SEGMENTATION AND COOPERATIVE FUSION
OF LASER RADAR IMAGE DATA*

M. Beckerman and F. J. Sweeney

Center for Engineering Systems Advanced Research

SPIE—The International Society for Optical Engineering, Conference on Sensor Fusion and Aerospace Applications II, part of SPIE’s International Symposium on Optical Engineering and Photonics in Aerospace Sensing, April 4–8, 1994, Orlando, FL

SEGMENTATION AND COOPERATIVE FUSION
OF LASER RADAR IMAGE DATA

Martin Beckerman and Frank J. Sweeney
Center for Engineering Systems Advanced Research
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6364

ABSTRACT

In segmentation, the goal is to partition a given 2D image into regions corresponding to the meaningful surfaces in the underlying physical scene. Segmentation is frequently a crucial step in analyzing and interpreting image data acquired by a variety of automated systems ranging from indoor robots to orbital satellites. In this paper, we present results of a study of segmentation by means of cooperative fusion of registered range and intensity images acquired using a prototype amplitude-modulated CW laser radar. In our approach, we consider three modalities - depth, reflectance and surface orientation. These modalities are modeled as sets of coupled Markov random fields for pixel and line processes. Bayesian inferencing is used to impose constraints of smoothness on the pixel process and linearity on the line process. The latter constraint is modeled using an Ising Hamiltonian. We solve the constrained optimization problem using a form of simulated annealing termed quenched annealing. The resulting model is illustrated in this paper in the rapid quenched, or iterated conditional mode, limit for several laboratory scenes.

1. INTRODUCTION

In segmentation, the goal is to partition a given 2D image into regions corresponding to the meaningful surfaces in the underlying physical scene. Segmentation is frequently a crucial step in analyzing and interpreting image data acquired by a variety of automated systems ranging from orbital satellites to indoor mobile robots. While considerable work has been done with intensity images, relatively few studies have been reported for range data, and fewer still results have been presented for pairs of registered range and intensity images. Range images have been segmented by Besl and Jain, using surface curvature estimates supplemented by an iterative region-growing method based upon variable-order surface fitting, and by Rimet and Cohen using a window-region-based approach with a maximum likelihood classifier. Hoffman and Jain segmented images into surface patches by means of a clustering algorithm, while Fan, Medioni and Nevatia segmented images by identifying surface patch boundaries. A hybrid, region plus edge method has been developed by Yokoya and Levine, and also by Jain and Nadabar. A summary of several earlier studies has been presented by Besl and Jain.

Coupled Markov random fields have been used to solve a variety of inverse imaging problems including range image segmentation. An important observation regarding coupled Markov random fields is that they provide a natural framework for fusing data from different modalities. As noted by Poggio, Gamble and Little, Chou and Brown, and Gamble et al, cooperative fusion of different sensory modalities can be performed best at surface boundaries. A similar observation on coincident discontinuities was made some time ago in a non-cooperative, "global and" approach to range and reflectance data fusion by Gil, Mitiche and Aggarwal. Recent work on coupled Markov random fields in which this philosophy has been adopted include the color segmentation investigation by Wright, and the laser range and reflectance edge classification study by Nadabar and Jain.

In this paper, we report results of a study of segmentation by means of cooperative fusion of registered range and intensity images acquired using a prototype amplitude-modulated CW laser radar camera. In the previous studies, depth and surface curvature or orientation were used to obtain results for a variety of table
scenes of complex machine parts and models. The goal in the present work is to segment laboratory scenes containing a number of large objects placed on the floor. The background, consisting of several walls and the floor, is not uniform and must be segmented, as well. In our cooperative fusion process, coupled Markov random fields are introduced for three sensory modalities - depth, surface normals and surface reflectance. In each of the modalities, coupled pixel and line processes are introduced, and appropriate energy terms are defined. The energy terms mediate the fusion process, model the appropriate smoothness and discontinuity constraints, and express the commonality conditions for segmentation.

The basic segmentation model is introduced in Sec. 2. We introduce the MRF model in Sec. 2.1, discuss Bayesian inferencing in Sec. 2.2 and the Gibbs sampler in Sec. 2.3. The surface orientation approach is discussed in greater detail in Sec. 3 and the Ising Hamiltonian is presented in Sec. 4. Our Ising model for the line process is free of many of the difficulties of the commonly encountered clique-potential-based approach. A sketch of the quenching algorithm is presented in Sec. 5, together with our results. A brief summary follows in Sec. 6.

2. SEGMENTATION MODEL

2.1 Coupled Markov Random Fields

Let us consider a discrete two-dimensional rectangular lattice, \( \Lambda \), and let \( X_{mn} \) denote a random variable defined on \( \Lambda \), which takes on the values \( x_{mn} \) at lattice sites \((m,n)\). We define a configuration of the lattice system \((X, \Lambda)\) to be a set of values of the random variable, one for each lattice site. This definition of the configurations \( \{x_{mn}\} \) can be written as

\[
X = \{x_{mn}\} = \{x_{mn}; (m,n) \in \Lambda\}
\]

If we now define a joint probability measure, \( p \), on the set of all possible configurations of the random variable over the lattice or any sublattice, then the system \((X, \Lambda, p)\) is a random field.

A Markov random field (MRF) is a random field for which: (1) the joint probability distribution has associated conditional probabilities which are local in character, that is, they obey the Markovian relationship

\[
p(X_{mn} = x_{mn}|\text{all others}) = p(X_{mn} = x_{mn}|X_{rs} = x_{rs}, r \neq mn, r \in N_{mn})
\]

where \( N_{mn} \) is a neighborhood system defined over the lattice; (2) The probability distribution is positive definite for all values of the random variable, and (3) the conditional probabilities are invariant with respect to neighborhood translations.

To construct the image model we introduce Markov random fields \((X, \Lambda, p)\) for our three sensor modalities. Each MRF has an identical structure, but codes different information. The pixel fields are then coupled to line fields describing the discontinuities in the images. The line, or dual, fields contain a pair of horizontal and vertical line elements for each lattice site. These elements are "on" (1) if the horizontal (vertical) pixel differences exceed the threshold setting; otherwise they are "off" (0). Simple, 2 x 1 first differences are used for the range and surface reflectance domains. The primary and dual lattices for the surface orientations will be discussed in shortly in Sec. 3. The overall structure of our image model is illustrated schematically in Fig. 1. Shown in this figure is a depiction of the pixel lattices for the three modalities. For each of these primary lattices there is a coupled dual lattice for the line process.
2.2 Bayesian Estimation

As is the case in inverse problems, we wish to determine the object fields, $f$, for each modality given the measured image fields, $g$. These arrays are related to one another through Bayes' formula

$$p(f|g) = \frac{p(g|f)p(f)}{p(g)} \tag{3}$$

The first term in the numerator is the likelihood, which describes the detection and imaging mechanism. We will assume that the detector response function is linear and can be set equal to unity. We will also assume that the point-spread function is space invariant and can be set to unity, as well. For segmentation purposes and mathematical convenience we assume an I. I. D. Gaussian noise model so that at each lattice site we have

$$p(g_i|f_i) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(f_i - g_i)^2}{2\sigma^2}\right) \tag{4}$$

The second term is the prior. This quantity plays a key role in achieving a physically meaningful segmentation. As already mentioned, the prior is used to model the piecewise-smoothness of the pixel process and the piecewise-linearity of the line process. These are local attributes, defined over the appropriate neighborhood systems. As done previously\(^1\) we employ a neighborhood system for the pixel process containing the eight first and second nearest neighbors and the accompanying line elements. We then exploit the Gibbs-Markov equivalence to cast the Markovian prior distributions into the form of a Gibbs distribution having an energy function, or Hamiltonian, composed of sums of local contributions from the lattice sites. At each site the Gibbs distribution for the prior takes the form

$$p(f_i) = \exp(-\alpha U(f_i)) \tag{5}$$

where $\alpha$ is a Lagrange multiplier which functions as a control parameter, and the energy term has the structure of a sum of two-body interactions over the neighborhood system.

There are several acceptable choices for the energy function. Two selections appropriate for the smoothing task are the entropic and quadratic differences potentials. We employed an entropic form in our previous study, but will select the quadratic type in the present work since it leads to a somewhat easier-to-calculate result in the zero temperature limit. At each lattice site we have
\[ U(f_i) = \sum_{j \in N_i} (f_i - f_j)^2 (1 - l_{ij}) \]  

where \( l_{ij} \) is the line element between the \( i \)th pixel and its four nearest neighbors, and is an effective line element between the \( i \)th pixel and its four next nearest neighbors.

2.3 The Gibbs Sampler

The posterior distribution which we wish to maximize is of the form known as the Gibbs sampler\(^7\). The sampler is written as

\[ p(f|g) = \frac{1}{Z} \exp\left(-\frac{U(f, g)}{T}\right) \]  

where \( T \) is another Lagrange multiplier, usually designated as a temperature, \( Z \) is the partition function, and the energy is the sum of the contributions from each lattice site of the form

\[ U(f_i, g_i) = \alpha U(f_i) + \frac{1}{2\sigma^2} (f_i - g_i)^2 \]  

Gibbsian maximum \textit{a posteriori} segmentation estimates are obtained using a form of simulated annealing which we term quenched annealing. The basic principle in quenched annealing is that in the absence of local minima and phase transitions we can rapidly lower the temperature. In some instances, we may directly go to the \( T = 0 \) limit, where the process is identical to finding the iterated conditional modes\(^2\) of the posterior distribution. We may start with slow cooling and then switch to a fast quench, and couplings between modalities can be temperature-dependent. To illustrate the utility where the process we will present results in this paper exclusively in the \( T = 0 \) limit. We examine the full quenched annealing approach in greater detail in a forthcoming publication.

Fig. 2. Range Image. The brighter the pixel the nearer the object surface to the sensor.  

Fig. 3 Reflectance Image. The brighter the pixel the the greater the surface reflectance.

3. SURFACE ORIENTATION

Amplitude-modulated CW imaging laser radar systems measure the time-of-flight indirectly by detecting the change in phase between transmitted and received signals. In the system under consideration a 820 nm
The dual lattice for surface orientation is constructed by thresholding the symmetric differences

\[ (\Delta f_x)_x^2 + (\Delta f_y)_y^2 > \tau_x \]  \tag{11} 

and

\[ (\Delta f_x)_y^2 + (\Delta f_y)_x^2 > \tau_y \]  \tag{12} 

with
We then define the representation to a (-1) form in order to constrain properties. The linear manifold through the site. As in our previous work, we define one dimensional horizontal and vertical neighborhood systems for the line sites. This allows us to model the linearity properties of the dual lattice in a straightforward manner. The linearity, \( \eta \), of a dual lattice site is defined as the number of active line elements within the pair of three-element line site triplets, or triads, \( L_j \), containing the primary line site plus its two nearest neighbors. More formally, we have

\[
\eta(l_j) = \sum_{i \in L_j} l_i
\]

(15)

The linearities defined in this manner vary from zero (empty) to three (fully occupied). The linearity property is analogous to the smoothness property of the pixel process. We use the linearity property to constrain the line process though a self-coupling term, and to cooperatively fuse the image data by means of lattice-lattice couplings.

4. LATTICE COUPLINGS AND FUSION

Let us recall that each dual lattice contains one horizontal and one vertical line site for each pixel lattice site. As in our previous work, we define one dimensional horizontal and vertical neighborhood systems for the line sites. This allows us to model the linearity properties of the dual lattice in a straightforward manner. The linearity, \( \eta \), of a dual lattice site is defined as the number of active line elements within the pair of three-element line site triplets, or triads, \( L_j \), containing the primary line site plus its two nearest neighbors. More formally, we have

\[
\eta(l_j) = \sum_{i \in L_j} l_i
\]

(15)

The linearities defined in this manner vary from zero (empty) to three (fully occupied). The linearity property is analogous to the smoothness property of the pixel process. We use the linearity property to constrain the line process through a self-coupling term, and to cooperatively fuse the image data by means of lattice-lattice couplings.

4.1 Ising Hamiltonian

In order to construct an appropriate Hamiltonian and sampling algorithm, we transform from the (0,1) representation to a (-1,1) form by introducing spin variables, \( s_i \), through the relation

\[
s_i = 2l_i - 1
\]

(16)

We then define the Hamiltonian, or energy function, for the line process to be

\[
U(s_i, s^{\wedge}_{i}) = \phi(s_i, s^{\wedge}_{i})s_i
\]

(17)

where \( \phi \) is the dual lattice coupling potential

\[
\phi(s_i, s^{\wedge}_{i}) = \sum_{j \neq i} J_{ij}s_js_j + h_i(s^{\wedge}_{i})
\]

(18)
4.2 Exchange Interaction

This lattice potential contains two terms. The first term is the familiar ferromagnetic nearest-neighbor interaction of the Ising model. The exchange integral, $J_{ij}$, assumes the value -1 if $i$ and $j$ are nearest-neighbors, and is zero otherwise. We may write this term as

$$\sum_{j \neq i} J_{ij} s_j s_i = - \sum_{j \neq i} s_j s_i = -s_i (s_{i-1} + s_{i+1})$$  \hspace{1cm} (19)$$

The contribution of the ferromagnetic term to the total energy is minimal whenever the triads of line sites are either fully occupied or empty, and is maximal when the sites are half occupied. There are two important differences between this Hamiltonian and conventional clique-based definitions. First, the large numbers of clique terms and accompanying parameters are eliminated in favor of a simpler set of couplings. Difficulties in parameters assignments for clique potentials have been noted and discussed by Nadabar and Jain\textsuperscript{11}. Second, isolated singletons and doublets are equally discouraged, in contrast to typical clique weightings.

4.3 External Field

The second term mediates the fusion process. This quantity represents the coupling to an external field, in this case, the field produced by the corresponding triads of spin elements for one or more of the other sensor modalities. There are two natural ways of constructing the external fields. We may take the external field as that formed by the primary spin element, $s_i$, of the coupled lattice

$$h_i (s_i^+) s_i = -s_i^+ s_i$$  \hspace{1cm} (20)$$

or we may introduce the field produced by the full triad of line site spin elements

$$h_i (s_i^+) s_i = -s_i^+ s_i [2 + s_i^+ (s_i^{i-1} + s_i^{i+1})]$$  \hspace{1cm} (21)$$

This second field has the property that it is zero if the nearest neighbor spin elements are oriented opposite to the primary spin element, and is maximal when all three elements are aligned in the same direction. The bracketed term is non-negative, and the multiplicative factor defines the overall sign of the field. The cancellation by opposing nearest neighbors eliminates contributions from isolated, noisy elements.

Fig. 6. External Field Coupling Scheme.
5. CALCULATIONS AND RESULTS

In the segmentation calculations, we adopted a cyclic coupling scheme for the external fields. This set of couplings is illustrated in Fig. 6. The T = 0 segmentation algorithm is, as follows:

1. Initialize all dual lattices using the initial pixel values.
2. Select a coupled primary (pixel)-dual lattice.
3. Select a lattice site.
4. Replace the pixel grey level with the mode for the neighborhood.
5. Go back to step 3 and continue the raster scan.
6. Initialize the dual lattice using the updated pixel values.
7. Select a dual lattice site.
8. If flipping the spin lowers the energy then flip the spin; else do not flip.
9. Go back to step 7 and continue the raster scan for the dual lattice.
10. Repeat steps 2 to 9 for the next pair of coupled lattices.
11. Iterate, repeating steps 2 to 10 over the three sets of lattice pairs until finished.

In this algorithm, we set the noise variances to unity. The calculations are fairly insensitive to settings for the control parameter for the prior, as long as it is large compared to unity. Convergence of the algorithm to the desired results occurs in as few as five or six iterations.

To complete the segmentation process we combine the edge maps from the dual lattices for the three sensor modalities. The corresponding segmentation criteria is homogeneity in depth, reflectance and orientation. The joint edge map for the scene shown in Figs. 2 and 3 is displayed in Fig. 7, and the coloring is presented together with the edge map in Fig. 8. We observe that all major surfaces have been labeled. The floor and walls have been segregated from one another and from the objects. No attempt has been made to prune or remove any of the details in these images. Small homogeneous regions such as the bands on the wall partitions, and the dark strip near the floor-wall boundary have been identified, and regions of unit pixel width have been found in the badly images oblique object surfaces.

![Fig. 7. Edge Map of the Laboratory Scene.](image1.jpg)

![Fig. 8. Colored Image Plus Edge Map.](image2.jpg)

A second scene, containing a storage drum in addition to the previous object set is presented in Figs. 9 and 10. The surface orientation primary lattices for the scene are shown in Figs. 11 and 12, and the segmentation results are displayed in Figs. 13 and 14. Again, we find that all surfaces have been labeled.
6. SUMMARY AND CONCLUDING REMARKS

To summarize, we have segmented and segregated a variety of images of laboratory scenes taken with an amplitude-modulated CW laser radar. In our approach, we have cooperatively fused data from three modalities - depth, surface reflectance and surface normals. In examining the edge maps from the three modalities, we found that the depth and reflectance maps were sharp. The surface orientation edge maps were less sharply delineated, but provided the essential crease edge information for the segmentation.

In our coupled Markov random field model, we constructed an Ising Hamiltonian for the line process. Several obvious generalizations are possible. The success of the algorithm provides encouragement that this MRF approach can be implemented in realistic situations without encountering problems with large numbers of arbitrary model parameters associated with clique potentials. We are further encouraged by the ability of the rapid quenching limit of the simulated annealing algorithm to produce useful results. As mentioned above, we will discuss this aspect further in a forthcoming publication.
7. ACKNOWLEDGEMENT

This research was sponsored by the Engineering Research Program of the Office of Basic Energy Sciences of the U.S. Department of Energy under contract No. DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

Fig. 13. Edge Map of the Laboratory Scene.

Fig. 14. Colored Image Plus Edge Map.

8. REFERENCES


