Exact Inference in Robots Using Topographical Uncertainty Maps

Josh Susskind, John Hershey, Javier Movellan

Abstract

Many problems in social robotics require a real-time combination of incoming sensor information and prior information about likely behaviors of the objects in the world. For example, tactile, visual and acoustic information may all inform a distribution of beliefs about the location of humans with whom the robot may want to interact. When sensory information is not available, the uncertainty in this distribution should increase in a principled manner to reflect the fact that people are not static objects.

Bayesian filtering provides a principled approach to solve these problems and has thus become a method of choice in robotics. Most of the Bayesian filtering methods applied to robotics rely on analog hypothesis spaces and find approximate solutions to the resulting non-linear filtering problem using Monte-Carlo approximations (i.e., particle filters). Unfortunately, particle-filters tend to be very inefficient, thus greatly limiting the applicability of the approach. We propose an alternative approach based on digitizing the hypothesis space into a large number of hypotheses (on the order of 100,000). The approach has not been tried in the past because, in principle, solving the filtering equations requires order n-squared operations per time step, where n is the number of hypotheses. This means that as the hypothesis space expands, solving the filtering equations becomes rapidly prohibitive.

We show that in many problems, one can make use of the spatial-temporal structure of the hypothesis space and we propose an algorithm to solve the filtering operations in order n operations, vastly reducing the computational strain on the system. In practice, this allows handling hundreds of thousands of hypothesis in real time. We illustrate how the algorithm works for the problem of tracking human faces in real time. In this problem, possible object locations and scales (states) arrange in a three-dimensional topology (two dimensions for location and one for scale). Rectangular convolution kernels capture movement uncertainty over scales and locations. Interestingly the resulting architecture resembles the functional architecture of primary visual cortex, suggesting

explanations for the computational role of forward, lateral and top-down connections in V1.