

# Information Self-Structuring: Key Principle for Learning and Development

Max Lungarella

*Department of Mechano-Informatics  
University of Tokyo, 113-8656 Tokyo, Japan*  
maxl@isi.imi.i.u-tokyo.ac.jp

Olaf Sporns

*Department of Psychology  
Indiana University, Bloomington IN 47405, USA*  
osporns@indiana.edu

**Abstract** - Intelligence and intelligence-like processes are characterized by a complex yet balanced interplay across multiple time scales between an agent's brain, body, and environment. Through sensor and motor activity natural organisms and robots are continuously and dynamically coupled to their environments. We argue that such coupling represents a major functional rationale for the ability of embodied agents to actively structure their sensory input and to generate statistical regularities. Such regularities in the multimodal sensory data relayed to the brain are critical for enabling appropriate developmental processes, perceptual categorization, adaptation, and learning. We show how information theoretical measures can be used to quantify statistical structure in sensory and motor channels of a robot capable of saliency-driven, attention-guided behavior. We also discuss the potential importance of such measures for understanding sensorimotor coordination in organisms (in particular, visual attention) and for robot design.

**Index Terms** – embodied interaction, information theory, information structure

## I. INTRODUCTION

Good sensory stimulation may not only be crucial for the proper development of neural structures and internal representations, but may also help bootstrap the emergence of cognitive abilities. Of pivotal importance may be the presence of informational structure in the sensory data. Plausibly, statistical regularities can be exploited by neural circuits to aid the stabilization of matching neural connections incorporating recurrent statistical features. It is therefore important to investigate the origin of such regularities. Are they exclusively the result of “pre-processing” performed by the sensory apparatus? Are they the outcome of neural information processing carried out by the brain? Or are other mechanisms involved?

We argue that informational structure in the sensory experience of (natural and artificial) embodied systems is at least in part the result of effectively coordinated motor activity. Such self-generated motor activity may have a powerful influence in shaping informational structure and ensuring high quality sensory information. An immediate implication of this view is that embodiment plays an important role as a necessary complement to neural information processing: the agent's control architecture (e.g. nervous system) attends to and processes streams of sensory stimulation, and ultimately generates sequences of motor actions which in turn guide the further production and

selection of sensory information. In other words, “information structuring” by motor activity and “information processing” by the neural system are continuously linked to each other through sensorimotor loops.

The importance of statistical structure for perception and cognition has not gone unnoticed. It has been argued, for instance, that the statistical regularities of the environment are highly relevant for an information processing brain and that their exploitation can promote survival [1]. The core idea is that because not all sensory signals are equally likely, sensory systems are not only optimized for processing and “efficiently coding” frequently occurring signals, but they are also particularly well adapted to the statistical structure of the natural environment. This idea has been thoroughly explored in attempts to relate the properties of cortical neurons to the amplitude distribution of spatial frequencies in natural images [2], to explain the shape of neuronal fields in the early stages of the auditory systems of frogs [3], and to better understand the form of the receptive fields of V1 neurons [4]. The research effort has been mainly geared towards the analysis of visual data obtained from static visual images: intensity statistics, color statistics, spatial correlations, space-time statistics, sparse coding [5], independent components analysis [6], wavelet analysis [7], and analyses in the frequency domain, such as power spectral analysis [8].

In comparison, research applying statistical and informational analyses to data streams sampled directly by organisms or robots is scarce. Notable in this respect is recent work by Betsch et al. [9] who have used second-order statistics and wavelets to analyze videos of natural scenes recorded by a camera attached to a cat's head. Their results show an enhanced occurrence of horizontal orientations compared to average orientations, and faster changes in natural stimuli of contour positions compared to orientations. Information-theoretical measures have also been used to quantify data streams obtained from robotic systems [10-14]. These studies demonstrate quite clearly that different sensorimotor strategies can lead to different amounts of statistical structure in visual input streams. For example, visual tracking can generate high mutual information and complexity in visual inputs.

The main question addressed in this paper is whether and to what extent the sensory information generated by and structured through coordinated motor activity can be quantified. To get a better grasp on the mechanisms underlying such structuring of sensory information, we will

employ a set of information theoretic measures to analyze sensorimotor data sampled by a robotic system capable of saliency-driven attention-guided movements. The tools presented here may aid in a further investigation of the “self-structuring of sensory inputs through embodied action” as a key principle for learning and development in robots and natural organisms. Given the importance of input statistics for neural function, a quantitative characterization of the relationship between environmental regularities and neural processing could also help to provide novel experimental tools as well as a theoretical framework for understanding the functional properties of neurons and neural systems.

## II. METHODS

In the context of this paper, we will refer to “structure” as statistical dependencies or relationships between sensing elements (receptors) which are measured using quantitative approaches from information theory. This section introduces and briefly discusses the set of statistical and information theoretical tools employed to analyze data generated by our robotic system. The measures may be applied at various levels (e.g. sensory, neural, behavioral), and at multiple time scales (e.g. learning, developmental, and evolutionary).

### A. Entropy and mutual information

One of our goals is to quantify the amount of informational structure present in sensorimotor data sets. An obvious candidate method is entropy. Entropy is a general information theoretic functional that measures the average uncertainty (or information) of a variable [15]. In its multivariate form, it can also be used to identify nonlinear relationships between multiple variables. Typically, the discrete entropy of a time series is calculated according to the equation:

$$H(X) = -\sum p(x) \log p(x),$$

$p(x)$  being the first order probability density function. Note that for equiprobable events the entropy is maximal, and  $H(X) = \log N$ , where  $N$  is the number of observations. The joint entropy of two variables,  $X$  and  $Y$ , is defined as:

$$H(X, Y) = -\sum \sum p(x, y) \log p(x, y).$$

Here,  $p(x, y)$  is the second order probability density function. It must be noted that the calculated (observed) entropy has a functional dependence on the amount of data and the type of discretization chosen, and differs from the true entropy. In particular, for real-world time series of finite length the entropy is biased and may be affected by a systematic error. We compensated for this bias by adding a corrective term according to Roulston [16].

Another important information theoretical measure we employed is mutual information. Mutual information is a generalized form of correlation, analogous to the linear correlation coefficient, but also sensitive to nonlinear dependencies between variables. In terms of probability density functions, the mutual information of two discrete variables,  $X$  and  $Y$ , can be expressed as:

$$MI(X, Y) = -\sum \sum p(x, y) \log \frac{p(x)p(y)}{p(x, y)}.$$

Intuitively, the mutual information is high if both  $X$  and  $Y$  have high entropy (high variance), and are highly correlated (high covariance); it is zero if  $X$  and  $Y$  are statistically independent, and thus  $p(x, y) = p(x)p(y)$ . In other words, mutual information measures the deviation from statistical dependence of the two variables.

Similar to the difference between observed and true entropy, due to the finite size of the data sets, a difference exists between observed mutual information and true mutual information. This systematic error can be partly corrected by subtracting a term from the observed mutual information [16].

Despite being an important information theoretic measure, mutual information neither contains dynamical nor directional information. A somewhat arbitrary, but nevertheless important improvement is obtained by introducing a time delay between the time series observations. The time delayed mutual information between two time series  $X_t$  and  $Y_t$  is an asymmetric entity, and is defined as  $MI(\tau) = MI(X_t, X_{t-\tau})$ , where  $\tau$  denotes the delay.

### B. Integration and complexity

Two additional information theoretical measures employed were integration and complexity. Both capture global aspects of the statistical dependence between and within sensorimotor data sets, as well as the information distribution among such sets.

Generally speaking, integration is a global estimate of the amount of statistical dependence within a given system or set of elements  $X = \{x_1, x_2, \dots, x_N\}$  provided by the difference between the individual entropies of the elements and the joint entropy of the entire set [17]:

$$I(X) = \sum_i H(x_i) - H(X).$$

Any amount of statistical dependence between the elements will express itself in a reduction of their joint entropy and thus in a positive value for  $I(X)$ . If all elements are statistically independent, their joint entropy is the sum of the individual entropies of the elements, and  $I(X) = 0$ . Thus, integration quantifies the total amount of structure or statistical dependency present in the sensorimotor data, essentially serving as a multivariate generalization of mutual information.

The interplay between segregation and integration leaves characteristic signatures in the patterns of statistical interactions and dependencies. Their combination exhibits “interesting” structure present at different levels of scale – a hallmark of complexity [17,18]. Less complex systems or data sets contain no statistical structure (e.g. ideal gas) or contain structure only at one level which simply repeats (e.g. crystal). To calculate the complexity for a data set, the spectrum of average integration across all levels of scale must be derived. The complexity can be expressed as:

$$C(X) = H(X) - \sum_i H(x_i | X - x_i).$$

Previous work [17,18] has demonstrated that  $C(X)$  is high for systems that effectively combine functional segregation and integration, for instance, by incorporating specialized elements capable of global (system-wide) interactions. On the other hand,  $C(X)$  is low for random systems, or for systems that are highly uniform (or, in other words, systems that lack either global integration or local specialization).

### C. Dimension reduction

The sensory space is defined as the ensemble of all possible configurations that the sensory system can assume, and the motor space is defined as the space of all possible motor states or motor actions. The combinatorial explosion of the number of states in both systems is apparent. The color camera used in our experiments, for instance, has a resolution of 240x320 pixels and a color depth of 24 bit and is thus theoretically capable of displaying something like  $10^{555000}$  different activity patterns. In general, the number of possible states increases dramatically with the number of dimensions (of sensory and motor space). It is thus desirable to reduce the dimensionality of sensory and motor data while maintaining its informational content.

One information reduction technique we employed is Principal Components Analysis (PCA). It is a popular and powerful tool used to represent a multivariate (potentially high-dimensional) data set with a linear combination of a limited number of uncorrelated variables called principal components. The eigenvectors of the covariance matrix of the data identify the directions of the principal components of the original data, with statistical significance given by the eigenvalues of the covariance matrix.

Although classical (linear) techniques for dimension reduction such as PCA and Multidimensional Scaling (MDS) are simple to implement, real-world data often contains nonlinear structure that such methods are not able to resolve. We chose therefore to use a rather novel nonlinear technique combining major algorithmic features of PCA and MDS: the isometric feature mapping algorithm [19]. Isomap is guaranteed – given a sufficient number of data points – to discover the nonlinear degrees of freedom of complex data sets.

## III. EXPERIMENTAL SETUP

Our experimental platform was an active vision system capable of saliency-driven attention-guided head/camera movements. The vision system consisted of a color CCD video camera mounted on a pan-tilt platform facing an LCD video screen displaying “synthetic” or “natural” video input (Fig. 1A). Synthetic video input consisted of a background array of pixels to which color values were randomly assigned, and a single foreground object, a bright red square. This visual stimulus was re-created at a frequency matching the frame processing rate of the visual model, resulting in the visual impression of a “flickering” de-correlated background with a red object floating on top. The object was gradually translated across the array in a random path. “Natural” video input

consisted of short segments of brightly lit outdoor film scenes containing various combinations of animate and inanimate objects, displayed by a standard DVD player/LCD TV combination at normal speed.

Pan and tilt servos were controlled by serial line commands issued from a Matlab simulation of a saliency-driven visual attentional system. Color images sampled by the CCD camera were processed at a frame rate of approximately 8.5 Hz, which was sufficient to allow real-time interactions with human operators, with fast-moving stimuli, or other kinds of rapidly changing visual input.

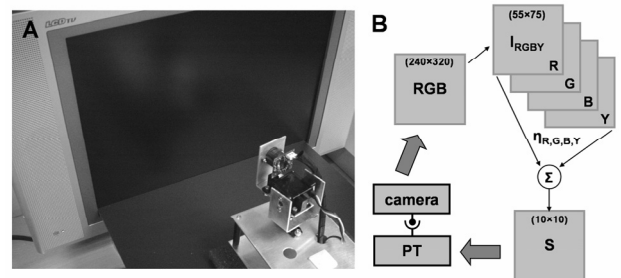


Fig. 1: Experimental setup (A) and model design (B). See text for details.

The purpose of the system was to direct attention – and thus processing resources – to particular locations in space. In doing so, the vision system was able to select stimuli from among several simultaneously present according to their “behavioral relevance”, or saliency. The design of the color system (Fig. 1B) closely followed Itti et al. [20]. The raw visual image was sampled at a resolution of 240×320 pixels, down-sampled to 55×75 pixels, and luminance-scaled according to standard formulae. The scaled components were then used to compute color opponent maps for  $R$ ,  $G$ ,  $B$  and  $Y$  (“yellow”) according to:  $R_{opp} = R - (G+B)/2$ ,  $G_{opp} = G - (R+B)/2$ ,  $B_{opp} = B - (R+G)/2$ ,  $Y_{opp} = (R+G)/2 - B - |R-G|$ . Subsequently, an opponent threshold was applied, followed by a “winner-take-all” mechanism resulting in a color intensity map  $I_{RGBY}(R, G, B, Y)$  that recorded each pixel’s thresholded intensity of the dominant colors  $R$ ,  $G$ ,  $B$  and  $Y$ . From this map a color saliency map was created as  $S_{col} = \eta_R I_{RGBY}(R) + \eta_G I_{RGBY}(G) + \eta_B I_{RGBY}(B) + \eta_Y I_{RGBY}(Y)$ . The scaling factors  $\eta_R$ ,  $\eta_G$ ,  $\eta_B$ , and  $\eta_Y$  quantify the relative saliency of each of the four color components, and were normalized such that their sum was 1. We set  $\eta_R = 1.0$ , which resulted in a strong preference of the active vision system for the color red. A more versatile active vision system would have to take into account additional feature channels selective for oriented contrast boundaries, motion, or disparity.

The color saliency map  $S_{col}$  was then “block-averaged” to derive a map with lower spatial resolution (10×10). The global maximum of the activation of the 10×10 saliency map determined the spatial location to which eye/camera movements were directed. The spatial coordinates of the

maximum were eventually transformed into servo motor commands resulting in immediate fixation of the camera on this location. If the visual scene was stationary, no further camera movements occurred. If the visual scene consisted of moving objects or scenes, then the camera system acquired and tracked whichever object was most salient, that is, produced the highest activation within the saliency map.

#### IV. ANALYSIS

All video segments ran for 120 seconds during which 1000 frames were recorded, and 1000 motor commands were issued by the active vision system. In the “foveation” condition (designated “fov” throughout all figures), the active vision system guided the video camera towards regions containing maximally salient red pixels, “in sync” with current visual input. The result was coordinated visuo-motor activity in which the video camera acquired and tracked regions of red inputs. In the “random” condition (“rnd”), a previously recorded visual scan path was substituted as a motor signal, while the corresponding video segment was played with a time delay of 20 seconds. Our approach effectively decoupled video input and motor output, while maintaining very similar video content and identical movements. It also ensured that any differences in informational measures between “fov” and “rnd” conditions were only due to the presence/absence of coupling between sensory and motor data streams.

The saliency encoding of color allowed us to analyze raw sensory data in the color domain (RGB channels of the CCD camera). Visual data was converted from 8-bit resolution to 5-bit resolution to provide more stable statistical estimates for univariate and multivariate state distributions. (Note that all results reported in this paper are qualitatively identical over a wide range of state spaces, from 8-bit down to 3-bit). Then, cumulative maps for intensity and entropy were calculated for each pixel of the  $55 \times 75$  visual array (Fig. 1B).

For both synthetic and natural video input, we observed an increase of the cumulative intensity of the color red near the center of the visual field (the subsequent analysis is carried out for the color red only). In addition, the entropy of pixels in this region was significantly decreased compared to more peripheral regions (Fig. 2). This decrease in entropy reflected the prevalence of high measurements of “red” near the center relative to less even and hence more entropic distributions in the periphery. For both synthetic and natural visual input, the region of higher intensity and lower entropy in the visual input disappeared in the “random” condition (Fig. 2). We conclude that the measured intensity and entropy effects are the result of sensorimotor coupling between visual scene and vision system.

By taking the first temporal derivative of pixel intensities, we observed a reduction of the entropy levels of central pixels (this reduction was not as pronounced as it was for absolute pixel values). In addition, in the “foveation” condition, the mutual information between adjacent pixels was significantly increased in the central versus peripheral parts of the visual field. No significant differences existed in the “random”

condition. Qualitatively similar results were obtained for synthetic and natural video input (data not shown). Using the first temporal derivative also led to a greater stationarity of the data sets and a more Gaussian state distribution.

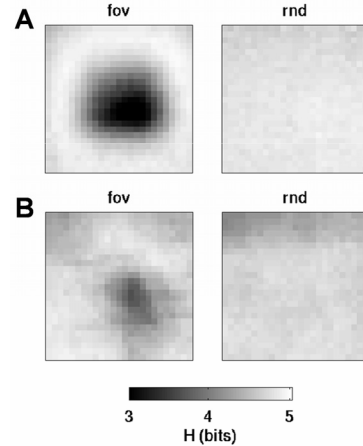


Fig. 2: Entropy. Maps of central  $25 \times 25$  pixel regions obtained from one representative experiment. (A) “Synthetic video”. (B) “Natural video”. Entropy values (mean, S.D) for  $5 \times 5$  pixel patches at the center and in the periphery of the ‘fov’ map are  $3.20 \pm 0.27$  vs.  $4.37 \pm 0.62$  bits (A) and  $3.98 \pm 0.22$  vs.  $4.50 \pm 0.30$  bits (B) (comparisons are significant at  $p < 0.05$ ). Corresponding ‘rnd’ central patches are  $4.82 \pm 0.04$  (A) and  $4.69 \pm 0.03$  (B).

Our hypothesis that coordinated sensorimotor activity can induce changes in informational measures that capture global patterns or statistical interactions is confirmed by Figure 3. The figure demonstrates that for natural video data sampled in the “foveation” condition, significant differences exist between central and peripheral parts of the visual field. Both, integration and complexity were much larger for non-overlapping  $5 \times 5$  pixel patches near the center (Fig. 3A; 3B,C, left panel). “Random” data did not show such significant differences (Fig. 3B,C, right panel). A similar conclusion can be drawn from Figure 4, which displays the time course of integration and complexity calculated using a moving window of 100 time steps. Temporal fluctuations in these informational measures reflect fluctuations in visual input. High values of integration and complexity correspond to frames with good foveation and salient objects. By contrast, poor foveation and absence of salient targets go together with low integration and complexity.

We analyzed the central  $5 \times 5$  region of the visual input for evidence of dimension reduction by comparing the “foveation” and the “random” condition for both synthetic and natural video input. We performed standard PCA and applied the “isomap” algorithm to data samples of central patches of  $5 \times 5$  pixels to determine how well the sampled sensory data can be embedded in a low-dimensional space. Figure 5 displays residual variance as a function of dimensionality, with “foveation” behavior resulting in lower residual variance, i.e. better embedding in low-dimensional space, as compared to “random” behavior. Note that while some of the structure is linear and is detected by standard PCA, additional structure is

nonlinear and only detected by “isomap”. The sharp inflection of the isomap-curve for natural video inputs suggests that the true dimensionality of the visual data gathered with “foveation” behavior is near  $d=2$ .

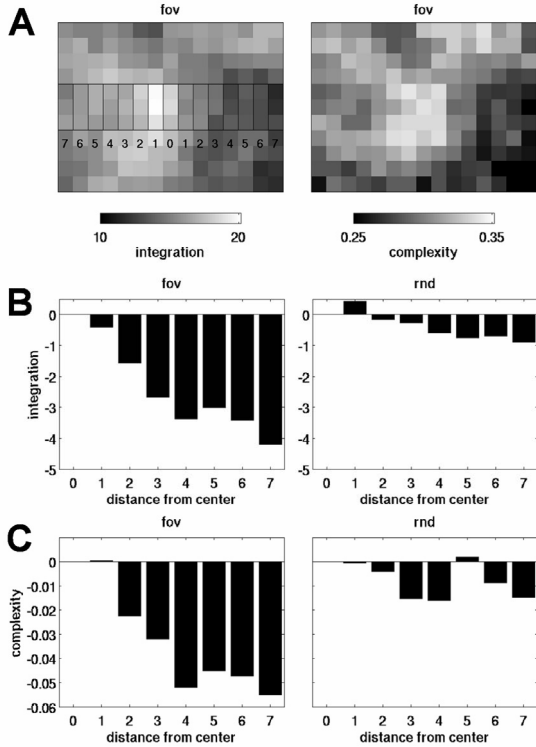


Fig. 3: Integration and complexity for “natural video” data. (A) Integration and correlation map covering the entire visual field, obtained by calculating  $I(X)$  and  $C(X)$  for non-overlapping  $5 \times 5$  pixel patches. Data is averaged over 4 representative experiments. (B) Decrease in integration and (C) decrease in complexity with distance from the center of the visual field, plotted as the difference  $I_{\text{periph}} - I_{\text{center}}$  and  $C_{\text{periph}} - C_{\text{center}}$ . Locations of distance bins are indicated in (A), left panel. Note that in the ‘fov’ condition both integration and complexity are significantly higher in the center versus the periphery, but no significant difference is found in the ‘rnd’ condition.

Our approach can be extended to include neural and sensorimotor data. “Foveation” and “random” conditions produce different distributions in the saliency map  $S_{col}$ , characterized by a significant drop in entropy of “foveation” over “random” data for both synthetic and natural video input (2.2058 bits versus 6.1271 bits for synthetic video and 4.7259 bits versus 6.3007 bits for natural video; data from one representative example each). The entropy of maximal activations represents a measure of the amount of spatial structure in the neural activity pattern of the saliency map  $S_{col}$  generated by sensorimotor coordination.

To quantify the amount of structure in sensorimotor data sets, we analyzed the mutual information between pixels in the central  $5 \times 5$  region of the visual array and the first derivative of the pan-tilt motor signal. In both synthetic and natural video data we found that the mutual information between the sensory and motor variables was higher in “fov”

than in “rnd” conditions ( $0.2089 \pm 0.0103$  versus  $0.1699 \pm 0.0099$  for synthetic video and  $0.3249 \pm 0.0112$  versus  $0.2281 \pm 0.0124$  for natural video input; data from one representative example each). Thus, sensory and motor states are mutually more informative and predictive during coordinated as compared to un-coordinated behavior.

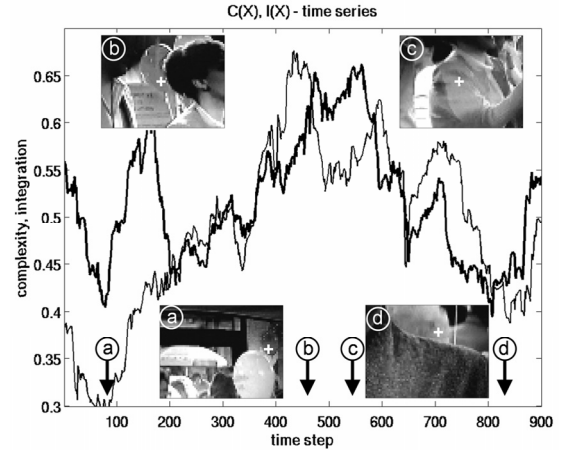


Fig. 4: Time course of integration and complexity.  $I(X)$  and  $C(X)$  are calculated from central  $5 \times 5$  pixels using a moving window of 100 time steps. Note that integration and complexity tend to show correlated increases and decreases over the 120 seconds of the experiment. Sample frames collected when  $I(X)$ ,  $C(X)$  were low (frames a and d) show video frames with poor foveation and lacking salient targets. Frames sampled at high  $I(X)$  and  $C(X)$  show good foveation and salient targets (frames b and c).

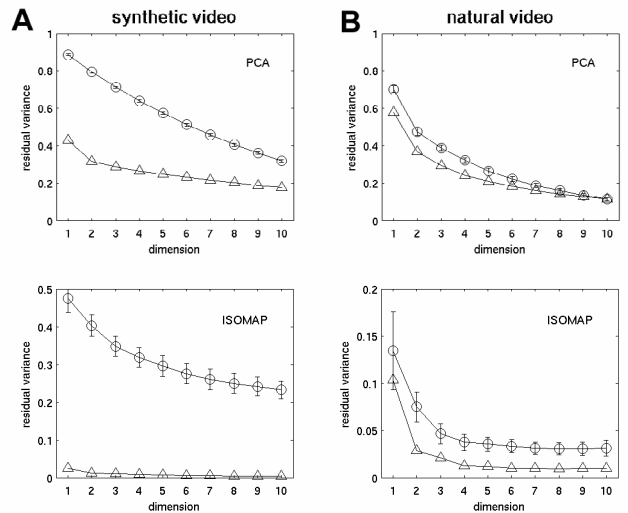


Fig. 5: Dimension reduction in visual data due to “foveation” behavior. “Random” condition: open circles. “Foveation” condition: open triangles.

## V. DISCUSSION

Natural organisms and robots are dynamically coupled to their environments, with sensory activity determining motor outputs, and motor activity selecting and shaping the statistics of the sensory input. In this paper, we argued that embodied agents “self-structure” sensory and motor information by

coordinated and dynamic interaction with their local environment. To test this hypothesis, we introduced a set of quantitative measures aimed at capturing uni- and multivariate statistical regularities present in sensory and motor time series. We exemplified their use by means of a simple robotic application.

Because the foveation behavior of our experimental system was governed by color only, we chose to limit our analysis to the structure of channels carrying color information. This choice also helped us to keep the application simple and straightforward. A more complex robotic system would probably require us to take into account other sensory channels as well, or structural features within the control architecture (neural areas) encoding higher-order sensory features, such as visual orientations, textures, object boundaries, or faces.

Embodied interactions impose statistical structure not only on “raw pixels” within primary sensory channels, but also (and perhaps more powerfully so) on neural activity patterns far removed from the sensory periphery. We predict that embodied systems operating in a highly coordinated manner generate information and additional statistical regularities at all hierarchical levels of their control architectures, including but not limited to the immediate sensory input. Such effects have been demonstrated in the development of visual and neuromodulatory circuits of neurorobotic systems, e.g. [21]. The methods exposed in this paper can be readily extended to address these issues, e.g. by analyzing time series data of V1 or of inferotemporal cortex in behavioral conditions that correspond to our “coordinated” and “un-coordinated” regimes.

How embodied is the system used in the present study? The fact that our active vision system is mechanically non-redundant and immobile relative to a fixed frame of reference (the LCD screen), poses obvious constraints on the complexity of the interaction. We hypothesize that a robotic system interacting more closely with its surrounding environment (e.g. by roaming around, or by manipulating foveated objects) could generate higher levels of correlations and mutual information at or around objects, thus promoting, better recognition and object segmentation.

In this paper, we quantified the contribution of embodied interaction in structuring information for a simple robotic application. Will it be possible to perform this type of analysis in more complex systems (e.g. humanoid robots, or real organisms)? Especially intriguing will be to use informational measures for classifying modes of embodiment according to their impact on information processing, or for designing robots “optimal” with respect to the amount of structure in input and sensorimotor spaces, e.g. by using informational measures as fitness functions in evolutionary algorithms. This study represents only a first step and more methods need to be developed, validated and added to our repertoire of quantitative tools. We hope that these methods will contribute to a quantitative framework guiding future studies of embodied interaction.

## VI. ACKNOWLEDGEMENTS

M.L., who was supported by research grant no. 04413 of the Japanese Society for the Promotion of Science, is also indebted to Y. Kuniyoshi for many fruitful discussions. O.S. was supported by NIH/NIDA grant 1R21DA15647-01.

## VII. REFERENCES

- [1] Barlow, H.B. (2001). The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences*, 24:602-607.
- [2] Field, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4:2379-2394.
- [3] Rieke, F., Bodnar, D.A. and Bialek, W. (1995). Naturalistic stimuli increase the rate and efficiency of information transmission by primary auditory afferents. *Proc. R. Soc. London B*, 262:259-265.
- [4] Webber, C.J. (1991). Competitive learning, natural images, and the self-organization of cortical cells. *Network: Computation in Neural Systems*, 2:169-187.
- [5] Olshausen, B. and Field, D.J. (1996). Emergence of simple-cell receptive-field properties by learning a sparse code for natural images. *Nature*, 381:607-609.
- [6] Bell, A.J. and Sejnowski, T.J. (1997). The independent components of natural scenes are edge filters. *Vision Research*, 37:3327-3338.
- [7] Field, D.J. (1999). Wavelets, vision and the statistics of natural scenes. *Phil. Trans. R. Soc. Lond. A*, 857:2527-2542.
- [8] van der Schaaf, A. and van Hateren, J.H. (1996). Modeling the power spectra of natural images: statistics and information. *Vision Research*, 36(17):2759-2770.
- [9] Betsch, B.Y., Einhauser, W., Kording, K.P. and Konig, P. (2004). The world from a cat’s perspective – statistics of natural videos. *Biological Cybernetics*, 90:41-50.
- [10] Lungarella, M. and Pfeifer, R. (2001). Robots as cognitive tools: Information-theoretic analysis of sensory-motor data. *Proc. of the 1<sup>st</sup> Int. Conf. on Humanoid Robots*, 245-252.
- [11] Lungarella, M., Pegors, T., Bulwinkle, D. and Sporns, O. (2005). Methods for quantifying the informational structure of sensory and motor data. *Neuroinformatics* (in press).
- [12] Sporns, O. and Pegors, T.K. (2003). Generating structure in sensory data through coordinated motor activity. *Proc. of the Int. Conf. on Neural Networks*, 2796.
- [13] Sporns, O. and Pegors, T.K. (2004). Information-theoretic aspects of embodied artificial intelligence. *Embodied Artificial Intelligence*, LNCS 3139. Heidelberg: Springer-Verlag, 74-85.
- [14] Tarapore, D., Lungarella, M. and Gomez, G. (2005). Quantifying patterns of agent-environment interaction. *Robotics and Autonomous Systems* (in press).
- [15] Shannon, C.E. (1948). A mathematical theory of communication. *The Bell Systems Technical Journal*, 27:379-423.
- [16] Roulston, M.S. (1999). Estimating the errors on measured entropy and mutual information. *Physica D*, 125:285-294.
- [17] TONI, G., Sporns, O. and Edelman, G.M. (1994). A measure for brain complexity: Relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA*, 91:5033-5037.
- [18] TONI, G., Edelman, G.M. and Sporns, O. (1998). Complexity and coherency: Integrating information in the brain. *Trends in Cognitive Sciences*, 2:474-484.
- [19] Tenenbaum, J., de Silva, V. and Langford, J. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319-2323.
- [20] Itti, L., Koch, C. and Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(11):1254-1259.
- [21] Alexander, W. and Sporns, O. (2003). An embodied model of reward conditioning. *Adaptive Behavior*, 11(2):143-159.