

Chapter 1

Introduction

Visual modeling in this book is addressed mainly from the computational viewpoint. It concerns how to define an objective function for the optimal solution to a vision problem and how to find the solution. The reason for defining the solution in an *optimization* sense is due to various uncertainties in vision processes. It may be difficult to find the perfect solution and we usually look for some optimal solution to optimally satisfy certain constraints.

Contextual constraints are ultimately necessary in the interpretation of visual information. A scene is understood in their spatial and visual context of the objects in it; the objects are recognized in the context of object features at a lower level representation; the object features are identified based on the context of primitives at an even lower level; and the primitives are extracted in the context of image pixels at the lowest level of abstraction. The use of contextual constraints is indispensable for a capable vision system.

Markov random field (MRF) theory provides a convenient and consistent way of modeling context dependent entities such as image pixels and other spatially correlated features. This is achieved through characterizing mutual influences among such entities using MRF probabilities. The practical use of MRF models is largely ascribed to the equivalence between MRFs and Gibbs distributions established by Hammersley and Clifford (1971) and further developed by Besag (1974). The MRFs-Gibbs equivalence gives an explicit formula for the joint distributions of MRFs. This enables us to model vision problems by a *mathematically* sound yet tractable means for the image analysis in the Bayesian framework (Grenander 1983; Geman and Geman 1984). From the *computational* perspective, the local property of MRFs leads to algorithms which can be implemented in a local and massively parallel manner. Furthermore, MRF theory provides a foundation for multi-resolution computation (Gidas 1989).

For the above reasons, MRFs have been widely employed to solve vision problems at all levels. Most of the MRF models are for low level processing. These include image restoration and segmentation, surface reconstruction,

edge detection, texture analysis, optical flow, shape from X, active contours, deformable templates, data fusion, visual integration, and perceptual organization. The use of MRFs in high level vision, such as for object matching and recognition, has also emerged in recent years. The interest in MRF modeling in computer vision is still increasing, as reflected by books as well as journal and conference papers published in recent years¹.

MRF theory tells us how to model the *a priori* probability of contextual dependent patterns, such as a class of textures and an arrangement of object features. A particular MRF model favors its own class of patterns by associating them with larger probabilities than other pattern classes. MRF theory is often used in conjunction with statistical decision and estimation theories, so as to formulate objective functions in terms of established optimality principles. *Maximum a posteriori* (MAP) probability is one of the most popular statistical criteria for optimality and in fact, has been the most popular choice in MRF vision modeling. MRFs and the MAP criterion together give rise to the MAP-MRF framework adopted in this book as well as in most other MRF works. This framework, advocated by Geman and Geman (1984) and others, enables us to develop algorithms for a variety of vision problems systematically using rational principles rather than relying on *ad hoc* heuristics. See also introductory statements in (Mardia 1989; Chellappa and Jain 1993; Mardia and Kanji 1994).

An objective function is completely specified by its *form*, *i.e.* the parametric family, and the involved *parameters*. In the MAP-MRF framework, the objective is the joint posterior probability of the MRF labels. Its form and parameters are determined, in turn, according to the Bayes formula, by those of the joint prior distribution of the labels and the conditional probability of the observed data. "A particular MRF model" referred in the previous paragraph means a particular probability function (of patterns) specified by the functional form and the parameters. Two major parts of the MAP-MRF modeling is to derive the form of the posterior distribution and to determine the parameters in it, so as to completely define the posterior probability. Another important part is to design optimization algorithms for finding the maximum of the posterior distribution.

This book is organized in four parts in accordance with the motivations and issues brought out above. The first part (Chapter 1) introduces basic notions, fundamentals and background materials, including labeling problems, relevant results from MRF theory, optimization-based vision and the systematic MAP-MRF approach. The second part (Chapters 2 – 5) formulates various MRF models in low and high level vision in the MAP-MRF framework and studies the related issues. The third part (Chapters 6 – 7) addresses the problem of MRF parameter estimation. Part four (Chapters 8 – 9) presents

¹There are numerous recent publications in this area. They are not cited here to keep the introductory statements neat. They will be given subsequently.