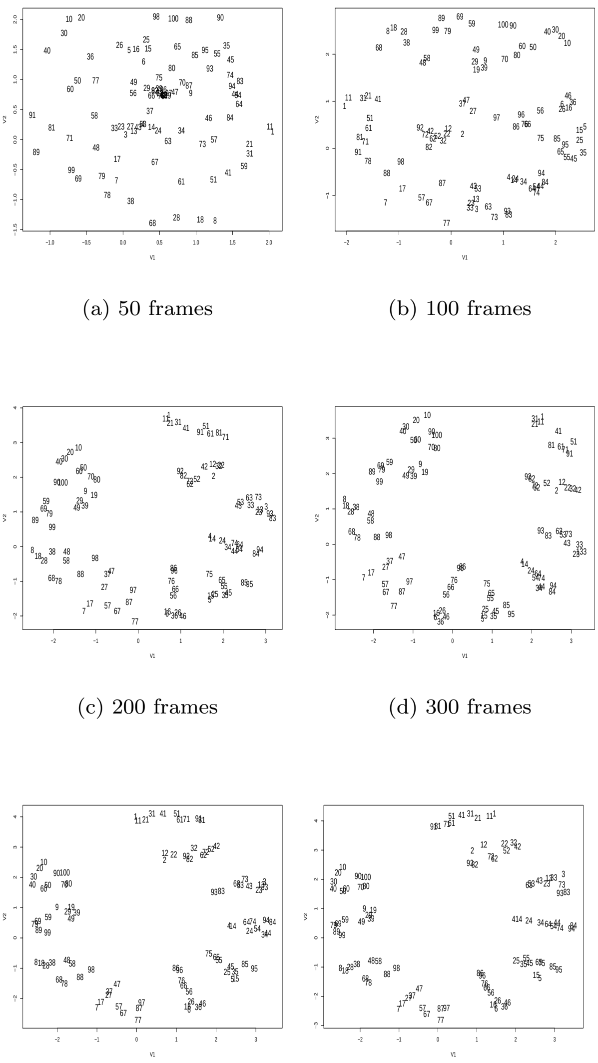
(a) 50 frames (b) 100 frames

(c) 200 frames (d) 300 frames

(e) 400 frames (f) 500 frames

Figure 4: Metric projections of the sensors after 50, 100, 200, 300, 400, and 500 frames of visual data in the normal office environment.

Now consider Figure 5 with metric projections for the visual data from the environment with only vertical contours. After 50 frames no real order can be found. After 100 frames it is possible to see that the sensors start to become clustered in a number of groups. After more frames these clusters become



(a) 50 frames (b) 100 frames

(c) 200 frames (d) 300 frames

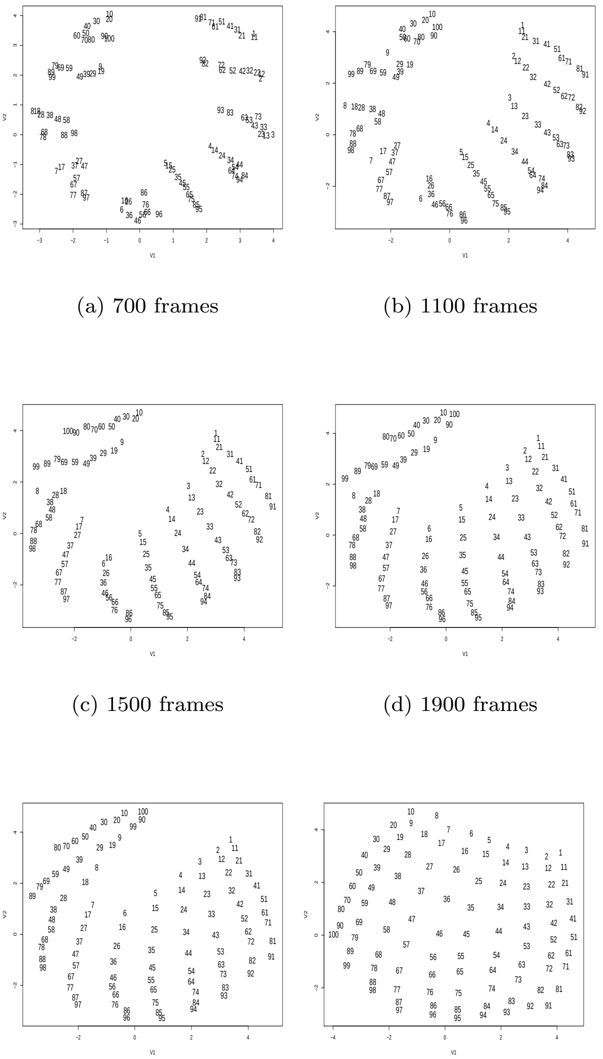
(e) 400 frames (f) 500 frames

Figure 5: Metric projections of the sensors after 50, 100, 200, 300, 400, and 500 frames of visual data in the vertical environment.

more and more noticeable and after 500 frames the sensors are grouped in a horse-shoe shape of 10 clusters with 10 sensors in each. This is very different from the metric projection formed by the frames from the office environment. Looking closer at these clusters we find that each cluster correspond to one column in the sensor layout in Figure 3. For example, the upper right end of the horse-shoe contains the sensors 1,11,21,. . . , 91 which are the sensors of the leftmost column of the layout of visual sensors in Figure 3. If we follow the horse-shoe shape starting from sensors 1,11,21,. . . , 91 we nd that the next cluster is all sensors in the second column (ending with 2) and so forth. Finally the leftmost cluster of the horse-shoe shape contains all sensors of the rightmost column of sensors (10, 20, . . . , 100) in Figure 3.

What is the reason for this clustering of columns of sensors in the vertical environment? In an environment with only vertical contours and a uniform background all sensors with the same horizontal position will at each time step return the same value. Thus the informational distance between these two sensors will be close to or 0.0. In the results presented in Figure 5 we note that the distance between the sensors in the same column is larger than 0.0. This is due to the fact this is data from the real world and hence the light is di erent in di erent parts of the visual eld and not all lines are aligned at exactly the same angle. The shape of the group of clusters is dependent on the distribution of vertical lines within the environment and the width of the lines and the robot's distance to the lines. In this particular experiment the vertical lines were often thinner than the width of the visual eld (10 pixels). Thus we nd that the informational distance between the two outermost columns is shorter than the distance between for example the leftmost column and the middle columns. This is the reason for the horse-shoe shape of the clusters.

Now consider a situation where the AIBO after 600 time steps move from the environment with only vertical contours to a normal o ce environment with contours of all angles. In Figure 6 metric projections are shown of the visual sensors after the robot has moved from the vertical environment to the o ce environment. After 700 time steps the sensors of each vertical column are still clustered together even though they have started to move apart. The longer time that the AIBO has spent in the normal environment the more the metric projection looks like Figure 4 which is the layout of sensors in the normal environment. This is an example of *unfolding* of sensors where the sensors are separated in the metric projection when they distinguish di erent information.



(a) 700 frames (b) 1100 frames

(c) 1500 frames (d) 1900 frames

(e) 2100 frames (f) 2300 frames

Figure 6: Metric projections of the sensors after 700, 1100, 1500, 1900, 2100, and 2300 frames of visual data in the normal environment where the robot moved after 600 frames from the vertical environment.

4. Conclusions

In this paper we have tried to capture some of the properties of experiments done with real animals in environments with oriented contours in experiments using a real robot. A SONY AIBO robot has moved around in an environment with only vertical contours. Metric projections have been created from the visual sensors that show the informational relationships between the sensors. The metric projections show that the sensors from the same column in the visual eld are clustered together which means that the informational distance between them is small, and would in an simulation without noise be 0.0. We also showed how the sensors unfold in the metric projections when the robot moves from the vertical environment to an o ce environment with contours of all angles.

The methods and results in this paper can be applied to robots that adapt and evolve over time, where the robot is constrained by the available number of sensors and the processing power available to process the information from the sensors. First of all is it important to note that all robots interact in some environment and that knowledge about this environment usually can be utilized to optimize the robots' sensors. Consider for example the fact that most mammals have more neurons selective for vertical and horizontal contours than contours of other orientations (Callaway, 1998). This can be explained by the fact that most environments contain more contours of those orientations (Coppola et al., 1998). Thus, the mammal visual system has been adapted by finding generic properties of many different environments. Similarly, is it possible to adapt the visual system of robots by studying the properties of the environment that the robot will interact in. This can for example be done using the sensory reconstruction method (Olsson et al., 2004b) or by evolution of the sensory layouts, see for instance (Olsson et al., 2004a). This was illustrated in this paper where the sensors were clustered in ten groups and most of the sensors could have been discarded or moved to other positions on the robot.

There are several issuses that need to be considered when designing a robot with sensors that can adapt and be optimized to a certain environment as discussed above. First of all there is the obvious question of how to actually design a sensoric system that can adapt to a specific environment. Another important issue is how this specialisation might affect future adaptations to other environments. For example, consider a robot with a sensoric system optimized for the environment with only vertical contours. How can the robot detect that it has moved to a more complex forrest environment using sensors adapted to the restricted vertical environment? This is related to the trade-offs between redundancy and novelty that any designer of a sensoric system is faced with (Olsson et al., 2004a). These are all questions that we intend to investigate in future work.

Acknowledgements

We wish to express our gratitude to the anonymous reviewers for their helpful and insightful comments.

References

Callaway, E. M. (1998). Visual scenes and cortical neurons: What you see is what you get. *Proceedings of the National Academy of Sciences*, 95(7):3344–3345.

Chapman, B., Stryker, M. P., and Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. *Journal of Neuroscience*, 16(20):6443–6453.

Coppola, D., Purves, H. R., McCoy, A. N., and D., P. (1998). The distribution of oriented contours in the real world. In *Proceedings of the National Academy of Sciences*, volume 95, pages 4002–4006.

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

Dusenbery, D. B. (1992). *Sensory Ecology.* WH Friedman & Co, first edition.

Goodhill, G., Finch, S., and Sejnowski, T. (1995). Quantifying neighbourhood preservation in topographic mappings. *Institute for Neural Computation Technical Report Series, No. INC-9505.*

Krzanowski, W. J. (1988). *Principles of Multivariate Analysis: A User's Perspective.* Clarendon Press, Oxford.

Nehaniv, C. L. (1999). Meaning for observers and agents. In *IEEE International Symposium on Intelligent Control / Intelligent Systems and Semiotics, ISIC/ISIS'99*, pages 435–440. IEEE Press.

Olsson, L., Nehaniv, C. L., and Polani, D. (in press, 2004a). Information trade-offs and the evolution of sensory layouts. In *Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9) September 12-15th 2004 Boston, Massachusetts, USA*. MIT Press.

Olsson, L., Nehaniv, C. L., and Polani, D. (in press, 2004b). Sensory channel grouping and structure from uninterpreted sensor data. In *2004 NASA/DoD Conference on Evolvable Hardware June 24-26, 2004 Seattle, Washington, USA*. IEEE Computer Society Press.

Pierce, D. and Kuipers, B. (1997). Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92:169–229.

Polani, D., Martinetz, T., and Kim, J. (2001). An information-theoretic approach for the quantification of relevance. In Kelemen, J. and Sosík, P., (Eds.), *Advances in Artificial Life, 6th European Conference, ECAL 2001, Prague, Czech Republic, September 10-14, 2001, Proceedings*, Lecture Notes in Computer Science, pages 704–713. Springer Verlag.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656.

Wiesel, T. (1982). Postnatal development of the visual cortex and the influence of environment. *Nature*, 299:583–591.