ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002, Melbourne, Australia.

Pixon-based Image Segmentation with Markov random fields

Faguo Yang and Tianzi Jiang, *IEEE Member*

National Laboratory of Pattern Recognition, Institute of Automation

Chinese Academy of Sciences, Beijing 100080, P. R. China

E-mails: {fgyang,jiangtz}@nlpr.ia.ac.cn

Abstract

Image segmentation is an essential processing step for many image analysis applications. So far, there does not exist a general method that is suitable for all the image analysis applications. In this paper, we propose a novel pixon-based multi-resolution method for image segmentation. The key idea to our approach is that a pixon-based image model is combined with a MRF model under a Bayesian framework. In our method, we put forward a new pixon definition scheme that is more suitable for image segmentation than the "fuzzy" pixon scheme. Experimental results demonstrate that our algorithm performs fairly well and the computational cost decreased sharply compared to the traditional MRF-based algorithm.

1. Introduction

Image segmentation is the process of segmenting an image into several disjoint regions whose characteristics such as intensity, color, texture etc, are similar. It is a key step in early vision problem and it has been widely investigated in the field of image processing.

A large number of segmentation techniques are available in the literature. But, there does not exist a general algorithm that can excellently perform the segmentation task even for all light intensity images, which is the most common type. Available segmentation techniques include thresholding [1], region growing [2], clustering [3], classifier [4], neural network-based approaches [5], deformable models [6], MRF model-based approaches [7]. Thresholding does not take into account the spatial characters of an image. This makes it sensitive to noise. The primary disadvantage of region growing is that it requires manual interaction to obtain the seed point. Clustering algorithms do not require training data, but they do require an initial segmentation. Classifier methods are pattern recognition techniques to partition a feature space derived from the image using data with known labels. Neural network represents a paradigm for machine learning and can be used in a variety of ways for image segmentation. Deformable models are physically motivated, model-based techniques for delineating region boundaries using closed parametric curves or surfaces that deform under the influence of internal and external forces.

The methodology of using MRF models to the problem of segmentation has emerged later and has created a lot of interest. Markov random fields have been, and are increasing being used to model a prior beliefs about the continuity of image features such as region labels, textures, edges and so on [8]. The main disadvantage of MRF-based methods is that the objective function associated with most nontrivial MRF problems is extremely nonconvex and as such the minimization problem is computationally very taxing. To reduce the computational burden, some approaches based on multi-resolution techniques have been reported [9][10]. The essence of MRF frame work on multi-resolution is that it starts processing images at a coarse resolution, and then progressively refine them to finer resolution.

In this paper, we propose a novel image segmentation method, in which "pixon" concept is incorporated. The "pixon" concept is firstly put forward by Piña and Puetter while doing astrophysical image restoration [11]. The essence of "pixon" concept is that the spatial scale at each site of the image varies according to the information embedded in the image. Like the photographic grain size, we would record the picture information with large photographic grains in portions of the image with coarse structure. We need only fine grains when we need to record fine spatial structure. This means that the picture information can be dealt with by using variable sized cells, with the cell sizes set so as to capture the spatial information present. When doing image restoration and image reconstruction, "fuzzy pixon", a pixon definition scheme using local convolution, is used. In "fuzzy-pixon" scheme, once the kernel function is selected, the shape of the pixons does not vary, and only the size can change.

But, “fuzzy pixon” is not suitable for image segmentation. We put forward a new pixon definition scheme, whose shape and size can vary simultaneously. By incorporating the “pixon” concept, the segmented image can be expressed by a concise model, so the computational cost can be discreased greatly.

The outline of the paper is as follows. In section 2, we give a brief description of the pixon concept and our pixon definition scheme. Section 3 devotes to the introduction to a Markov random field defined on our pixon-based image model. In section 4, we explain the proposed image segmentation scheme. Section 5 gives out the experiment result and the conclusions are listed in section 6.

2. Description of pixon model

In this section, we will give a concise review of the pixon concept and our pixon definition scheme. The main idea of the pixon concept is that at each point in the image there is a finest spatial scale of interest and that there is no information content below this scale. To further explain the pixon concept, let us think about photographic grain sizes. We would record the picture information with large photographic grains in portions of the image with coarse structure. We need only have fine grains when we need to record fine spatial structure. This means that the image information can be dealt with by using variable sized cells, with the cell sizes set so as to capture the spatial information present. In a word, the pixons are cells with variable shape and size, which locally define the resolution of the data. The size, shape and position of all pixons over an image are collected into a pixon map, which gives a multi-resolution description of the image with various spatial scales. Because different parts of an image often do not exhibit a uniform spatial resolution, the pixon map, as a multi-resolution language, suggests itself. It gives the finest spatial scale at each portions of the image.

2.1. Fuzzy pixon

When doing astrophysical image restoration and reconstruction, Piña and Puetter gave out the fuzzy pixon definition scheme [11]. The image is modeled by the local convolutions of a pseudo-image and a pixon map. To formalize the definition, at each pixel in the image, the image is written as

$$\begin{aligned} Y(t) &= (K \otimes I_p)(t) \\ &= \int K_t(t,v) I_p(v) dv \end{aligned} \qquad (1)$$

where $K_t(t,v)$ is the pixon kernel function. One commonly used functional form is a circularly symmetric pixon shape, which is truncated paraboloid. In the fuzzy pixon definition scheme, once the pixon kernel function is selected, a pixon is completely determined by its size. So, the pixon shape does not vary and only the pixon size change according to the observed image. By picking a set of pixons, a pixon map gives a multi-resolution description of the image.

2.2. Our pixon definition scheme

The key aspect of any pixon definition scheme is the ability to control the number of degrees of freedom used to model the image. In other word, the pixon definition scheme should give out a optimum scale description of the observed image. The fuzzy pixon definition scheme is suitable for image restoration and reconstruction, but it is not fit for image segmentation. Our pixon definition scheme can be described as follows:

$$I = \bigcup_{i=1}^{n} P_i \qquad (2)$$

where I is the pixon-based image model; n is the number of pixons; P_i is a pixon, which is made up of a set of connected pixels. The mean value of the connected pixels making up of the pixon is defined as the *pixon intensity*. Both the shape and size of each pixon vary according to the observed image. After the pixon-based image model is defined, the image segmentation problem is transferred to a problem that labels the pixons. To determine the set of pixons, we must obtain the number of the pixons, the shape of the pixons. By maximizing $P(I|X)$, where I is the set of pixons and X is the observed image, we can obtain an optimum pixon set. Because the number of the pixons and the size of the pixons can not be known in advance, it is difficult to optimize $P(I|X)$. In our current implementation, we use the method in [12] to obtain the pixon-based image model and it is expressed by a graph structure. An example is illustrated in Fig1. In Fig1a, P_1, P_2, P_3, P_4 and P_5 represent five pixons, which form the pixon-based image model. The pixon-based image model can be expressed by the graph structure illustrated in Fig1b.

3. Markov random field on pixon-based image model

Our Pixon-based image model is represented by a graph structure $G = (T, E)$, where T is the finite set of vertices of the graph, which represents the pixons in the image model, and E is the set of edges of the graph, which indicate that the two pixons are neighbors [13]. To define a Markov Random Field on the pixon-based image

model, we firstly introduce the following:

1) Neighbors: Two pixons are called neighbors if there is an edge connecting them in the graph structure, which means that there is at least a pair of neighboring pixels belonging to the two pixons respectively.
2) Neighborhood system: a neighborhood system on the graph structure G can be expressed as $N=\{N_t \mid t\in T\}$, where N_t is a set of pixons in T which are neighbors of the pixon t.
3) Clique: a clique is a fully connected sub-graph of G.

One of a finite set S of labels will be assigned to each pixon in the image model. Such an assignment will be called a configuration, which is denoted by ω. We denote by ω_t the value given to the pixon t by the configuration ω. Let ω_A represent the configuration ω restricted to the subset A of T, which is a configuration on the smaller graph restricting T to pixons of A. A Markov random field is defined if and only if the two following conditions are satisfied:

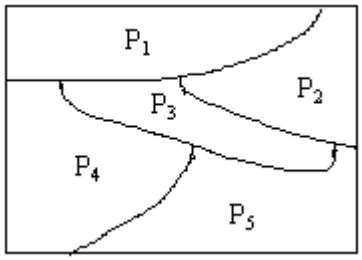
1) $P(\omega)>0$
2) $P(\omega_t \mid \omega_{T-t}) = P(\omega_t \mid \omega_{N_t})$

It can be proved that each Markov random field is a Gibbs distribution [14], which can be characterized by

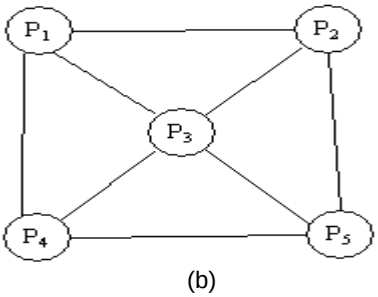
$$P(\omega)=\frac{1}{Z}e^{-U(\omega)} \tag{3}$$

where

$$Z=\sum_{\omega} e^{-U(\omega)} \tag{4}$$



(a)



(b)

Figure 1. (a) Pixon-based image model. (b) The corresponding graph representation.

$U(\omega)$ is an energy function defined on the set of all configurations ω by

$$U(\omega)=-\sum_{C} V_C(\omega) \tag{5}$$

Where C is a clique, and $V_C(\omega)$ is called clique potential associated with the clique C.

Thus, the joint probability $P(\omega)$ can be determined by specifying the clique potential functions $V_C(\omega)$. How to choose the forms and parameters of the potential functions for a specific problem is a major topic in MRF modeling.

4. Pixon-based image segmentation with MRFs

In this section, we will introduce our new pixon-based image segmentation algorithm in detail. Multi-resolution techniques are receiving considerable attention in various fields, such as image restoration, medical imaging and so on. In the field of image segmentation, there are a number of methods related to multi-resolution techniques. Pyramidal schemes and the multi-channel image modeling are two commonly used techniques. It can be seen that our pixon-based algorithm is a multi-resolution method.

4.1. New algorithm for image segmentation

Let X represent the observed image and Y the segmented image, which is a configuration ω of the graph structure representing the pixon-based image model. Let us assume that the noise in the observed image X is independent Gaussian white noise. In the Bayesian image segmentation framework, the segmented image is

$$Y^{*}=\arg\max_{Y} P(Y|X) \tag{6}$$

Where

$$P(Y|X)\propto P(X|Y)P(Y) \tag{7}$$

Since we assume the noise in the image is independent Gaussian white noise. And, if pixon P_j belongs to class K, then $g_{P_j}=u_k+n$, where g_{P_i} is the pixon intensity, u_k is the mean value of class K and n is the noise, $P(X|Y)$ can be written as follows:

$$P(X|Y)=\prod_{i=1}^{M}\prod_{P\in\Gamma_i}\frac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{(g_p-u_i)^2}{2\sigma_i^2}} \tag{8}$$

Where M is the number of class, Γ_i represents the i th class, g_p is the pixon intensity, which is the mean value

of the pixels making up of the pixon, u_i and σ_i is the mean gray level and the variance of the class i respectively. $P(X|Y)$ expresses fidelity to the observed image X.

The prior model is based on a Markov Random Field and assumes that the intensity profile is piecewise contiguous, and the adjacent pixons are likely to have similar constituents. The probability density of Y can be described as follows:

$$P(Y)=\frac{1}{Z}e^{-\frac{U(Y)}{T}} \tag{9}$$

Where $Z=\sum_{Y}e^{-\frac{U(Y)}{T}}$ is a normalizing constant and $U(Y)=\sum_{C}V_C(Y)$ is an energy function. We only consider the cliques including two pixons in this paper. So the clique potential $V_C(Y)$ is

$$V_C(Y)=V_{ij}(Y)=\frac{\eta_{ij}}{\left|g_{P_i}-g_{P_j}\right|} \tag{10}$$

if pixon P_i and pixon P_j have the same label, $\eta_{ij}=0$, or, $\eta_{ij}=1$. When the pixon intensities of pixon P_i and P_j are similar, and they have different labels, the clique potential associated with the pixon P_i and P_j will be large. This means that the pixons having similar pixon intensities have high probability to be labeled same.

Applying a logarithmic transformation to Eq. (7), we can obtain

$$\begin{aligned}P(Y|X)&\propto \ln[P(X|Y)]+\ln[P(Y)]\\&\propto -\sum_{i=1}^{M}\left(\sum_{P\in\Gamma_i}\frac{(g_p-u_i)^2}{2\sigma_i^2}+\ln\sigma_i\right)\\&\quad-\alpha\sum_{c\in C}V_c(Y)\end{aligned} \tag{11}$$

So, we define the energy function as follows:

$$\begin{aligned}E(Y)=&\sum_{i=1}^{M}\left(\sum_{P\in\Gamma_i}\frac{(g_p-u_i)^2}{2\sigma_i^2}+\ln\sigma_i\right)+\\&\alpha\sum_{c\in C}V_c(Y)\end{aligned} \tag{12}$$

Where Y represents the segmented image. Γ_i is the i th class; g_p is the pixon intensity of pixon P. u_i and σ_i are the mean value and variance of the i th class respectively. $V_c(Y)$ is the clique potential. α is the coefficient, which represents the trade-off between fidelity to the observed image and the smoothness of the segmented image. Through minimizing the energy function (12), we can obtain the segmented image.

4.2. Optimization

We assume that the class number M is known, and the parameter α is selected by experience. So, the parameters of the model denoted by θ only include the mean value u and the variance σ of each class. Given the observed image X and knowing the parameters θ, we need to find Y that minimizes the energy function $E(Y)$. Finding a global minimum of $E(Y)$, given all the configurations ω is a practically difficult task. The difficulty is further compounded by the fact that the model parameters necessary for the segmentation are not known. Hence, we attempt to find the optimal segmentation while estimating the optimal parameters for the segmentation. Therefore, the segmentation problem can be expressed in the following two-step processes:

$$Y^*=\arg\min_{Y}E(Y|X,\theta^*) \tag{13}$$

$$\theta^*=\arg\max_{\theta}P(X|Y^*,\theta) \tag{14}$$

In a word, the optimum segmentation and parameters for the models are needed to evaluate by the segmentation algorithm. The optimal parameters should predict image data with a maximum probability.

Given the segmented image and the observed image, we use maximum likelihood algorithm to estimate the mean and the variance of each class.

There are many local and global optimization techniques one can use to optimize the objective function, given the parameters of the model. The energy function is non-convex and has many local minima, so simulated annealing algorithm is often used to find the global minimum of the energy function. But the computational cost is very high. So we implement a sub-optimal algorithm to optimize the energy function. As long as a better initial solution is given, the sub-optimal algorithm also can find a satisfied solution. The proposed optimizing method can be stated as follows:

1) Initialize N_Γ, which represents the number of classes; NUM, which denotes the iteration number; $u_1,u_2\cdots u_{N_\Gamma}$ and $\sigma_1,\sigma_2,\cdots,\sigma_{N_\Gamma}$ according to an initial segmentation, which is obtained using K-means method; $i=0$, which is the iteration index.
2) Initialize our pixon-based image model, that is to assign a label k representing the k th class to each

pixon P, which minimize the expression $\left|g_p - u_k\right|$.

3) To each pixon P, we execute the following operation: For each pair of mean and variance $(u_k, \sigma_k)(k = 1, \cdots, N_\Gamma)$, we compute the value of the energy function. We assign pixon P a label, whose corresponding energy function value is minimal.
4) Repeat step 3 for L times. Then, according to the newly obtained segmented image, we re-estimate the mean value and variance of each class using a maximum likelihood algorithm.
5) $i = i + 1$. If $i < NUM$, go to step 3, otherwise stop.

5. Experimental results

In this section, we illustrate the efficiency and power of our approach and make a comparison with the traditional MRF-based method using three images viz.,

1) a synthetic noisy image with three labels and size 128 ×128 (see Fig2(a));
2) the Baboon image of size 512×512 with gray scale resolution of 256 gray scales (see Fig3 (a));
3) the Pepper image of size 512×512 with gray scale resolution of 256 scales (see Fig4 (a));

In order to see the pixon information of the images, we displayed the edge pictures of the pixons in Fig2 (b), Fig3 (b) and Fig4 (b) . In these images, white color represents the edges of the pixons. There are more pixons and the pixon size is smaller in the white parts of the edge images than those in the black parts. From these edge images of the pixons, we can see that there are more pixons and the pixon size is smaller in the parts of the original image, where there are more details. It demonstrates that the pixon-based image model is a multi-scale description of the image waiting for segmentation. Fig2 (c), Fig3 (c), and Fig4 (c) give out the segmentation results of the traditional MRF-based methods. The segmentation results of our approach are illustrated in Fig2 (d), Fig3 (d), and Fig4 (d). From these segmentation results, we can see that the segmentation results of our approach are comparable with that of the traditional MRF-based methods. Moreover, the computational burden of our approach is decreased greatly, which can be seen from table 1 and table 2. In table 1, the column I , column II and column Ⅲ are the number of pixons, the number of pixels and the ratio between them respectively. In table 2, the computational time required of the traditional MRF-based method and our approach, based on 550MHz processor are listed out by column I and column Ⅱ. And the column Ⅲ gives out the ratio between them. To be comparable, our approach and the traditional MRF-based method adopt the same optimization method described in section 4.2. From table 2, we can see that though it spends some computational time to create the pixon-based image model, the total computational time is decreased greatly. So, we can conclude that the computational cost decreased sharply when pixon concept is incorporated into the image segmentation process, and at the same time the segmentation results of our approach is comparable to those of the traditional MRF-based method. We can also see that images with more details need more pixons to describe it and therefore the pixon size is smaller.

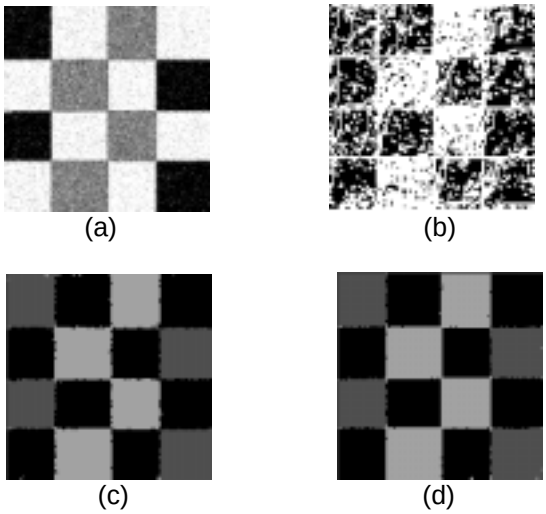


Figure 2. Segmentation results of synthetic image. (a) Original image. (b) Edge image of the pixons. (c) Traditional MRF method. (d) Our approach.

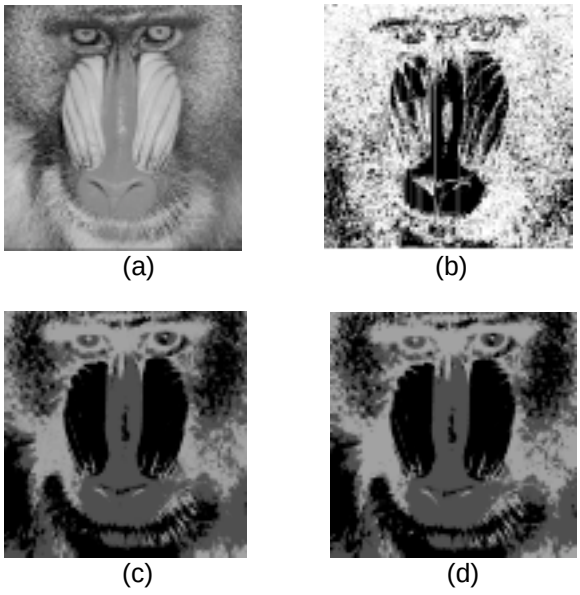
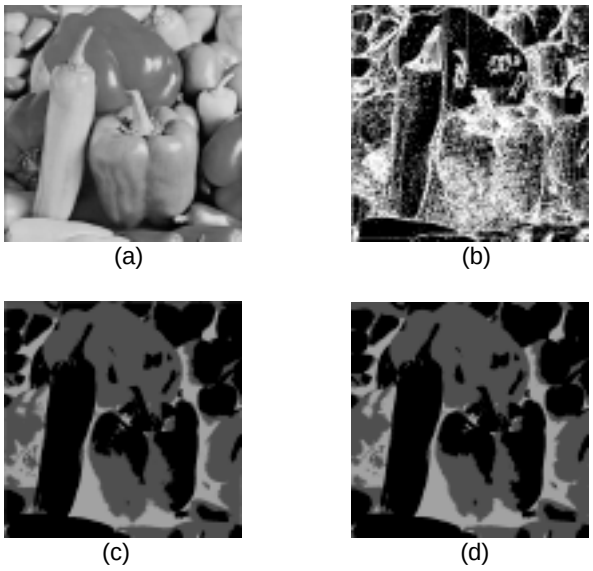


Figure 3. Segmentation results of the baboon image. (a) Original image. (b) Edge image of the pixons. (c) Traditional MRF method. (d) Our approach.

The parameters used in our experiments are as follows:

α in the energy function is 1500; the total iteration number NUM is 20; L that determine the frequency to re-estimate the parameters is 5, which means that after 5 iterations the mean value and the variance of each class are estimated again.



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

Figure 4. Segmentation results of the pepper image. (a) Original image. (b) Edge image of the pixons. (c) Traditional MRF method . (d) Our approachl.

Table 1. Pixon and pixel number

Images	I	II	III
Synthetic Image	3155	16384	19.3%
Baboon Image	57400	262144	21.9%
Pepper Image	16138	262144	6.2%
Cortex Image	3408	40000	8.52%

Table 2. Computational time

Images	I	II	III
Synthetic Image	8483	4096	48.2%
Baboon Image	133322	65904	49.4%
Pepper Image	132130	31494	23.8%
Cortex Image	15552	6159	39.6%

6. Conclusions and future work

In this paper, we proposed a novel image segmentation algorithm, which is based on the pixon concepts and the Markov random fields. In our algorithm, we put forward a new pixon definition scheme that is more suitable for image segmentation than the "fuzzy" pixon definition scheme. Because pixon concept is incorporated into our method, the computational cost decreased greatly compared to traditional MRFs-based image segmentation algorithm. Our algorithm has a high ability to resist noise, which can be seen from Fig2, because our method is also based on Markov random fields incorporating prior knowledge into the segmentation process. The results obtained indicate a promising direction for further research on pixon-based image segmentation. The experimental results demonstrate that our algorithm performs fairly well and the computational cost is decreased greatly.

In future, we will contunue our research on the following aspects: 1) Perfecting our pixon definition scheme for image segmentation; 2) Incorporating our pixon definition scheme into other traditional image segmentation method.

References

[1] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "A survey of thresholding techniques", *Comput. Vis. Graph. Im. Proc.*, vol. 41, 1988, pp. 233-260.

[2] J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms and applications in image segmentation", *Graph. Mod. Im. Proc.*, vol, 58,1996, pp. 246-261.

[3] G. B. Coleman and H. C. Andrews, "Image segmentation by clustering", *Proc. IEEE*, vol. 5, 1979, pp. 773-785.

[4] J. C. Bezdek, L. O. Hall, and L. P. Clarke, "Review of MR image segmentation techniques using pattern recognition", *Med. Phys.*, vol. 20, 1993, pp. 1033-1048.

[5] D. L. Vilarino, V. M. Brea, D. Cabello, and J. M. Pardo, " Discrete-time CNN for image segmentation by active contours", *Patt. Rec. Let.*, vol. 19, 1998, pp. 721-734.

[6] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours", *Int. J. Comp. Vision*., vol. 22, 1997, pp. 61-79.

[7] P. Andrey and P. Tarroux., "Unsupervised segmentation of Markov random field modeled textured images using selectionist relaxation", *IEEE Trans. Pattern Anal. Machine Intell., vol*. 20, 1998, pp. 252-262.

[8] Q. Lu and T. Jiang, "Pixon-based image denoising with Markov random fields", *Pattern Recognit.*, vol. 34, 2001, pp. 2029-2039.

[9] A. Sarkar, M. K. Biswas, and K. M. S. Sharma., "A simple unsupervised MRF model based image segmentation approach", *IEEE Trans. Image Processing*, vol. 9, 2000, pp. 801-812.

[10] I. Y. Kim and H. S. Yang, "A systematic way for region-based image segmentation based on Markov random field model", *Pattern Recognit. Lett.*, vol. 15, 1994, pp. 969-976.

[11] R. K. Piña and R. C. Pueter, "Bayesian image reconstruction: the pixon and optimal image modeling", *P. A. S. P.*, vol. 105, 1993, pp. 630-637.

[12] M. Suk, "A new image segmentation technique based on partition mode test", *Pattern Recognit.*, vol. 16, 1983, pp. 469-480.

[13] R. Kindermann, J. L. Snell, Markov Random Fields and their Application, Providence: American Mathematical Society, 1980.

[14] F. Spitzer, "Markov random fields and Gibbs ensembles", *Amer. Math.*, vol. 78, 1971, pp. 142-154.