Implementation a Different Segmentation Method as a Joint Prior Model to Reconstruction the Bayesian Image analysis

Salem M. Al. Gezeri\textsuperscript{1} & Kamla Elabd Othman\textsuperscript{2}

\textsuperscript{1}Department of Statistics, Faculty of Science, Benghazi University. \textsuperscript{2}Department of Statistics, Faculty of Arts and Science (Al-Marj), Benghazi University

Abstract

The basic concepts in the application of Bayesian approaches to image analysis have been introduced as advanced method. The Bayesian approach contains benefits in respect to image analysis and interpretation because it permits the use of prior knowledge concerning the situation under study. This paper use to investigate the application of some of the well-known procedures (determines number labels of image under several conditions like noise of image, resolution of image) for the Bayesian image analysis with segmentation as a joint prior model in order to estimate the Maximum Likelihood (ML). Markov random field with segmentation which resulted by mean of posterior (MP). This paper contains several sections. Firstly, includes introduction about the image analysis with, Bayes frame work and statistical background to Markov’s random field and its relationship through Markov Chain Monte Carlo. Secondly, section, which directly addresses solutions by using a principle of segmentation, which is a representation in threshold, is simply type introduced. Thirdly, presents the description of Experiment of Segmentation by depend on histogram also study of the factor (one prior) and the same model by adding segmentation (joint prior) based on techniques presented in the previously section are discussed . Fourthly, this section contains on the second prior implementation and simulation by using phantom data (Castle from south East Asia) also steps of estimation. Finally, result of experiment as well as estimation and summary.

Keywords: Bayesian image analysis, Markov random field, Posterior Mean, image segmentation, joint prior model
تطبيقات أسلوب التجزئة بطريقة مختلفة كنموذج قبلي مشترك لإعادة بناء الصورة في نماذج بيز

د. سالم محمد القزيري 1 و كاملا العبد عثمان 2

(1) قسم الإحصاء، كلية العلوم، جامعة بنغازي
(2) قسم الإحصاء، كلية الآداب والعلوم المرج، جامعة بنغازي - ليبيا

الملخص:

Bayesian approach

في هذا البحث تم تقديم المفاهيم الأساسية لتطبيق أسلوب بيز لتحليل الصور، ولذا الأسلوب فرائد كبيرة بالنسبة لتحليل الصور كطريقة متقدمة لأنها تسهم باستخدام التوزيع القياسي أي المعلومات القيمية (Prior Distribution) المطلقة بالحالة المدروسة. وفي هذا البحث تم تحديد عدة إجراءات منها (عدد علامات للصور تحت شروط مثل التشويش، وعدد الوضوح) وتستخدم هذه النظرية للدراسة والتحقق من تطبيق بعض العمليات المعروفة لتحليل الصور بطريقة بيز للتجزئة أو التقيم للصور (Segmentation Image) كنموذج قبلي مشترك (Joint Prior) للموزع القياسي المشترك (Posterior Mean Distribution) كما أن التجزئة باستخدام نظرية حقل ماركوف أثبتت بنتائج باستخدام متوسط التوزيع البعدي (Posterior Mean Distribution) خلفية عن تحليل الصور، وأيضاً تعريف نظرية الحقل العشوائي وعلاقتها بسائل ماركوف وموتي كارلو في إطار نظرية بيز. ثانياً في هذا الجزء، الذي علّق الحول باستخدام أسسيات الانقسام أو التجزئة والتي تمثلت في استخدام اصطلاحات وهو (Threshold). ثالثاً، قدم وصف للتجزئة الانقسام بالاعتماد على المدرج التكراري وأيضا دراسة العامل (المعلومات القيمية) للموزع بالإضافة معلومات ثانية (Joint Prior) للنموذج وباختصار المعلومات ثانية التي تم تعريفها في الأجزاء السابقة وأيضاً مناقشة أسلوب المحاكاة لبيانات المتصلة عليها لقبة من جنوب شرق آسيا. رابعًا، تضمن هذا الجزء على تطبيق المعلومات الثانية وخطوات التقدير، بالإضافة إلى التقدير النهائي لشكل الصور. وأخيراً نتائج التجربة والخلاصة.
1.1 Introduction:

This paper is focusing on using statistical image analysis techniques to analyze data image by use segmentation technique. The aim of this paper is to study and develop statistical estimation methods of invert archaeology for underground data. The output (image data) generated by the resultant methods should provide an image of the site (information about location of image) presented in the form of buried features.

1.2 Bayes’ Framework

In the statistical inference the problem can be stated as having an unknown unobserved quantity of interest \( \theta \) assuming values in set denoted by \( \Theta \). \( \theta \) can be a scalar, a vector or a matrix. Until now the only relevant source of inference was provided by the probabilistic description of the observations. We will formalize the use of other sources of information in statistical inference. This will be defined the Bayesian approach to inference. Let \( H \) (for history) denote the initial available information about some parameter of interest. Assume further that this initial is expressed in probabilistic terms. It can then be summarized through \( p(\theta/H) \) and if the information content of \( H \) is enough for our inferential purpose, this is all that is needed. In this case the description of our uncertainty about \( \theta \) is complete. Depending upon the relevance of the question we are involved with may not be sufficient and, in this case, it must be increased. The main tool used in this case is experimentation. Assume a vector of random quantities \( X \) related to \( \theta \) can be observed providing further information about \( \theta \). (If \( X \) is not random then a functional relationship relating it should be exist. We can then evaluate the value of \( \theta \) and the problem is trivially solved). Before observing \( X \), we should know the sampling distribution of \( X \) given by \( p(X/\theta,H) \) where the dependence on , central to our argument, is clearly stated. After observing the value of, the amount of information we have about \( \theta \) has changed form \( H \) to \( H^* = H \cap (X = x) \). In fact, \( H^* \) is subset of \( H \) (a refinement on \( H \) was performed ), see (Bayes, 1763).

Now the information about \( \theta \) is summarized by \( p(\theta/X,H) \) and the only remaining question left is how to pass from :

\[
p(\theta/X,H) = \frac{p(\theta,X/H)}{p(X/H)} = \frac{p(X/\theta,H)}{p(X/H)} \quad (1.1)
\]

Where

\[
\int p(X,\theta/H)d\theta = p(X/H) \quad (1.2)
\]

The result presented above is known as Bayes’ theorem. This was introduced by the (Rev Thomas Bayes in two papers in (1763, 1764), Published after his death, as mentioned in (Barnett, 1973). As we can see the function in the denominator does not depend upon \( \theta \) and so, as far the quantity of interest \( \theta \) is concerned it is just a constant. Therefore, Bayes theorem can be rewritten in is more form that is usual

\[
P(\theta/X) \propto P(X/\theta) P(\theta) \quad (1.3)
\]
2.1 Segmentation principle

In many instances, Markov random fields provide good models and Metropolis-like Monte Carlo methods yield respectable facsimiles of the texture. We combine MRF texture models, for the pixel process, with an segmentation "texture label process" in order to segment and label a scene consisting of patches of natural textures. The image model thereby involves both a pixel process, of grey level intensities, and a label process, whose components identify the texture type of each picture element in the scene. (Derin and Elliot, 1987; Cohen and Cooper, 1987) By use of two image model. In this section, we define the specific models we will use for $X^P$, $X^L$ where these two components represent the pattern of gray values and the set of labels identifying the texture types present in the image respectively. This will be done by specifying the conditional density. In The image process comprises a pixel process and a label process, $X=\{X^P, X^L\}$. As usual, the pixels sites form an $N \times N$ (and $N$ are row column matrices) square lattice, say $S^P$. For each pixel site there is a matching label site, and thus graph associated with the image model has sites $S = S^P \cup S^L$ where $S^L$ is just like $S^P$ the element of $S^P$ and $S^L$ index the components of $X^P$ and $X^L$, respectively, so that $X^P = \{X^P_x\}$ and $X^L = \{X^L_s\}$, the range of the labels depend upon the actual number of texture in scene, thus assuming this number to be known a prior. Let $M$ be number of texture that are to be modeled. Then $X^L_s \in \{1, 2, ..., M\}$, $\forall s \in S^L$. We shall develop the image model by first assuming that the texture type is fixed say "$L" and constant over the scene. Conditional on $X^L_s \in \{1, 2, ..., M\}, \forall s \in S^L$. Then the $X^L_s$ representation grey levels:

$$
\pi(X^L) = \exp[-k \sum_{1 \leq s \leq c} n_{kl}]
$$

(2.1)

Where the $n_{kl}$ represents the number of distinct neighbors pairs colored $(k, l)$ and $k$ is a positive interaction parameter the process

$X^P$ Is a Gaussian Markov random field:

$$
\pi(X^P / X^L) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{1}{2\sigma^2} (X^P - D(X^L))^2\right\}
$$

(2.2)

The formulation of the joint prior which include of$X^P, X^L$ then the models of additive joint prior that is $\pi(X^P, X^L) = \pi(X^P / X^L) \pi(X^L)$

$$
= (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{1}{2\sigma^2} (X^P - D(X^L))^2\right\} \exp[-k \sum_{1 \leq s \leq c} n_{kl}]
$$

(2.3)

$$
= (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{1}{2\sigma^2} (X^P - D(X^L))^2\right\} - k \sum_{1 \leq s \leq c} n_{kl}
$$

(2.4)

The mathematical model finally above is representation by the Markov random fields and segmentation, which uses structure in rebuilding the (bad picture) such as algorithms.

2.1.1 Types of Segmentation

Thresholds, Edges finding, Points, and Lines detection.

2.2.2 Thresholding

Thresholding essentially involves turning a color or grayscale image into a 1-bit binary image. This is done by allocating every pixel in the image either black and white, depending on their value. The pivotal value that is used to decide whether any
given pixel is to be black or white is the threshold. Gray level thresholding is the simplest segmentation process. Many objects or image regions are characterized by constant reflectivity or light absorption of their surface thresholding is computationally inexpensive and fast. Thresholding can easily be done in real time using specialized hardware.

3.1 The description of Experiment of Segmentation

In the classical approaches the estimation of a unique source of information is representation by Maximum Likelihood Function which is used in estimated image analysis as an old approach not give good results when experimented by two the scientists: (Shepp and Vradi ,1982). Given an image which they wanted to estimate as a dependent figure image where each pixel in the image depended on another pixel, not a clear image. And in this method picture estimation was proved sample information became insufficient because the model is complicated it contains (Blur + Noise); therefore, it needs supplementary information such as (prior) MRF in applying Bayesian image analysis. If you want to prove that wherever available a new source of information wherever executed improves estimation of picture to adjoin previous information size (MLE). In this technique, we will consider the availability of a second source of information in addition to Markov Random Field (MRF) which is represented in the colors of a maps and this map will be invented to through the fundaments of segmentation.

\[
\pi(X^p, X^l) = \pi(X^p / X^l) \pi(X^l) \tag{3.1}
\]

\[
= [(2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{\frac{1}{2\sigma^2} (X^p - D(X^l))^2\right\} - k \sum_{1\leq c} n_{kl}] \tag{3.2}
\]

Now we describe briefly the estimation of posterior probabilities. Several approaches for finding the best approach to the best obtained images estimates to remove the blur and noise which wrap the true image featured. In the classical approaches we depend on the information from one source which is represented in Markov Random Field (MRF) and the cumulative sample information Maximum Likelihood Function (MLE) when given the posterior, in the equation it is clear that the estimates technique in the dominator where the identification of the old value by value which is picked up from simulation (MCMC) . The principle idea to Markov Chain Monte Carlo algorithm (MCMC) is initially proposed by matrix to picture and the estimation of locations one by one on the basic proposal of a new value and accepted if executed to improvement of estimation and rejected vice versa. For a pixel which we want to estimate and new value by value we want to compare with old values for the same pixel it follows that:

\[
K = \frac{\text{Likelihood +Prior(MRF)(New)}}{\text{Likelihood +Prior(MRF)(Old)}} \tag{3.3}
\]

where K is a symbol of posterior if K >1 then that means the new value has a larger chance than the old value because the good quality of the posterior gives good results , If K< 1,so we make a selection of the value from Uniform Distribution that is drawn random sample from (0 -1) we cast a coin if the selected value >1 we choose the new value if selected value <1 we choose the (value ) old value and so on.
This technique is named the Metropolis Hastings Algorithm. In this technique, we will consider the availability of a second source of information in addition to Markov Random Field (MRF) which is represented in colors map, and this map will be concocted through the fundament of segmentation. For each value in the matrix of picture proximal color in colors of map help to arrive quickly to the real values. There are many approaches to the procedure of segmentation but the simple one is threshold. The histogram numbers of labels or numbers of colors will be chosen from the picture from which it takes its color.

3.1.1 Histograms

Histogram normalization or contrast enhancement involves rescaling the grey level values so that the full range of values depend on the true color of image and also involves many colors in the image. Means in general, shape the histogram through subdivision labels to multi-color and concrete little regions that means if the blocks are big then the bias is large and the variance is small.

![Histograms](image.png)

Figure (1) histograms for (5 - 10 - 20 -40) labels by R 1.6.1 code

The most crucial issue is the choice of the numbers of labels or the numbers of colors (L). The main question to be answered is to ask how wide the local neighborhood should be that the local approximation will be a good one. If one takes a very small number of the labels, the modeling bias will be small since the number of data points falling in this local neighborhood is also small but the variance will be large. On the other hand, taking large numbers of the labels will create a large modeling bias depending on the underlying function. This means the bandwidth governs the complexity of the model through the tradeoff between quantities of bias and variance. Improvement of the image (Warp image) and increasing estimates quality, and then the other source is segmentation which makes the segmentation image to regions color different and homogenous in shape. The segmentation corresponds to Gaussian Distribution as follows: 

\[ \frac{1}{Z} e^{-\frac{1}{2}(X^p - X^l)^2} \]

where Z is a constant and \(X^p\) is the pixel value, \(X^l\) the label value, preservative to first prior (MRF) then we think position of pixel depends on label place measured by difference \((X^p - X^l)(X^p - X^l)\) if variation to
both is small that is most excellent and this means that pixel position belonging to same label is good in the equation (5.2) obviously in this equation:

\[
K = \frac{[\text{Likelihood } \times (\text{Prior}(MRF)(\text{Segmentation}))](\text{New})}{[\text{Likelihood } \times (\text{Prior}(MRF)(\text{Segmentation}))](\text{Old})}
\]

If \( K > 1 \) then second prior is stronger (likelihood function and main prior ) so this improves posterior probability If \( K < 1 \) then the difference is large and we ignore the proposed value for pixel and keep the old proposed value until given the best value that agrees with value label, the proposed value (new) which has a large chance and accepts Markov Random Field with Segment not to be discredited improvement value to estimation. If \( K = 1 \) it has no diversion in old and new value for pixel, one should be chosen from one of them. Abiding the principle problem, numbers of chromaticity indicators (labels) which must be dependent, therefore, the researcher depends on procedure experiments by several users’ different numbers to chromaticity indicators and from the study of factors which affect it and the selection and choice of the best between them. In the experiment we take on the truth which decreases numbers of color to six levels. At the same time we make segmentation procured on the image degradation which represented by shareholding as a simple type of the segmentation by using the manual technique of depiction histogram. Where we take the truth image and reduce it into true six images where the image consists of six colors by reducing the color image moving average approach six time and measure variance in each image from starting image in the end image notice decrease in variance we obtain smallest variance and can able this to clear in image feature. In the below figure is shown the six images as follows with graphic (plot the variance).

![Figure (2) Truth image (Phantom data)](image-url)
Figure(3) six images with graphic (plot the variance)

In addition, the blur and noise for each sixth true image we try in every one time to determine value of noise by 0.03 for all pictures. And do not show the image feature in additive first, until it starts to show features in last image. As following the images shown which are obtained from the data and application of the segmentation technique by using the histogram on each data image by R ,1.6.1 code to levels four (low, medium , high and very high labels).

3.2 The Study of the Factor

To clarify the quality of studying the factor during experiment simulation when dealing with color images, and how to measure quantitatively the color differences between any two arbitrary colors. Experimental evidence suggests that taking the original picture made by extracting six pictures where bottom by suing (moving average approach) decreased picture color that put the principle values for original picture, create stepwise in each stage measure the variance in each one time until obtaining final overall on small variance to sixth picture, through FORTRAN code4.0 through effect. Point spread function (PSF) restored one time in R code1.6.1 by accomplishment additive noise on three levels (0.1 , 0.3, 0.5) where the picture which has clear parameters not effect noise such remainder other picture which has less clear parameters; this means less variance whose identified data ,and by manual technique numbers of labels or numbers of colors to picture for sixth data picture and partial it to several levels (Low, Medium ,High ) and study the best appropriate level for a good estimation picture.

4.1 The second prior implementation

There are a couple of things that have an effect on the use of segmentation such as other information source and this depends on the normal figure of the color in a given picture for the data with power and weak location it, and since asymptotical in light intensity as well. There is another factor which is very important and should be taken into consideration; that is the number of labels which is chaffing to present as a color
map expression about a second source information (second prior). Before beginning in chaffy process, the joint model form to prior distribution and which composes segmentation absolute mullion it, and still makes reconnaissance experiment aim to expose the impact on all of them: the degree of derangement in given data, the degree of color variance, the number of labels which must be chaffed.

4.2 Simulation

Completed data accommodation mimics real picture, taken from one of the chronicle positions (South East Asia) which is like (a barn or a castle) from highest, it consists of a circle form that which exists inside many squares and blotches which is the interpretation secondary or part which does not concern calendar account, because to control in input color variance consist of truth where credence completes three different levels of color variation as following: low variance, normal variance, high variance, and this expression of by language image processing intensity contrast. In this part we consider three main factors affecting the segmentation information which are: truth resolution - noise and number of labels using different resolution variance (same noise & same number of labels).

Variance truth data variance truth data

Figure (4) representation by truth dependence three levels for variation as following:
2-Using different noise (same resolution & number of labels).

Figure (5) a representation different for the same numbers of labels and resolution image.

2. Thresholding Manually (medium noise) - Using a different number of labels (same resolution & same noise).

(a) Data (medium noise)
2 Labels (b) 10 labels (c)

15Labels (d) 25 labels(e)

With different levels (labels)

Figure (6) to show different labels manually by using R code

<table>
<thead>
<tr>
<th>No of column</th>
<th>No of rows</th>
<th>Depth</th>
<th>Extent</th>
<th>Inclination</th>
<th>Earth magnetic field (EMF)</th>
<th>Height of lower sensor</th>
<th>Height Of upper sensor</th>
<th>Smoothing parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>80</td>
<td>80</td>
<td>0.2</td>
<td>0.25</td>
<td>65</td>
<td>65000</td>
<td>0.2</td>
<td>0.7</td>
<td>3000</td>
</tr>
</tbody>
</table>

Table (1) the information about the site which choose to image

4.3 Steps of Estimation

At this point it is mannered in each level a set of labels or (partitions of colors) procedure which is addressed in the table (5.2), in addition to the site in the sequence that has been given. Smoothing parameter is fixed in each experiment level of label three levels to the scaling parameter (σ ) and the variance and we have used the potential function sixth type which was proposed by (Allum ,1997). This reason agrees excellent results and types of the relationship between the pixel and it neighbours of the second order of MRF which is represented through Gaussian's distribution. Since the second order of MRF contains a whole shape that includes a lot of information about the shape from side pixels under study and keep on the bounding and edging of the shape until giving a lot of parameters and features to image. In this experiment we have used the FORTRAN code 4. to estimate the unidentified truth ,the number of iterations in experiment ,and iterations will burn period for sample mean (posterior mean ).
5.1 Results of experiment

In this experiment, the smoothing parameter is a fixed selection and while a small number of colors are fifth labels with (low level) scaling parameter ($\sigma$) noise, with variance ($\sigma^2$) and the same label with intermediate (variance, scaling parameter) ,also no difference in label but to high (variance, scaling parameter) selection tenth and twentieth of labels colors (medium) with low (variance, scaling parameter intermediate (variance, scaling parameter),high(variance, scaling parameter) selection fortieth of labels with low(variance, scaling parameter ), high (variance, scaling parameter ) following estimated the image ,calculating the mean square error of estimate:

$$MSE = \frac{\sum(y - E(y))^2}{N}$$

Where number of pixels , where they represent the data i.e. wrap image where represents known image ,where the number of a good label that has less mean square error (MSE) ,and we show as following in the table.

The subsequent table (1) demonstrates the summary for the results of experiment. This consists of the three types of noise, the three types of variance, the four types of labels of segmentation and the matching mean square error for all estimates of recalled previous. We note in the table value which has less mean square error at low noise with high variance to image also has high labels that value (0.00628).

<table>
<thead>
<tr>
<th>Noise</th>
<th>Variance</th>
<th>5 Labels</th>
<th>10 Labels</th>
<th>20 Labels</th>
<th>40 Labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Low</td>
<td>0.56681</td>
<td>0.20956</td>
<td>0.17922</td>
<td>0.13335</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>0.09816</td>
<td>0.05290</td>
<td>0.03477</td>
<td>0.01210</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.02343</td>
<td>0.01697</td>
<td>0.01305</td>
<td>0.00628</td>
</tr>
<tr>
<td>Medium</td>
<td>Low</td>
<td>0.51018</td>
<td>0.22691</td>
<td>0.13855</td>
<td>0.08757</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>0.1996</td>
<td>0.10063</td>
<td>0.07134</td>
<td>0.04455</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>0.14789</td>
<td>0.08091</td>
<td>0.05873</td>
<td>0.01269</td>
</tr>
</tbody>
</table>
5.2 Estimation
1-There are many factors that affect the using of segmentation, as a new source for information but the most important one is the truth variance that follows noise adversity and number of choice labels.

2-Notice in general where truth variance increases the value decreases of Mean Square Error (MSE).

3-Data noise adversity effects estimation fit for real pictures, when noise increases accessibility is impossible to get a good estimation until in real picture case one high variance.

4-In general, whenever we increase numbers of picked labels, the information volume will be available from second prior distribution model it is segmentation except here the base is not specified to actualize in a noised picture case plus noise.

5-We applied completed joint prior distribution model on simulated data except in case application on actual data not available with information about real picture variance and not on degree of noise additional, we can't compute mean square error (MSE) therefore, admonish to selection large numbers from labels if discerned to researcher the picture clearly which is representation to data either in case noise data admonish the researcher of chaffy numbers of ten labels.

5.3 Introduction
In the dissimilar and advanced statistical approaches and levels in research are implements to arrive at the major objectives in this research. Where these kinds of studies are advantaged by highly developed statistical techniques which help in good estimation command to give clear parameters picture to judge, we hope to reach this result.

5.4 Summary
The objective of this paper is to obtain the best image estimation by using advanced models, which have been proved a good solution of this way. Image segmentation is an essential preliminary step in most automatic pictorial pattern recognition and scene analysis applications. As indicated by the variety of examples presented in the previous chapters, the choice of one segmentation technique over another is dictated mostly by the special characteristics of the problem being considered these methods discussed in this chapter, although far from exhaustive are representations of techniques commonly used in practice.
1- As the information about unknown image increases the quality of estimation increases to

2- There is a direct relation between the variance of intensity of the feature and the quality of estimation.

3- There is a direct relation between levels of noise and the value of mean square error (MSE).

4- A number of labels should be adopted. This depends on the spatial variability of the data image, using different numbers of labels (either less or more) leads to bad estimation of the unknown truth.

5- The resulted images consist of a few regions and each has a unique constant gray value, furthermore, these few gray values are corrupted by additive Gaussian noise so the gray value observed at each pixel has a Gaussian distribution.

6- The segmentation technique has been referred to as the restoration problem because we are restoring the original pixel labels from noisy observation.

7- Until the image in this research consist of a single object on a background, the object and background will contain several gray values not just one.

8- The segmentation technique (new approach) is more complex here because the random variable in Y is not conditionally independent given X so the likelihood term cannot be decomposed as a simple product.

References.


Communicated by Mr. Price ,in a letter to John Canton,A.M.F.R.S." , 


