Spatial distance dependent Chinese restaurant processes for image segmentation

Soumya Ghosh[1], **Andrei B. Ungureanu**[2], **Erik B. Sudderth**[1], and **David M. Blei**[3]

[1]Department of Computer Science, Brown University, {sghosh,sudderth}@cs.brown.edu
[2]Morgan Stanley, andrei.b.ungureanu@gmail.com
[3]Department of Computer Science, Princeton University, blei@cs.princeton.edu

Abstract

The *distance dependent Chinese restaurant process* (ddCRP) was recently introduced to accommodate random partitions of non-exchangeable data [1]. The ddCRP clusters data in a biased way: each data point is more likely to be clustered with other data that are near it in an external sense. This paper examines the ddCRP in a spatial setting with the goal of natural image segmentation. We explore the biases of the spatial ddCRP model and propose a novel hierarchical extension better suited for producing "human-like" segmentations. We then study the sensitivity of the models to various distance and appearance hyperparameters, and provide the first rigorous comparison of nonparametric Bayesian models in the image segmentation domain. On unsupervised image segmentation, we demonstrate that similar performance to existing nonparametric Bayesian models is possible with substantially simpler models and algorithms.

1 Introduction

The Chinese restaurant process (CRP) is a distribution on partitions of integers [2]. When used in a mixture model, it provides an alternative representation of a Bayesian nonparametric Dirichlet process mixture—the data are clustered and the number of clusters is determined via the posterior distribution. CRP mixtures assume that the data are exchangeable, i.e., their order does not affect the distribution of cluster structure. This can provide computational advantages and simplify approximate inference, but is often an unrealistic assumption.

The *distance dependent Chinese restaurant process* (ddCRP) was recently introduced to model random partitions of non-exchangeable data [1]. The ddCRP clusters data in a biased way: each data point is more likely to be clustered with other data that are near it in an external sense. For example, when clustering time series data, points that closer in time are more likely to be grouped together. Previous work [1] developed the ddCRP mixture in general, and derived posterior inference algorithms based on Gibbs sampling [3]. While they studied the ddCRP in time-series and sequential settings, ddCRP models can be used with any type of distance and external covariates. Recently, other researchers [4] have also used the ddCRP in non-temporal settings.

In this paper, we study the ddCRP in a spatial setting. We use a spatial distance function between pixels in natural images and cluster them to provide an unsupervised segmentation. The spatial distance encourages the discovery of connected segments. We also develop a region-based hierarchical generalization, the rddCRP. Analogous to the hierarchical Dirichlet process (HDP) [5], the rddCRP clusters groups of data, where cluster components are shared across groups. Unlike the HDP, however, the rddCRP allows within-group clusterings to depend on external distance measurements.

To demonstrate the power of this approach, we develop posterior inference algorithms for segmenting images with ddCRP and rddCRP mixtures. Image segmentation is an extensively studied area, which