Berthouze, L., Kaplan, F., Kozima, H., Yano, H., Konczak, J., Metta, G., Nadel, J., Sandini, G., Stojanov, G. and Balkenius, C. (Eds.)
Proceedings of the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems
Lund University Cognitive Studies, 123. ISBN 91-974741-4-2

Segmentation Stability: a Key Component for Joint Attention

Jean-Christophe Baillie* **Matthieu Nottale***
*ENSTA / Cognitive Robotics Theme
Laboratory of Electronics and Computer Engineering,
32, boulevard Victor, 75739 Paris Cedex 15 France

Abstract

It is now well established that joint attention is a key capability for socially interacting robots (Brooks et al., 1999, Kaplan and Hafner, 2004, Scassellati, 1999, Itti, 2003). It is also a key component for epigenetic robotic applications in general. This subject has been widely discussed and we present here one specific technical improvement for joint attention which relies on image segmentation.

In the new Talking Robots (Baillie, 2004) experiment that we have started, following a successful reimplementation of the Sony's Talking Heads (Steels, 1998) experiment on Aibo ERS7, we try to have two robots interacting to evolve a shared repertoire of synchronized behaviors, or "games".

This leads to a dynamic version of the Talking Heads, where the interaction protocol, or language game, is not predefined in the agents. As usually with experiments involving symbol grounding and social interaction, and for these experiments in particular, it is essential that the two robots can establish joint attention and share a common representation of their surrounding environment, while they are looking at the same scene. Many techniques can be used to achieve this stable shared representation. We have chosen to focus on segmentation-based algorithms, which provide interesting shapes together with vectors of features that can be easily extracted and proved useful in the context of our experiment.

However, when using segmentation algorithms, a slight change in the viewpoint, or even the residual camera noise, is enough to significantly change the result of the image partition, possibly leading to completely different perceptions of the same scene.

We have developed an original measure to assess the stability of a segmentation algorithm and we have used it on a set of algorithms to automatically determine the most stable partition for a given scene. This approach is different from classical methods used to estimate the quality of a segmentation algorithm, where the result of the algorithm is compared to an ideal perfect segmentation done by hand. In our approach, the measure is done automatically, involves only stability considerations and could lead to interesting improvements whenever joint attention using segmentation is required.

We quickly present in the poster the background of the Talking Robots experiment and why joint attention and image segmentation stability is an important issue for us. We introduce then two stability measures and show some results on natural scenes from our experiments in the lab and test scenes used to control image parameters. The influence of several image characterizations (noise, number of objects, luminosity,...) is carefully reviewed. An example (fig.1) is given below, showing the influence of noise on typical images.

Using the fact that a given algorithm can be ranked according to its stability score, which is calculated online assuming that the scene itself is static, a general method of algorithm switching is introduced and used in the experiment with different kind of algorithms: region growing, recursive histogram splitting, CSC (Priese and Rehrmann, 1993) and split & merge (CIS). We show how this method significantly improves the convergence speed in the experiment and conclude on the generality of our approach to facilitate certain aspects of joint attention, when it relies on segmentation.

References

Baillie (2004). Grounding symbols in perception with two interacting autonomous robots. *Proceedings of the 4th International Workshop on Epigenetic Robotics*, 117.

Brooks, R. A., Breazeal, C., Marjanovic, M., Scassellati, B., and Williamson, M. M. (1999). The cog project: Building a humanoid robot. *Lecture Notes in Computer Science*, 1562:52–87.

Itti (2003). Visual attention. In Arbib, M. A., (Ed.), *The Handbook of Brain Theory and Neural Networks, 2nd Ed.*, pages 1196–1201. MIT Press.

Kaplan, F. and Hafner, V. (2004). The challenges of joint attention. *In: Proceedings of the 4th International Workshop on Epigenetic Robotics*.

Priese, L. and Rehrmann, V. (1993). A fast hybrid color segmentation method. In *DAGM-Symposium*, pages 297–304.

Scassellati, B. (1999). Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot. *Lecture Notes in Computer Science*, 1562:176–195.

Steels, L. (1998). The origin of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103:133–156.

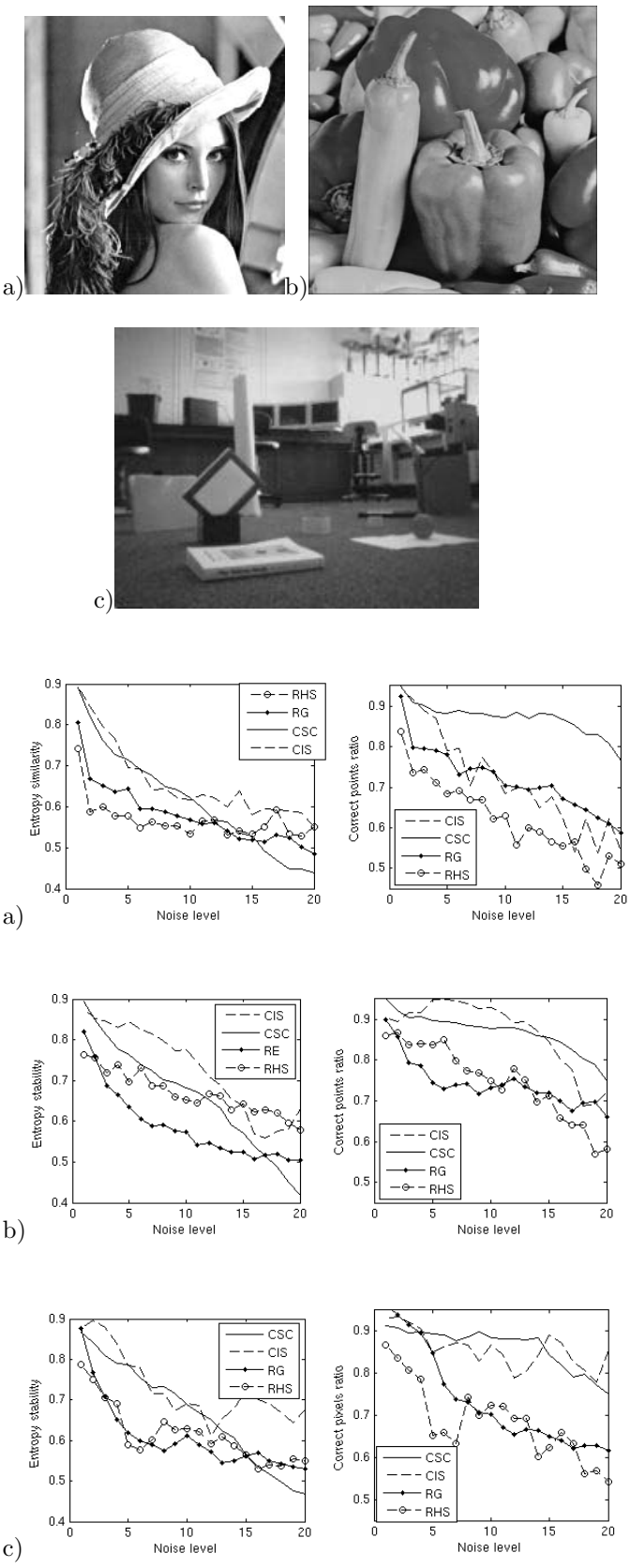

Figure 1: Some stability measure results for image Lena, peppers and an image of our lab, against varying noise level (standard deviation of pixel graylevel). Most stable images have a stability of 1.