Statistical Modeling and Conceptualization of Visual Patterns

Song Chun Zhu (Department of Computer Science, The Ohio State University)
szhu@cis.ohio-state.edu

Abstract

The objective of perceptual organization (grouping, segmentation and recognition) is to parse generic natural images into their constituent components which are respectively instances of a wide variety of visual patterns in our visual environment. These visual patterns are fundamentally stochastic processes governed by probabilistic models which ought to be learned from the statistics of natural images. In this paper, we divide existing models into four categories: *descriptive models, causal models, generative models, discriminative models*, and review the objectives, principles, theories, and typical models in each category along the progress in studying natural image statistics. The central theme of this epistemological paper is to study the relationships between the four types of models and to pursue a unified mathematical framework for the conceptualization and modeling of various visual patterns. Indeed, many vision concepts can be rigorously defined only in the context of explicit mathematical models. It is also desirable that under this mathematical framework statistical models for various visual patterns form a "continuous" spectrum – in the sense that they belong to a serial of probability families and are learned under one principle. These statistics models and concepts should amount to a visual language which is essential for building effective, robust, and generic vision systems. There is still a long way to go before such a general visual language can be established, nevertheless, the mathematical framework becomes increasingly clear in recent years. This paper is an attempt to summarize our current understanding of the framework, for the purpose of effective communication between various image search streams.