

Sensorimotor Models of Space and Object Geometry

Jeremy Stober *

Department of Computer Science
The University of Texas at Austin
stober@cs.utexas.edu

Abstract

A baby experiencing the world for the first time faces a considerable challenging sorting through what William James called the “blooming, buzzing confusion” of the senses. With the increasing capacity of modern sensors and the complexity of modern robot bodies, a robot in an unknown or unfamiliar body faces a similar and equally daunting challenge.

Addressing this challenge directly by designing robot agents capable of resolving the confusion of sensory experience in an autonomous manner would substantially reduce the engineering required to program robots and the improve the robustness of resulting robot capabilities. Working towards a general solution to this problem, this work uses *distinctive state abstractions* and *sensorimotor embedding* to generate basic knowledge of sensor structure, local geometry, and object geometry starting with uninterpreted sensors and effectors.

1 Introduction

Sensorimotor development is the first stage of constructivist theories of child development. During this period infants construct egocentric representations of the world that serve as a conceptual basis for later learning. For developing robot agents the challenge is clear. Starting with what William James termed “the blooming, buzzing confusion” of the senses generate useful egocentric representations of the world. With this in mind consider the following *sensorimotor geometry problem*:

Design a developmental process that, for any roughly humanoid robot, starting only with a basic set of sensor primitives and motor reflexes, progresses through a period of sensorimotor develop-

ment that results in knowledge of body, sensor and object location and geometry.

In robotics, advances in simultaneous localization and mapping (SLAM) have allowed robotic agents to acquire spatial knowledge of the environment along with accurate estimates of a robot’s pose within the environment.

Unfortunately, these methods require that an engineer specify sensor and action models, as well as the target representation of space. The area of developmental robotics includes algorithms that enable agents to acquire foundational knowledge that is normally “engineered in” under more traditional paradigms. A developmental approach to acquiring geometric knowledge of space and objects avoids this need to specify models in advance.

If such an approach exists then robots can understand new sensors or changes to existing sensors and engineers avoid having to provide models. The completed and proposed work presented here focuses on the important sub-problem of learning about geometry of sensors, space, and objects from sensorimotor experience.

2 Related Work

There are a number of existing approaches to solving the sensorimotor geometry problem. Pierce and Kuipers [1997] used dimensionality reduction methods to learn geometry of robot sensors and motors. Philipona and O’Regan [2010] developed a sensorimotor approach using ISOMAP for learning spatial structure from uninterpreted sensors and effectors, and Bowling et al. [2007] developed action respecting embedding (ARE) while working towards a solution to the problem of *subjective localization* (solving SLAM problems without sensor and action models).

These methods all suffer shortcomings. For example, dimensionality reduction alone may not be sufficient for learning spatial and geometric concepts beyond sensor organization as these methods are sensitive to the policy used to control the agent during data collection. Advanced approaches such as ARE are based on maximum-variance unfolding, which has scaling problems on large datasets and may not cope with the firehose of experience. This work presents an alternative general method of extracting geometric information from raw sensory experience through interaction with

*This work has taken place in the Intelligent Robotics Labs at the University of Texas at Austin and at the University of Michigan. Research in the Intelligent Robotics Labs is supported in part by grants from the National Science Foundation (IIS-0713150 to UT Austin and CPS-0931474 to UM) and from the TEMA-Toyota Technical Center to UM.

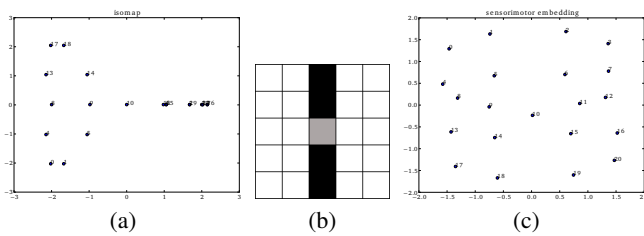


Figure 1: In (b) the agent’s task is to efficiently navigate to the central distinctive state. (a) shows an embedding generated using ISOMAP with local distances given by the magnitude of local actions. (c) shows the improved result of using sensorimotor embedding, which computes distances between states by comparing sequences of actions generated by an optimal policy.

the environment that addresses the drawbacks of existing approaches.

3 Acquiring Geometric Knowledge

For an agent starting with only raw sensory experience and the ability to interact with the environment, acquiring geometric knowledge of the environment is a difficult task. The proposed solution to the problem proceeds in several steps.

The first step is designed to manage the “firehose of experience.” The agent forms a distinctive state abstraction in order to divide sensory experience into manageable chunks [Provost *et al.*, 2006]. This method of abstraction has been used to solve high-diameter reinforcement learning problems and provides a concrete set of policy improvement goals for developing agents. During the second step, an agent learns accurate policies for acquiring and moving between distinctive states. In the final step shown in Figure 1, an agent infers geometric properties of the world using manifold learning methods and dynamic time warping by analyzing action sequences generated by learned policies.

4 Completed Work

Completed work has shown how to derive sensor geometry from raw interactive experience using sensorimotor embedding [Stober *et al.*, 2009]. In these experiments, the agent had access to a foveated (non-uniform) array of sense elements. The agent also had the ability to perform ballistic actions known in human visual processing as saccades. Using sensorimotor embedding agents could learn the structure of a foveated retina by analysing learned saccade policies.

Since retinal geometry is derived from learned saccade policies, adjustments to the saccade policy that governs eye movement, in response to degradation or other changes, also results in adjustments to the agent’s understanding of retinal geometry. In experiments agents using sensorimotor embedding could adapt to even complete inversions of the visual field as well as forms of simulated macular degeneration.

5 Proposed Work

Proposed work will focus on applying the principles of sensorimotor embedding to two new problem domains. The first,

subjective localization, has already been mentioned as part of related work. Since the proposed work and related work share a common platform, direct comparison between sensorimotor embedding, ARE, and other manifold learning approaches will be possible.

The second area of proposed work is applying sensorimotor embedding to the problem of learning visual 3D object models. Recent work in computer vision has focused on generating object models using homographies between distinctive views [Savarese and Fei Fei, 2007]. Instead of homographies between distinctive views, a sensorimotor approach to building 3D object models would use *policies* that allow an agent to bring about distinctive views.

Extending this work to use *interactions* between a robot and an object leads naturally to considering *self-generated views*, a phenomenon that has also been studied in humans. Preliminary results indicate that the process of forming sensorimotor embedding models of objects results in viewpoint biases that mimic the biases found in data collected from humans.

6 Conclusion

Distinctive state abstractions manage the firehose of experience and provide concrete policy improvement goals for agents with uninterpreted sensors and effectors. Using sensorimotor embedding an agent can extract geometric knowledge from sequences of actions generated using learned policies for navigating to and between distinctive states. Since representations of geometry are derived from actions in sensorimotor embedding, the resulting geometric representations are naturally calibrated to the agent’s own body.

References

- [Bowling *et al.*, 2007] M. Bowling, D. Wilkinson, A. Ghods, and A. Milstein. Subjective localization with action respecting embedding. *Robotics Research*, 28:190–202, 2007.
- [Philipona and O’Regan, 2010] D. Philipona and J.K. O’Regan. The sensorimotor approach in CoSy: The example of dimensionality reduction. *Cognitive Systems*, pages 95–130, 2010.
- [Pierce and Kuipers, 1997] D. Pierce and B. Kuipers. Map learning with uninterpreted sensors and effectors. *Artificial Intelligence*, 92(1-2):169–227, 1997.
- [Provost *et al.*, 2006] J. Provost, B. Kuipers, and R. Miikkulainen. Developing navigation behavior through self-organizing distinctive-state abstraction. *Connection Science*, 18(2):159–172, 2006.
- [Savarese and Fei Fei, 2007] S. Savarese and L. Fei Fei. 3D generic object categorization, localization and pose estimation. In *IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007.
- [Stober *et al.*, 2009] J. Stober, L. Fishgold, and B. Kuipers. Learning the sensorimotor structure of the foveated retina. In *Proceedings of the Ninth International Conference on Epigenetic Robotics*, 2009.