Journal of the Korean Society of Remote Sensing, Vol.15, No.2, 1999, pp.147~158

Efficient Methodology in Markov Random Field Modeling : Multiresolution Structure and Bayesian Approach in Parameter Estimation

Myung-Hee Jung and Eui-Seok Hong

Dept. of Image Processing, Anyang University

피라미드 구조와 베이지안 접근법을 이용한 Markove Random Field의 효율적 모델링

정명희 · 홍의석

안양대학교, 영상처리학과

Abstract : Remote sensing technique has offered better understanding of our environment for the decades by providing useful level of information on the landcover. In many applications using the remotely sensed data, digital image processing methodology has been usefully employed to characterize the features in the data and develop the models. Random field models, especially Markov Random Field (MRF) models exploiting spatial relationships, are successfully utilized in many problems such as texture modeling, region labeling and so on.

Usually, remotely sensed imagery are very large in nature and the data increase greatly in the problem requiring temporal data over time period. The time required to process increasing larger images is not linear. In this study, the methodology to reduce the computational cost is investigated in the utilization of the Markov Random Field. For this, multiresolution framework is explored which provides convenient and efficient structures for the transition between the local and global features. The computational requirements for parameter estimation of the MRF model also become excessive as image size increases. A Bayesian approach is investigated as an alternative estimation method to reduce the computational burden in estimation of the parameters of large images.

Key Words : Markov Random Field, Multiresolution framework, Bayesian, Multivariate Normal Density

요 약 : 지표면에 대한 다양한 정보를 제공해 주는 원격탐사기법은 수 십년 동안 우리의 환경을 관찰하고 이해하는데 중요한 역할을 해왔다. 이러한 원격탐사 자료를 이용하는데 다양한 디지털 영상처리기법이 도입되어 자료에서 관찰되는 여러 가지 특성을 모형화하고 처리하는데 매우 유용하게

Received 2 April 1999 ; Accepted 10 May 1999.

활용되어져 왔다. 화소들 간의 공간적 관계를 고려하는 Markov Random Field (MRF) 모형은 텍스처 모델링이나 영상분할 및 분류와 같은 여러 분야에서 많이 이용되는 모형으로 이것에 기초한 다양한 알고리즘이 발표되었다.

보통 원격탐사 자료는 그 크기가 매우 크고 시간적 간격을 두고 변화를 관측해 가는 경우에는 분석해야 할 자료의 양이 매우 방대하다. 이러한 자료를 처리하는데 걸리는 시간은 처리해야 할 자료의 양과는 비선형적 관계에 있다. 본 논문에서는 MRF를 이용하여 원격탐사 자료를 처리할 때 걸리는 시간을 단축하기 위한 방법론들이 연구되었다. 이를 위해 논리적 구조로 영상을 피라미드형태로 감소하는 크기로 분석하는 multiresolution 구조가 고려되었는데 이는 영상의 거시적 특징과 미세한 특징을 효율적으로 분석할 수 있는 방법을 제공해 준다. 영상의 크기가 커질수록 파라미터 추정 또한 복잡하고 많은 시간을 요하게 된다. 본 논문에서는 이를 위해 Bayesian 방법을 이용하여 원격탐사 영상과 같은 크기가 큰 영상의 MRF 모형의 파라미터를 효율적으로 추정할 수 있는 방법이 제안되어 있다.

1. Introduction

Remote sensing has become an important tool for monitoring a wide variety of targets on the earth's surface, such as vigor, type, and quantity of vegetation, growth of urban areas, and ocean circulation and so on , providing an attractive alternative to ground surveys because remotely sensed imagery provides measurements of the characteristics of a large area and contains various levels of information on landcover. The relatively low cost of satellite images makes large scale environmental monitoring over long time periods possible and our environment has been better understood from these data for decades.

In many problems using the remotely sensed data, digital image processing methodology has been usefully employed to extract information in the observed data characterizing features in the data. Many effective data analysis techniques have been developed using theoretically well-defined approaches in processing and modeling the remotely sensed data. Among these approaches, the use of stochastic models has increased recently, resulting in development of many practical algorithms in the field of image classification, feature extraction, image restoration, and so on. Statistical approaches often apply the Markov random field (MRF) to problems in texture modeling and classification (Cross and Jain, 1983; Dubes and Jain, 1989; Solberg, 1999) and to researches in segmentation and restoration (Dubes and Jain, 1989), permitting the introduction of spatial context. MRF extends Markovian dependence from 1-D to 2-D or 3-D general setting exploiting spatial adjacency relationships.

Remotely sensed data are very large in nature, especially in the problem requiring temporal data over time period. The time required for process of increasing larger images is not linear. For example, processing of a 1024×1024 image using conventional relaxation-type algorithms is greater than 4 times the time for a 512×512 image, because it requires more iterations for convergence.

One way to reduce the computational burden is to reduce the volume of input data. This idea is naturally related to the concept of multiresolution. Representation of an image at multiple resolutions provides convenient and efficient structures for the transition between the local and global features. Furthermore, it often produces better results such as in the early work on edge detection (Jolin and Rosenfeld, 1994). The basic concept is that if the objects or regions of interest are relatively large compared to the pixel size, then the probability of the groups of contiguous pixels being in the same class is much greater than the probability of being in different classes. Many works integrated with multiresolution structure have been explored extensively in a variety of applications and turned out to be useful (Lakshmanan and Derin, 1980; Burt *et al.*, 1981; Rosenfeld, 1984).

In this study, the methodology to reduce the computational cost is explored in the utilization of the Markov Random Field. According to the Markovian property, global features in the image are explained only through the local interactions between the neighboring pixels, which results in the slow computational algorithms. A multiresolution MRF framework would provide an efficient approach that images are processed at a coarse resolution and then progressively refined to the finer resolutions instead of analysis of all the pixels at the full resolution. This would allow global features to propagate relatively quickly. However, there is a fundamental difficulty with the consistent model description at multiple levels since the local dependency assumed in the MRF model is modeled in only one level. Reduced resolution version of images obtained by general sampling or block-to-point type resolution transformation turns out to lose Markovian property except a very few sampling schemes (Jeng, 1989). Therefore, non-Markov fields are to be approximated by MRFs to model the images as MRFs at multi levels. In this study, adapting the approach of normalized group study, a markovain approximation is obtained by bond-moving approximation between the distant pixels.

Parameter estimation is also one inherent difficulty of using MRF models because of the normalizing function in a Gibbs distribution. Instead of computationally complicated classical maximum likelihood estimation method, some alternative techniques such as coding method (Cross and Jain, 1983) or Least Square Error Method (Derin and Elliot, 1987) are suggested. As image size increases, the computational requirements for the parameter estimation of MRF models also become excessive. In this research, to reduce the computational complexity in estimation of the parameters of large images, a Bayesian approach is explored as an alternative estimation method based on the MRF assumption that a global feature is defined through local characteristics.

The paper is organized as follows: processing and modeling images at the multiresolution framework to reduce the computational burden is investigated in Section II. Section III describes the alternative Bayesian approach for the parameter estimation. Section IV includes the evaluation of the proposed algorithms and some results. Finally, Section V has some conclusions.

2. Utilization of Multiresolution MRF framework

It is noted that a unique Gibbs Random Field exists for every MRF and vice versa, as long as the GRF is defined in terms of cliques on a neighborhood system, which is known as the Hammersley-Clifford theorem (Cressie, 1993). This equivalence provides a simple, practical way of specifying MRFs by specifying potentials instead of local characteristics which is one of the difficulties with MRF. It makes the Gibbs distribution a useful model, especially in the context of image modeling and processing. The Gibbs random field is defined as follows:

A random field X = $\{X_i\}$ defined on lattice has a Gibbs Distribution or equivalently is a Gibbs Random Field (GRF) with respect to set of all the cliques, if and only if its joint distribution is of the form

$$P(X = \omega) = \frac{1}{Z} e^{-U(\omega)/T} \quad (1)$$

$U(\omega) = \sum_{all\ c} V_c(\omega)$, where the clique function, or potential function, $V_c(\omega)$ is associated with each clique and depends only on the values at sites in clique c. Various types of clique functions can be selected to formulate a wide variety of Gibbs distributions for both discrete and continuous random fields. T, "temperature", is adapted to isolate the most probable states under the Gibbs distribution by gradually reducing it according to some cooling algorithm (Geman and Geman, 1984).

The neighborhood system, the associated clique types, and the clique functions, $V_c(\omega)$'s, in the Gibbs distribution are adequately selected to represent spatially meaningful continuity according to the application. In this study, the clique function to test with for the proposed model is selected as follows:

For single pixel clique

$V_c(\omega) = \alpha_k$ if x_i in c is equal to k

For clique of "dir" type with two pixels

where the parameters assigned to each clique type are defined according to 4 types of direction

$$V_c(\omega) = \begin{cases} -\beta_{dir} & \text{if all } x_i \text{ in the clique } c \text{ are equal} \\ \beta_{dir} & \text{otherwise} \end{cases} \quad (2)$$

between pixels (↔ ↕ ↗ ↘). Parameters, α and β's, control the percentage of each class and clustering of pixels in each direction, respectively.

As shown in Eq. (1), a GRF describes the global properties of an image in terms of local properties. With this idea, many relaxation-type algorithms based on a Monte Carlo computation theory are developed based on a Bayesian framework such as Gibbs Sample (Geman and Geman, 1984). It provides a coherent approach for processing of images, but the computational cost for processing of large images is expensive. The time required to simulate increasing larger images is not linear. For example, processing of a 1024×1024 image requires greater than 4 times the time for the simulation of a 512×512 image, because it requires more iterations for convergence. The computational complexity for an $N \times N$ image is defined as in $O(N^2T(N))$, where $T(N)$ is a function of the number of iterations required for convergence. $T(N)$ is unknown, but at least is a linear function when N is small.

A multiresolution MRF framework defined in a stack of images would support efficient algorithms that start processing of images at a coarse resolution and then progressively refine them to finer resolutions, which would allow global features such as region information to propagate quickly. The problem in developing a multiresolution MRF work is to provide consistent model descriptions for MRFs at multiple resolutions since a simple resolution transformation loses the Markov property. Therefore, approximation of non-Markov fields by MRFs is needed to model the images as MRFs at multiple resolutions. In this study, a Markov approximation for a coarser resolution field is obtained by relating the parameters corresponding to the MRF approximation at coarse resolutions to the parameters at the fine resolution such that the interaction between distant pixels is decimated as in the approach of normalized group study (Gidas, 1989).

For the multiresolution approach, two major steps are to be considered: resolution transformation and projection. During the resolution transformation, a finite sequence of coarser and coarser grids, $L^{(0)} \rightarrow L^{(1)} \rightarrow L^{(2)}$ ... are iteratively generated, where each $L^{(l)}$ is obtained from the previous grid $L^{(l-1)}$ by coarsening it. Here, reduced-resolution versions of a given image are defined in an exponentially tapering 'pyramid' of arrays of size 2^{N-1} by 2^{N-1}, 2^{N-2} by 2^{N-2}, ... , 2 by 2. The internal communication in this structure is comprised of two different kinds of links: intra-level link (horizontal) and inter-level links (vertical). In inter-level links, the corresponding sequence of coarse grid energy functions is assumed to be $U^{(0)} \rightarrow U_H^{(1)} \rightarrow U_H^{(2)}$... depending on the positive parameter H, which reflects the decimation of bonding energy between the distant pixels. A coarse grid $U_H^{(l)}$ is obtained iteratively from the posterior U as follows: For the selected conditional probability $P^{(l)}(x^{(l)} \mid x^{(l-1)})$, which is the probability of the gray level configuration of $L^{(l)}$, $x^{(l)}$ given the gray level configuration of $L^{(l-1)}$, $x^{(l-1)}$ and a selected positive constant H, $U_H^{(l)}$ is iteratively computed from $U_H^{(l)} = H^{-1}U^{(0)}$ and the energy function at the coarse levels is obtained by

$$e^{-U_H^{(l)}(x^{(l)})} = \sum_{\{x^{(l-1)}\}} e^{-U_H^{(l-1)}(x^{(l-1)})} P^{(l)}(x^{(l)}|x^{(l-1)}) \qquad (3)$$

which represents the reduced intra-level commutation between pixels at increasing distance.

Fig. 1 shows examples of sampling schemes, where the remaining pixels are denoted by the black dot and the non-sampled pixel by the empty dot. Actually, a special sampling schemes where a random field subsampled at the coarse resolution still preserves the Markovian property are very limited. For example, in the scheme (a) presented in Fig. 1, a subsampled random field at the level 1 is still Markovian field for the first order neighborhood since the nonsampled points are independent (Jeng, 1989). However, the resulting coarse image doesn't correspond to a uniform spatial grid and has the neighboring system changed. In this study, the sampling scheme (b) was selected . This sampling scheme can also satisfy the conditional probability constraint on the projection process. The resolution transformation is defined as

$$X_{i,j}^{(l)} = X_{2i,\,2j}^{(l-1)}$$

and so,

$$X_{i,j}^{(l)} = X_{2_i^l,\,2_j^l}^{(0)}.$$

The projection step starts at the top level and proceeds iteratively downward to obtain the bottom level image: first, an image is processed at a level l and transmited to the level l–1, then successively to the full resolution image $L^{(0)}$. The projection process is constrained by the realization $x^{(l)}$ via the equation

$$P^{(l)}(x^{(l)}|x^{(l-1)}) = \max_{\{x^{(l)}\}} P^{(l)}(x^{(l)}|x^{(l-1)}). \qquad (4)$$

For the discrete-valued field, a special, but useful, class of conditional probability can be

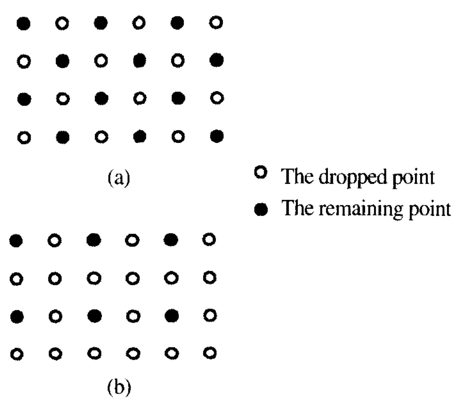


Fig. 1 Examples of sampling type resolution transformation.

defined in the following. If $L^{(l)} \subset L^{(l-1)}$, the conditional probability can be selected as follows:

$$P^{(l)}\left(x^{(l)} | x^{(l-1)}\right) = \prod_{i \in L^{(l)}} \delta_{x_i^{(l)} x_j^{(l-1)}} \quad (5)$$

where

$$\begin{cases} \delta_{ij} = 1 & \text{if } i = j \\ \delta_{ij} = 0 & \text{if } i \neq j \end{cases}$$

This allows the configuration on $L^{(l-1)}$ induced by (4) to be easily determined.

The simplest approach to guarantee this is that the son nodes corresponding to the father node are all assigned the same value as that of the father node,

$$X_{i,j}^{(l-1)} = X_{[i/2],[j/2]}^{(l)}$$

where $[x]$ takes the integer value of x. This tends to simplify the boundaries. As an alternative approach, a probabilistic decision considering a reduced set of neighbors can be used. First, according to the constraint (4), the value of the father node is assigned to the son node sampled in the resolution transformation, that is,

$$X_{i,j}^{(l)} = X_{2i,2j}^{(l-1)}.$$

Then, the remaining son nodes are determined by the conditional probability given the reduced neighbors. Thus, as an example of reduced neighbors, the class value of a node in the boundary area whose father node is adjacent to a node of different class is decided by the joint probability,

$$\begin{aligned} &P\left(x_{i,j+1}^{(l-1)}, x_{i+1,j}^{(l-1)}, x_{i+1,j+1}^{(l-1)} | x_{i,j}^{(l-1)}, x_{i+2,j}^{(l-1)}, x_{i,j+2}^{(l-1)}\right) = P\left(x_{i,j+1}^{(l-1)} | x_{i,j}^{(l-1)}, \right. \\ &\left. x_{i,j+2}^{(l-1)}\right) P\left(x_{i+1,j}^{(l-1)} | x_{i,j}^{(l-1)}, x_{i+2,j}^{(l-1)}\right) \times \\ &P\left(x_{i+1,j+1}^{(l-1)} | x_{i,j}^{(l-1)}, x_{i+1,j}^{(l-1)}\right) \quad (6) \end{aligned}$$

where each probability is estimated from the data.

3. Bayesian Estimation of Parameters for Large Images

As described before, MRF models a *priori* beliefs about the continuity of image features and the global features are defined through the local interactions. With this concept, a Bayesian approach is suggested as an alternative estimation method to reasonably reduce the computational burden in estimation of the parameters of large images.

In the MRF model, each pixel is assumed to have a local interaction with the specified neighbors that is independent of the location in the image, but depends only on the intensity values of the given pixel and its neighbors. This local interaction is usually assumed to be constant over the entire image depending on the pixel values and is defined through the parameters of the Gibbs distribution. This implies that in some sense, the values of the parameters reflect the averaged values of local interactions between neighboring pixels across the entire image.

This idea can allow a Bayesian approach for estimating parameters of large images. Subsets of the image, whose sizes are not too small, can be reasonably assumed to have the same spatial relationships as that of the entire image, if the interaction is actually constant across the image. That is, when subsets of the same size or different sizes are extracted from the image, each estimated parameter follows the parametric rule which describes the whole image.

Suppose that the parameter set estimated from each subset is $\mathbf{q}_i$. Then, $f(\mathbf{q})$ can be obtained from these $\{\mathbf{q}_i\}$ as a parametric function of $\mathbf{q}$ which describes the local interaction over the entire image. When the parameter set is estimated directly from the whole image, it is considered to be a deterministic value describing the local interaction of the scene. In the Bayesian approach investigated in this study, the parameter space which describes the local interaction observed in the entire image is estimated as a parametric function. A possible parameter set $\mathbf{q}^*$ which describes the given image is generated from this $f(\mathbf{q})$ and then, the images are processed with the Gibbs distribution $P(X \mid \mathbf{q}^*)$. Actually, some other parameter space may share similar characteristics and the estimation scheme may correspond to some point of this alternative parameter space. In this study, the multivariate normal distribution is utilized to obtain $f(\mathbf{q})$: The p-dimensional multivariate normal density with mean vector μ and positive definite covariance matrix Σ is of the form,

$$f(\mathbf{q}|\theta) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\exp(-\frac{1}{2}(q-\mu)^{\prime}\,\Sigma^{-1}(q-\mu)) \quad (7)$$

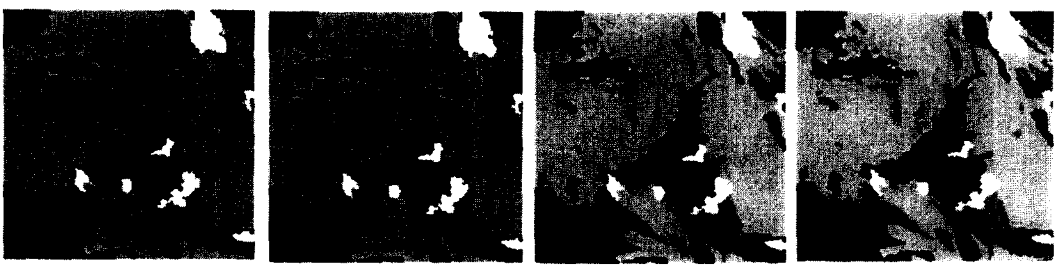
This is denoted as $N_p(\mu, \Sigma)$ where p is the number of parameters of the model. The parameter sets estimated from n subsets of the image, $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$, represent a random sample from the distribution for the parameter space corresponding to the spatial relationship between the pixels. Maximization of the likelihood function or log-likelihood function with respective to μ and Σ, for given the data $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$, yields the maximum likelihood estimate.

4. Results

1) Utilization of Multiresolution Framework

In this section, the utilization of the pyramid structure was explored with a discrete Markov random field model proposed in Section II. Even though the multiresolution MRF paradigm has difficulty with the consistent model description for MRFs at multiple resolutions as described before, it is still an attractive framework since it provides the method to process the image considering the characteristics at different scales.

Fig. 2 compares the original full resolution image and reconstructed images from a coarse image. Fig. 2 (a) is a 1024×1024 subset obtained from the classified image of the TM data covering the Great Victoria Desert, Australia. This image was considered to be $L^{(0)}$ and then reduced in resolution to $L^{(2)}$ (256×256 in size) by the sampling scheme $X_{i,j}^{(l)} = X_{2^l i, 2^l j}^{(0)}$, which is shown in Fig. 2 (b). Then, $L^{(0)}$ was reconstructed from the coarse image $L^{(2)}$ according to the constraint described in the Section II. Fig. 2 (c) was obtained by the simple projection that the son nodes are all



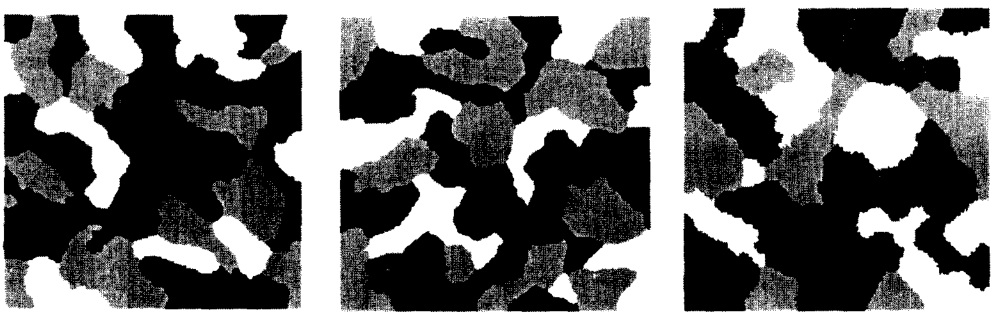
(a) (b) (c) (d)

Fig. 2. Comparison of the original image and reconstructed image: (a) 1024x1024 original image (b) 256x256 coarse image ($L^{(2)}$) (c) Reconstructed full resolution image by simple projection (d) Reconstructed full resolution image by probabilistic projection

assigned the same value as that of the father node, where the blocky boundaries are shown. Fig. 2 (d) is an image reconstructed by the probabilistic projection considering the neighboring pixels. Even though the reconstructed full resolution images have a difficulty with recovery of the exact boundary information, this multiresolution framework could be still useful since they still have the global features of region information. Then, for the boundary variation, the fuzzy approach can be applied (Jung, 1997).

In the next example, the multiresolution MRF framework is evaluated with the image simulated at the full resolution and the image simulated at a coarse resolution with the coarse grid energy function and then refined from it to the full resolution. Fig. 3 (a) is a 1024×1024 image simulated with β=(1.0, 1.0, 1.0, 1.0), whose estimated parameter set is with β=(1.03, 1.07, 1.07, 1.01). Fig. 3 (b) was obtained by simulation with the estimated parameter set at the full resolution, whose estimated values are β=(0.99, 1.04, 0.94, 0.96). Fig. 3 (c) was obtained by simulation at a coarse resolution, Level 2, with the estimated parameter set β=(1.03, 1.07, 1.07, 1.01) and H=1.1 which was estimated from the coarse grid energy function and projection to successively finer resolutions. The estimated values of the resulting full resolution image are β=(0.87, 0.91, 0.91, 0.97). Here, it should be noted that the selected parameter estimation method, the Least Square Error method (LSQR) (Derin and Elliot, 1987), might be one possible factor for the difference in the estimated values from the original values since the LSQR method uses the number of the 3x3 configurations existent in the scene for computing the joint probability of a pixel and its neighbors. That is, the change occurred around boundary affects the estimation of parameters using the LSQR method, even though the images all have similar global features.

Next, the time for the full resolution process and multiresolution framework is compared. The total CPU time to obtain 1024× 1024 full resolution images, (a) and (b) on a Sun SparcUltra 1-170E was 9.6 hours and for (c) 4.3 minutes. Since (c) has global characteristics which is visually similar to (a) and requires much less computational time, the multiresolution framework could be extremely useful.



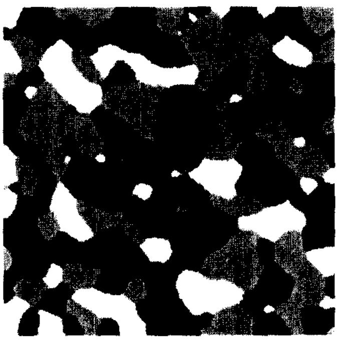
(a) (b) (c)

Fig. 3. Comparison of images simulated at full resolution and at a coarse resolution using projection scheme: (a) 1024x1024 original image simulated with β=(1.0 1.0 1.0 1.0) (b) Image simulated at full resolution (c) Image refined from the image simulated with the estimated parameters at Level 2

2) Bayesian Approach in Parameter Estimation for Large Images

The Bayesian approach in parameter estimation was investigated as a means of reducing the computational burden for large images. The method was implemented and applied to simulated scenes.

First, Fig. 4 (a) is 1024 × 1024 image simulated with parameters α_i = 0.0 and $\beta_{i,j}$ = 1.0 and 1000 iterations. Fig. 4 (b) correspond to the four nonoverlapping subsets of 512 × 512 in size. Table 1 compares the estimated parameters of the entire image and its four subsets. The distribution of the parameters, $f(\beta)$, can be obtained from the estimated parameter set $\{\beta_i\}$ of the four subsets. In this example, a univariate normal distribution function for each β_i, i = *hor, ver, ne, nw* was used to represent $f(\beta)$ instead of the multivariate normal



(a) 1024 × 1024 image

Table 1. Comparison of the estimated parameter sets of the whole image and its four subset images

	β_{hor}	β_{hv}	β_{d1}	β_{d2}
Image (a)	1.12	1.16	1.08	1.03
subset (1)	1.00	1.05	0.91	0.98
subset (2)	0.96	0.89	1.01	0.92
subset (3)	0.96	0.98	0.91	1.02
subset (4)	1.08	0.97	1.00	1.00

distribution function because the estimated covariance between the β_i's is extremely small. The representative functions, $f(\beta)$, are:

$\beta_{hor} = N(1.00, 0.002)$

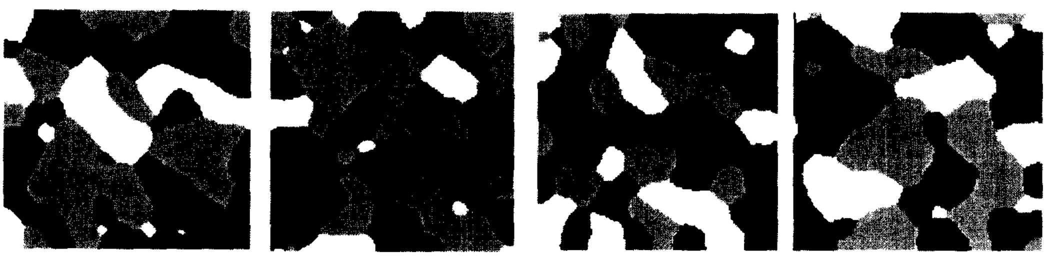
$\beta_{ver} = N(0.98, 0.002)$

$\beta ne = N(0.96, 0.002)$

$\beta_{nw} = N(0.98, 0.001)$

Bayesian estimators which can possibly describe the local interaction over the image can be generated from these functions. Here are some examples of realization generated by these $f(\beta)$'s,

The images generated with the parameters estimated directly from the whole image and the parameters obtained from these functions are shown in Fig. 4 for the sake of comparison. Fig. 4 (a) was generated with parameters estimated from the image in Fig. 4 (a), β = (1.12,1.16,1.08,1.03). The parameters estimated from this simulated image are β = (1.11,1.21,1.15,1.13). Fig. 4 (b) is generated as one realization of parameters obtained from the distribution, $f(\beta)$'s, β = (1.02,1.08,1.08,1.01). The



(b) 512 × 512 nonoverlapping subsets of image (a)

Fig. 4. Example: a full image and its four subset images

parameters estimated from this image are β = (1.13,1.15,1.03,1.08). The two images have visually similar characteristics. All images were obtained with 1000 iterations. Here, it should be noted that one reason for difference in estimated parameters from the actual parameters might be due to the fact that only 1000 iterations were performed. Using this methodology, the parameters of large images can be obtained as a possible parameter set to describe the local interaction over the whole image.

The Bayesian parameter estimation method was applied to the TM classmap image shown in Fig 3. Table 2 lists the estimated parameters of the entire image and its four non-overlapping 512× 512 subsets. From these estimated parameters of the subsets, the parametric function of β which describes the local interaction over the entire image was obtained using multivariate normal distribution. The resulting functions, $f(\beta)$, are defined with

$$\mu = (0.987, 0.955, 0.680, 0.822)$$

$$\Sigma = \begin{bmatrix} 0.014 : 0.004 : 0.002 : 0.010 \\ 0.004 : 0.006 : 0.008 : 0.001 \\ 0.002 : 0.008 : 0.014 : 0.002 \\ 0.010 : 0.001 : 0.002 : 0.034 \end{bmatrix}$$

Table 2. Comparison of the estimated parameter sets of the original image and its four subset images

	β_{hor}	β_{hv}	β_{d1}	β_{d2}
entire class map	0.91	0.91	0.61	0.76
subset (1)	0.89	0.91	0.54	0.63
subset (2)	1.10	0.89	0.71	0.90
subset (3)	0.85	0.93	0.61	0.67
subset (4)	1.11	1.09	0.86	1.09

Here are some examples of realization generated by the parametric function:

β= (0.96, 1.06, 0.73, 0.70), (0.90, 0.97, 0.69, 0.71), (0.89, 0.98, 0.70, 0.75) (1.02, 0.98, 0.66, 0.85), (0.85, 0.97, 0.67, 0.75)

The parameters estimated from the full image are statistically considered as one realization of the function based on the test at α = 0.05.

5. Conclusions

A range of remotely sensed data from different sensors is widely available in many applications. The Markov Random Field (MRF) models characterize the stochastic properties observed in the scene such as a large scale characteristic of a

(a) (b)

Fig. 5. Illustration of the effect of Bayesian estimation (a) 1024x1024 image simulated with parameters estimated from the whole image: Estimated parameters β = (1.11,1.21,1.15,1.13) (b) 1024x1024 image simulated with Bayesian estimators: Estimated parameters β = (1.13,1.15,1.03,1.08)

scene like region formation, or continuous random variation and is usefully employed in many problems providing a theoretically robust approach. However, computational cost is getting very expensive as image size increases. For instance, Gibbs Sampler or the exchange algorithm based on the Metropolis algorithm requires much computational time given the large size of such data sets.

In this research, efficent methodology to reduce the computational burden for large images was investigated. Multiresolution MRF framework would support efficient algorithm that initiate processing images at a coarse resolution and then progressively refine them to finer resolutions, which results in reducing the computational cost. As shown in Section VI, where multilevel framework was tested with the simulated data, it is noted that it requires much less computational time and could be a very useful approach. Here, a non-Markov field at a coarser resolution field was approximated to a MRF by coarsening the energy function to reflect the decimated local interaction. In addition, the Bayesian approach was explored as a reasonable alternative method in parameter estimation of large images based on the idea that the global features are defined in term of local interaction in Markov Random Field. Actually, the parameter set estimated from a subset of an image follows the parametric rule which describes the whole image. In this study, the parameter space which describes the local interaction observed in the entire image was obtained as a parametric function.

References

Burt, P. J., T. Hong, and A. Rosenfeld, 1981. Segmentation and estimation of region properties through cooperative hierarchical computation, *IEEE Trans. on Systems, Man and Cybernetics*, 11(12).

Cressie, N. A. C., 1993. *Statistics for spatial Data*, John Wiley & sons.

Cross, G. R. and A. K. Jain, 1983. Markov Random Field texture model, *IEEE Trans. Pattern Anal. Machine Intell.*, 5: 25-39.

Derin, H. and H. Elliot, 1987. Modeling of segmentation of noisy and textured image using Gibbs random field, *IEEE Trans. Pattern Anal. Machine Intell.*, 9:39-55.

Dubes, R. C. and A. K. Jain, 1989. Random field models in image analysis, *Journal of Applied statistics*, 16:131-164.

Geman, S. and D. Geman, 1984. Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of Images, *IEEE Trans. Pattern Anal. Machine Intell.*, 6: 721-741.

Gidas, B., 1989. A Renormalization Group approach to Image Processing Problems, *IEEE Trans. Pattern Anal. Machine Intell.*, 11(2):164-180.

Jeng, F., Sampling of Markov Random Fields, 1993, *J. visual Communication and image representation*, 3(3): 225-229.

Jolin, J. M. and A Rosenfeld, 1994, *A pyramid Framework for Early Vision*, Kluwer Academic Publishers.

Jung, M., 1997. *Contextual Simulation of Landscape based on Remotely Sensed Data*, Ph.D. Dissertation, University of Texas at Austin.

Lakshmanan, Sridhar and Haluk Derin, 1980. Gaussian Markov Random Fields at Multiple Resolutions, *Random Fields Theory and Applications*, Academic Press:131-157.

Rosenfeld, A., 1984. Some Useful Properties of

Pyramids, *Multiresolution Image Processing and Analysis,* New York:2-5.

Solberg, A. H., 1999. Contextual Data Fusion Applied to Forest Map Revision, *IEEE Trans. Geoscience and Remote Sensing,* 37(3):1234-1243.