Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing 26 (2008) 361–371
www.elsevier.com/locate/imavis

Segmentation of color images via reversible jump MCMC sampling

Zoltan Kato *

University of Szeged, Institute of Informatics, P.O. Box 652, H-6701 Szeged, Hungary

Received 25 February 2005; received in revised form 5 June 2006; accepted 8 December 2006

Abstract

Reversible jump Markov chain Monte Carlo (RJMCMC) is a recent method which makes it possible to construct reversible Markov chain samplers that jump between parameter subspaces of different dimensionality. In this paper, we propose a new RJMCMC sampler for multivariate Gaussian mixture identification and we apply it to color image segmentation. For this purpose, we consider a first order Markov random field (MRF) model where the singleton energies derive from a multivariate Gaussian distribution and second order potentials favor similar classes in neighboring pixels. The proposed algorithm finds the most likely number of classes, their associated model parameters and generates a segmentation of the image by classifying the pixels into these classes. The estimation is done according to the Maximum A Posteriori (MAP) criterion. The algorithm has been validated on a database of real images with human segmented ground truth.

Keywords: Unsupervised image segmentation; Color; Parameter estimation; Normal mixture identification; Markov random fields; Reversible jump Markov chain Monte Carlo; Simulated annealing

1. Introduction

MRF modeling and MCMC methods are successfully used in different areas of image processing. In fact, the simplest statistical model for an image consists of the probabilities of pixel classes. The knowledge of the dependencies between nearby pixels can be modeled by a MRF. Such models are much more powerful [1,2], even if it is not easy to determine the values of the parameters which specify a MRF. If each pixel class is represented by a different model then the observed image may be viewed as a sample from a realization of an underlying label field. Unsupervised segmentation can therefore be treated as an *incomplete data problem* where the color values are observed, the label field is missing and the associated class model parameters, *including the number of classes*, need to be estimated. Such problems are often solved using MCMC procedures. Although the general theory and methodology of these algorithms are fairly standard, they have their limitations in case of problems with parameters of varying dimension. Recently, a novel method, called reversible jump MCMC (RJMCMC), has been proposed by Green [3]. This method makes it possible to construct reversible Markov chain samplers that jump between parameter subspaces of different dimensionality. In this paper, we will develop a RJMCMC sampler for identifying multi-variate Gaussian mixtures. In particular, we will apply this technique to solve the unsupervised color image segmentation problem in a Markovian framework.

Due to the difficulty of estimating the number of pixel classes (or clusters), unsupervised algorithms often assume that this parameter is *known a priori* [4,5]. When the number of pixel classes is also being estimated, the unsupervised segmentation problem may be treated as a *model selection problem* over a combined model space. Basically, there are two approaches in the literature. One of them is an exhaustive search of the combined parameter space [6,7]: segmentations and parameter estimates are obtained via an iterative algorithm by alternately sampling the label field based on the current estimates of the parameters.

* Tel.: +36 62 546 399; fax: +36 62 546 397.
E-mail address: kato@inf.u-szeged.hu
URL: http://www.inf.u-szeged.hu/~kato/

0262-8856/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2006.12.004

Then the maximum likelihood estimates of the parameter values are computed using the current labeling. The resulting estimates are then applied to a model fitting criterion to select the optimum number of classes. Another approach consists of a two step approximation technique [1,8]: the first step is a coarse segmentation of the image into the most likely number of regions. Then the parameter values are estimated from the resulting segmentation and the final result is obtained via a supervised segmentation.

Our approach consists of building a Bayesian color image model using a first order MRF. The observed image is represented by a mixture of multivariate Gaussian distributions while inter-pixel interaction favors similar labels at neighboring sites. In a Bayesian framework [9], we are interested in the *posterior distribution* of the unknowns given the observed image. Herein, the unknowns comprise the hidden label field configuration, the Gaussian mixture parameters, the MRF hyperparameter, and the number of mixture components (or classes). Then a RJMCMC algorithm is used to sample from the whole posterior distribution in order to obtain a MAP estimate via simulated annealing [9]. Until now, RJMCMC has been applied to univariate Gaussian mixture identification [10] and its applications in different areas like inference in hidden Markov models [11], intensity-based image segmentation [12], and computing medial axes of 2D shapes [13]. The novelty of our approach is twofold: first, we extend the ideas in [10,12] and propose a RJMCMC method for identifying multi-variate Gaussian mixtures. Second, we apply it to unsupervised color image segmentation. RJMCMC allows us the direct sampling of the whole posterior distribution defined over the combined model space thus reducing the optimization process to a single simulated annealing run. Another advantage is that no coarse segmentation neither exhaustive search over a parameter subspace is required. Although for clarity of presentation we will concentrate on the case of three-variate Gaussians, it is straightforward to extend the equations to higher dimensions.

2. Color image segmentation model

The model assumes that the real world scene consists of a set of regions whose observed color changes slowly, but across the boundary between them, they change abruptly. What we want to infer is a *labeling* ω consisting of a simplified, abstract version of the input image: regions has a constant value (called a *label* in our context) and the discontinuities between them form a curve – the contour. Such a labeling ω specifies a *segmentation*. Taking the probabilistic approach, one usually wants to come up with a *probability measure* on the set Ω of all possible segmentations of the input image and then select the one with the highest probability. Note that Ω is finite, although huge. A widely accepted standard, also motivated by the human visual system [14,15], is to construct this probability measure in a Bayesian framework [16–18]. We will assume that we have a set of observed (Y) and hidden (X) random variables. In our context, the *observation* $\mathscr{F} \in Y$ represents the color values used for partitioning the image, and the hidden entity $\omega \in X$ represents the segmentation itself. Hence the observed image $\mathscr{F} = \{\vec{f}_s | s \in \mathscr{S}, \forall i : 0 < \vec{f}_s^i < 1\}$ consists of three spectral component values at each pixel s denoted by the vector $\vec{f}_s$. Note that color components are normalized. Furthermore, a segmentation ω assigns a label ω_s from the set of labels $\Lambda = \{1, 2, \ldots, L\}$ to each site s.

First, we have to quantify how well any occurrence of ω fits $\mathscr{F}$. This is expressed by the probability distribution $P(\mathscr{F}|\omega)$ – the *imaging model*. Second, we define a set of properties that any segmentation ω must posses regardless the image data. These are described by $P(\omega)$, the *prior*, which tells us how well any occurrence ω satisfies these properties. For that purpose, ω_s is modeled as a discrete random variable taking values in Λ. The set of these labels $\omega = \{\omega_s, s \in \mathscr{S}\}$ is a random field, called the *label process*. Furthermore, the observed color features are supposed to be a realization $\mathscr{F}$ from another random field, which is a function of the label process ω. Basically, the *image process* $\mathscr{F}$ represents the manifestation of the underlying label process. The multivariate Normal density is typically an appropriate model for such classification problems where the feature vectors $\vec{f}_s$ for a given class λ are mildly corrupted versions of a single mean vector μ_λ [19,20]. Applying these ideas, the *image process* $\mathscr{F}$ can be formalized as follows: $P(\vec{f}_s|\omega_s)$ follows a three-variate Gaussian distribution $N(\vec{\mu}, \mathbf{\Sigma})$, each pixel class $\lambda \in \Lambda = \{1, 2, \ldots, L\}$ is represented by its mean vector $\vec{\mu}_\lambda$ and covariance matrix $\mathbf{\Sigma}_\lambda$. As for the *label process* ω, a MRF model is adopted [21] over a nearest neighborhood system. According to the *Hammersley–Clifford theorem* [9], $P(\omega)$ follows a Gibbs distribution:

$$P(\omega) = \frac{1}{Z}\exp(-U(\omega)) = \frac{1}{Z}\exp\left(-\sum_{C\in\mathscr{C}} V_C(\omega_C)\right), \tag{1}$$

where $U(\omega)$ is called an *energy function*, $Z = \sum_{\omega\in\Omega}\exp(-U(\omega))$ is the normalizing constant (or *partition function*) and V_C denotes the *clique potentials* of cliques $C \in \mathscr{C}$ having the label configuration ω_C. The prior $P(\omega)$ will represent the simple fact that segmentations should be locally homogeneous. Therefore we will define clique potentials V_C over pairs of neighboring pixels *(doubletons)* such that similar classes in neighboring pixels are favored

$$V_C = \beta \cdot \delta(\omega_s, \omega_r) = \begin{cases} +\beta & \text{if } \omega_s \neq \omega_r, \\ -\beta & \text{otherwise,} \end{cases} \tag{2}$$

where β is a hyper-parameter controlling the interaction strength. As β increases, regions become more homogeneous. The energy is proportional to the length of the region boundaries. Thus homogeneous segmentations will get a higher probability, as expected.

Factoring the above distributions and applying the Bayes theorem gives us the *posterior* distribution $P(\omega|\mathscr{F}) \propto P(\mathscr{F}|\omega)P(\omega)$. Note that the constant factor $1/P(\mathscr{F})$ has been dropped as we are only interested in $\hat{\omega}$