Parameter Estimation in MRF Line Process Models

Sateesha G. Nadabar Anil K. Jain
Department of Computer Science
Michigan State University
East Lansing MI 48824

Abstract

The specification of parameters of a Markov random field (MRF) prior distribution is one of the difficult problems in using these models for image restoration, edge detection and segmentation. Often, the admissible parameter space for the models is quite large which is difficult to explore by trial and error. We present a new scheme for the estimation of the MRF line process parameters which uses geometric CAD models of the objects in the scene. The models are used to generate synthetic images of the objects from random view points. The edge maps computed from the synthesized images are used as training samples to estimate the line process parameters using a least squares method. We show that this parameter estimation method is useful for detecting edges in range as well as intensity images.

1 Introduction

Markov random field (MRF) models have been successfully used to represent contextual information in many 'site' labeling problems. A site labeling problem involves classification of each site (pixel, edge element, region) into a certain number of classes based on an observed value (or vector) at each site. Contextual information plays an important role here because the true label of a site is assumed to be compatible with the labels of the neighboring sites. Markov random fields are appropriate models of context because they can be used to specify this spatial dependency or spatial a priori distribution.

An MRF prior model is specified in terms of certain parameters, called the clique parameters. These parameters correspond to the clique potential values of an equivalent Gibbs Random Field (GRF) representation [8] (Because of this parameterization of MRFs in terms of GRFs, we use the terms MRF and GRF interchangeably). In most of the previous work using MRF models, the parameters of the prior model have been determined in an ad hoc fashion. A significant limitation in specifying these parameters empirically is that, as the neighborhood size and consequently the number of parameters increases, the parameter space to be explored becomes unwieldy. Therefore, a systematic method for specification of these parameters is of significant interest.

Several schemes have been proposed in the computer vision literature to estimate the parameters of an MRF [4]. For MRFs defined on pixel sites, these schemes have been applied with considerable success, particularly when MRF models are used to model image texture. For MRFs defined on edge sites (line variables used to denote discontinuity between adjacent pixels), however, no parameter estimation technique is available. Silverman et al. [12] have obtained limited success with an analytical model to compute the penalty for edges.

In this paper, we propose a new method to estimate MRF line process parameters which is useful to find object boundaries in intensity or range images. We assume that geometric CAD models of some of the 2D or 3D objects which are expected to be present in the scene are available. The models of the objects are used to synthesize multiple images of the objects from different view points. The frequency of local edge configurations estimated from these synthesized edge maps is then used to estimate the line process parameters. This estimation scheme has been applied to detect edges in range and intensity images and its performance is demonstrated on several real images. The edge maps obtained using estimated parameters are better than edge maps using ad hoc parameters.

2 Background

This section introduces some notation, and defines a few preliminary concepts pertaining to MRFs. The site labeling problem along with a Bayesian solution
Consider the image lattice shown in Figure 1. The edge sites in the lattice indicate whether an edge exists between the two pixels it separates. A stochastic process defined on the edge lattice is usually referred to as a line process [8]. The line process is very useful in edge detection for representing natural assumptions about the structure of physical edges such as continuity and smoothness. Let 
\[ L = \{ L_1, L_2, \ldots, L_M \} \]
be the M-tuple random vector representing the labels at the M sites of an edge lattice, S. A raster scan ordering of the 2-dimensional lattice is used for notational convenience. Since we are interested in edge detection, each site can be assigned one of the two possible labels, \{E, NE\} depending on the presence or absence of an edge, respectively. Note that while a pixel process is measurable (observable), a line process is not. We use indirect measurements for edge sites in terms of measurements at pixels in the neighborhood of the edge site [2, 10]. Details about how the observations at edge sites are derived are given in [10]. Let vector \( Y = \{ Y_1, Y_2, \ldots, Y_M \} \) represent the indirect measurements at the M edge sites. We will represent realizations of random variables by the corresponding lower case symbols.

We will now introduce the Bayesian formulation of the edge labeling problem. We have two types of information available about the site labels: (i) the measurements \( Y \), and (ii) the spatial dependencies among the labels \( L_i \) represented by the probability distribution \( P(I) \). These two types of information can be combined optimally using the Bayes' rule. \( P(I) \) is assumed here to be an MRF distribution with the energy function,

\[
U(I) = \sum_{c \in C} V_c(I). \tag{1}
\]

Here, \( C \) is the set of all cliques with respect to the neighborhood system \( N \) used. We use the second-order neighborhood shown in Figure 2. Figure 3 shows all the cliques corresponding to this neighborhood system. A clique potential typically represents a penalty value attributed to a particular clique configuration. For example, a clique configuration that symbolizes a broken edge is typically attributed a higher clique potential than a clique configuration that represents continuity of edges.

Under certain assumptions [10], the \textit{a posteriori} probability distribution for the site labels \( L \), given the observations \( Y = y \) also has the form of a Gibbs random field

\[
P(I|y) = e^{-U(I|y)}/Z_y, \tag{2}
\]

where \( Z_y \) is a normalizing constant, and the corresponding energy function is

\[
U(I|y) = \sum_{i=1}^{M} [-\ln(f(y_i|l_i))] + \sum_{c \in C} V_c(I). \tag{3}
\]

The site labeling problem can now be stated as follows: given the observation vector \( y \), estimate the label vector \( I \) such that \( U(I|y) \) is minimized. In this paper, we have used the Highest Confidence First (HCF) algorithm proposed by Chou and Brown [2]. The HCF algorithm assigns a confidence value to each site which is the maximum amount of energy reduction that would be achieved by changing its label to
some other label. The site that has the highest confidence value associated with it is visited next and the confidence values for its neighboring sites are updated accordingly. Chou and Brown have demonstrated that for edge detection, the HCF algorithm performs better than other optimization algorithms. Our edge detection algorithm is summarized below.

1. For each edge site, compute the edge likelihood ratio, $L_e = \frac{P(y_t|e)}{P(y_t|ne)}$. For range images, this step consists of the following three operations.
   (a) Compute the jump edge likelihood ratio, $L_j$.
   (b) Compute the crease edge likelihood ratio, $L_c$.
   (c) Edge likelihood ratio, $L_e = \max(L_j, L_c)$.

2. Use HCF algorithm [2] to obtain the desired edge labeling.

The main parameters to be specified in this algorithm are the clique potential function values of the MRF model.

3 Estimation of Line Process Parameters from Object Models

For MRFs defined on edge sites, we face the following difficulties in parameter estimation: (i) Line variables are unobservable and therefore, the true line process configurations are seldom available; (ii) Even when such a true edge map is available, some edge configurations occur very infrequently; (iii) Techniques of simultaneous parameter estimation and labeling [11] are also not suitable here because the local joint probabilities estimated from the given, possibly incorrect, labeling are very unreliable. Because of these difficulties, the most commonly used method of specifying the parameters of the line process has been by trial and error.

We assume that a geometric model is available for some typical objects expected to be present in the scene. Our approach is to take a random training sample (set of edge maps) from the population of edge maps derived from all possible viewpoints of all the available models. It is important to note that a random sample here means an edge map from a randomly chosen viewpoint of a randomly chosen object model. We emphasize that the advantage of using the object models to synthesize the desired edge maps is the availability of true labels at the edge sites. By generating images from several random viewpoints of object models, we essentially obtain samples of edge maps which can be used for estimation of line process parameters.

3.1 Least Squares Formulation

For a second-order neighborhood, let $l_{k_i}$ represent the set of labels assigned to the neighbors of edge site $t$. Let $\Theta = [\theta_1, \ldots, \theta_p]^T$ be the parameter (clique potentials) vector. For an isotropic and homogeneous model with the second-order neighborhood shown in Figure 2, the number of line process parameters is 35, when all possible clique configurations are used as parameters. When the parameters are specified in an ad hoc fashion, however, a portion of the possible clique configurations is assigned zero potential in order to reduce the number of parameters of the model. In section 3.2, we show that there exists a parameterization with fewer parameters.

The following formulation is based on the treatment by Derin and Elliot [4]. Let $U(t, l_{k_i}, \Theta)$ be the sum of the potential functions for all cliques to which the edge site $t$ belongs.

$$U(t, l_{k_i}, \Theta) = \sum_{c \in G} V_c(l) = \Phi^T(t, l_{k_i})\Theta = \sum_{i=1}^p n_i\theta_i,$$

where $$\Phi^T(t, l_{k_i}) = [n_1, ..., n_p]^T$$

and $n_i$ represents the number of clique configurations of type $i$ in the neighborhood of site $t$.

The local conditional probability at site $t$ can be written as

$$P(l_t|l_{k_i}) = \frac{P(l_t, l_{k_i})}{P(l_{k_i})} = e^{-U(l_t, l_{k_i}, \Theta)} Z_t(l_{k_i}, \Theta).$$

where $Z_t(l_{k_i}, \Theta) = \sum_{l_t} U(l_t, l_{k_i}, \Theta)$.

Substituting $l_t = e$ and $l_t = ne$ in Eq. (5) and combining the two resulting equations, we get

$$e^{U(e, l_{k_i}, \Theta)} - e^{U(ne, l_{k_i}, \Theta)} = \frac{P(ne, l_{k_i})}{P(e, l_{k_i})}.$$

Taking logarithms on both sides and substituting for $U(l_t, l_{k_i}, \Theta)$ from Eq. (4), we have

$$(\Phi(e, l_{k_i}) - \Phi(ne, l_{k_i}))^T \Theta = \log \frac{P(ne, l_{k_i})}{P(e, l_{k_i})}.$$

The ratio $P(ne, l_{k_i})/P(e, l_{k_i})$ can be estimated from the edge maps synthesized from object models. By substituting each value of $l_{k_i}$ in Eq. (6), we will obtain 256 ($2^8$ possible neighborhood configurations) equations in $p$ unknowns. This overdetermined linear system of equations can be solved using the method of least squares.
3.2 A Canonical MRF Representation: Normalized Potential Functions

For a specific Markov or Gibbs random field, there could be many equivalent potential functions which specify the same distribution [9]. So an important question to ask is "Is there an equivalent MRF representation, for a given neighborhood, with the minimal number of free parameters?"

It turns out that such a representation exists and the associated potential functions are called the normalized potential functions which are defined as follows.

**Definition:** A potential function $V = \{ V_A : A \subseteq S \}$ is a normalized potential if $V_A(1) = 0$ whenever $s = \text{SN}$ for some $s \in A$.

The choice of the label \text{SN} in the above definition was arbitrary. It is also known that a unique normalization potential called the canonical potential exists for every MRF [9]. The clique configurations with non-zero canonical potentials are shown in Figure 4. All other clique configurations have at least one edge site with label \text{SN} and so are assigned zero potential according to the above definition. It is seen that only 8 parameters are needed to represent the MRF defined on the second-order neighborhood corresponding to the 8 configurations shown in Figure 4. Therefore, the number of line process parameters that need to be estimated is reduced considerably (from 35 to 8). These parameters are still estimated using the least squares method.

Figure 4: Clique configurations with non-zero canonical potentials and their values estimated from CAD models of 10 3D objects by the least squares method.

![Figure 4](image-url)

**Figure 4:** Clique configurations with non-zero canonical potentials and their values estimated from CAD models of 10 3D objects by the least squares method.

4 Edge Detection in Range Images

We view the problem of edge detection in range images as a site labeling problem. We use edge strength statistics computed from the range values observed at pixels in the neighborhood of the edge sites to determine the likelihood of edge labels. For details, see [10].

We created a CAD model database of 10 3D objects [8]. We generated 100 range images corresponding to 100 random views for each of the 10 model objects and computed the local conditional probabilities of edge labels.

The estimated values of the 8 line process parameters from the local characteristics obtained from all 10 CAD models are shown in Figure 4. Figure 5 shows the clique potential values used by Jain and Nadabar [10] which were chosen in an ad hoc fashion. Notice that the estimated value for the same clique configuration is drastically different from that selected in an ad hoc fashion. We must point out that these two sets of parameter values are not directly comparable, because the number of parameters in the two formulations are different.

Edge detection experiments were performed using range images obtained from the Technical Arts 100X scanner in our laboratory. We report here results on two range images. The same cliques and potential function values were used for all the images. Figure 6 shows the results for a 194 x 218 image of a pair of wooden blocks and Figure 7 shows the results for an 83 x 156 image of a rotationally symmetric industrial object. While all the jump and crease edges have been detected and localized quite accurately with both estimated and ad hoc clique potentials, it is seen that the edges detected using estimated clique potentials are visually better than the edges detected using ad-hoc clique potentials. Further, edges detected using the estimated normalized clique potentials contain fewer spurious edges than with ad-hoc clique potentials.

![Figure 5](image-url)

**Figure 5:** Ad hoc clique potential values used to obtain edges in Figure 6(d).
Edge Detection in Intensity Images

The parameter estimation technique introduced in Section 3 is applicable to edge detection in intensity images also. For intensity images, we synthesize edge maps from the polygonal models of the 2D objects in the database. The appearance of the objects from different viewpoints can be simulated by performing a two-dimensional rotation of the models. We used the algorithm proposed by Chou et al. [2] for edge detection.

The experiments were performed using a database of 10 2D objects. For each object in the database, 100 different edge images were generated. The values of the 8 canonical clique potentials are given in Figure 8. Comparing these parameter values with the values given in Figure 4, it is seen that some of the clique potential values are quite different in the two cases. This reflects the nature of edge configurations in the range images of 3D objects and intensity images of 2D objects.

Figure 9 shows the results of edge detection on a 128 x 128 image of a scene in a manufacturing plant (this image was provided to us by Odetics Corporation). The objects in this image are not included in the database of 2D objects used to estimate the clique potential values in Figure 8. The quality of edge detection using estimated parameters is seen to be comparable to that obtained using parameters specified in an ad hoc fashion. Since the objects in the scene are not included in the database of objects used to estimate the line process parameters, the results confirm that it is sufficient to use only typical object models instead of the actual object models present in the scene. This is because the line process prior model captures only the local structural characteristics of edges. It is also seen from Figure 9 that, MRF-based edge detection is slightly superior to the Canny edge detector [1] output because the latter results in oversmoothed edges especially near the corners. Canny edge detector also fails to detect low contrast edges.

6 Conclusions

We have presented a new off-line scheme for estimating parameters of MRF (GRF) line process when they are used to find edges in intensity or range images. This estimation method utilizes the geometric CAD models of 2D or 3D objects expected to be present in the scene. The use of CAD models here is significant, because CAD models are being increas-
Figure 9: Step edge detection results for the Odetix image: (a) input image; (b) edges detected using clique parameters estimated from 2D object models; (c) edges detected using parameters specified \textit{ad hoc}; (d) edges detected using the Canny edge detector (with hysteresis threshold values, 0.9 and 0.5 and width of the gaussian $\sigma = 1$ pixel).

...equivalently used for object representation in high-level vision problems [7]. These models are used to generate synthetic images of the objects from several random view points which provide a sample of edge maps for the MRF line process. The parameters are then estimated using linear least squares fitting. The idea of equivalent representation of an MRF using canonical clique potentials was used to reduce the number of unknown clique potential parameters. Experimental results presented on several range and intensity images indicate that our parameter estimation method is superior to the commonly used procedure of specifying the clique potential parameters in an \textit{ad hoc} fashion. The results also indicate that it is sufficient to use only typical objects to estimate the line process parameters.

References


