Contagion-Driven Image Segmentation and Labeling

A. Banerjee P. Burlina

Center for Automation Research
University of Maryland
College Park, MD 20742-3275

F. Alajaji

Mathematics and Engineering
Queen's University
Kingston, ON K7L 3N6, Canada

Abstract

We propose a segmentation method based on Polya's urn model for contagious phenomena. An initial labeling of the pixel is obtained using a Maximum Likelihood (ML) estimate or the Nearest Mean Classifier (NMC), which are used to determine the initial composition of an urn representing the pixel. The resulting urns are then subjected to a modified urn sampling scheme mimicking the development of an infection to yield a segmentation of the image into homogeneous regions. Examples of the application of this scheme to the segmentation of synthetic texture images, Ultra-Wideband Synthetic Aperture Radar (UWB SAR) images and Magnetic Resonance Images (MRI) are provided.

1 Introduction

Image segmentation is a fundamental problem in computer vision which has been extensively studied. With the advent of new image modalities such as Synthetic Aperture Radar (SAR) and Magnetic Resonance Imaging (MRI), research into methods of segmentation has attracted renewed interest. We describe a segmentation method using contagion urn schemes that rely on modified versions of the Polya-Eggenberger sampling process [1, 11].. This biologically inspired sampling procedure was originally designed to model the development of contagious phenomena.

For our segmentation purposes, we model a scene as being composed of distinct, contiguous regions, each of which is described by constant or homogeneous attributes such as intensity or texture. An image is a corrupted version of an underlying piecewise smooth scene [10]. A natural approach to delineating the regions in an image is to statistically estimate the attributes of the regions and use the descriptions to differentiate between the regions.

Techniques such as Maximum Likelihood (ML) use these descriptions to divide the image into regions. However, the ML estimate of the pixel labels tends to produce a speckly segmentation, and thus smoothing algorithms such as relaxation labeling (RL) [13, 9] or simulated annealing (SA) [6, 8, 12] are applied. These methods make use of contextual information by relating neighboring pixels to form the estimate of the pixel label. They allow local information to propagate via iterative processing [9].

This paper models images using contagion urn processes. The idea behind this method is similar to that of RL; it iteratively propagates local information by contagion. The motivation for employing urn schemes is twofold: First, urn processes can generate Markov chains as well as MRFs [7]. Second, urn schemes are of particular interest because they provide a natural probabilistic representation for the image labels. Therefore, they constitute an attractive generative process for the underlying image regions which exhibit strong spatial dependencies. The spatial dependencies of the pixel labels are captured by the contagious behavior which promotes segmentation of the image into regions. The urn process is analogous to relaxation labeling algorithms, except that the urn process is not deterministic, but stochastic. The urn sampling scheme is also iterative and can be performed in parallel at each site or pixel of the image.

This paper is organized as follows: The initial NMC and ML segmentations are presented in Section 2. The contagion-based smoothing process is then described in Section 3. In Section 4, the stochastic properties of the resulting image process are discussed. In Section 5 the relationships between the urn sampling scheme, relaxation labeling, and simulated annealing are examined. Finally, experimental results on texture, SAR and MR images are shown in Section 6.

2 Initial Segmentation

When no a priori information on the image statistics is available, general clustering algorithms such as NMC are usually applied. In the NMC method, an initial arbitrary labeling is used from which centroids of the feature vectors of each class are computed. Next, all samples are reclassified to the cluster corresponding to the nearest mean, and the centroids are recomputed. This process is iterated until a stopping criterion is met [5].

On the other hand, when a stochastic model for the image can be justified, it is possible to apply ML segmentation [3]. Here, the conditional distribution of the image is assumed to be a correlated, multivariate Gaussian using non-overlapping 3×3 windows.

The above schemes do not capture the homogeneity of regions since the ML and NMC methods estimate each pixel label independently. Contextual information is not taken into account in either method.

This drawback is addressed by using MAP segmentation. By assuming an MRF model for the prior probabilities of the labels, contextual information is incorporated into the MAP test [8]. The MAP estimate generally requires computationally expensive optimization algorithms such as SA. Instead, we propose to replace the annealing step by an urn contagion process to model the spatial dependencies between neighboring pixels. Our motivation for employing an urn scheme lies in its ability to generate MRFs.

3 Urn Sampling with Contagion

In this section, the concept of temporal and spatial contagion for image segmentation is introduced, and the general urn sampling scheme for pixel classification is outlined.

Polya [11] introduced the following urn scheme as a model for the spread of a contagious disease through a population. An urn originally contains T balls, of which W are white and B are black ($T = W + B$). Successive draws from the urn are made; after each draw, $1 + \Delta$ ($\Delta > 0$) balls of the same color as was just drawn are returned to the urn. Let $\rho = W/T$ and $\delta = \Delta/T$. Define the binary process $\{Z_n\}_{n=0}^{\infty}$ as follows:

$$Z_n = \begin{cases} 0, & \text{if the } n^{th} \text{ ball drawn is white;} \\ 1, & \text{if the } n^{th} \text{ ball drawn is black.} \end{cases}$$

It can be shown that the process $\{Z_n\}$ is stationary and non-ergodic [4, 11]. The urn scheme has infinite memory, in the sense that each previously drawn ball has an equal effect on the outcome of the current draw.

The urn sampling scheme proposed in this paper incorporates both temporal and *spatial* contagion. Instead of representing an image by a finite lattice of pixels, we consider an image as a finite lattice of urns. In the single-urn sampling described above, the effect of each sample propagates through time. For the lattice of urns, the sampled ball at each iteration must depend not only on the composition of the pixel's urn, but also on the compositions of the neighboring urns to encourage contagious behavior. Thus, we need to allow for spatial interactions at each time instant by involving the urns of the neighboring pixels in the determination of the newly sampled ball.

The following presentation considers an L-ary labeling problem. Let $I_n = [p_n^{(i,j)}]$ be an L-ary label image of size $H \times K$, where $p_n^{(i,j)} \in \{1, ..., L\}$ is the label of pixel (i,j) at iteration n, $n = 0, 1, \ldots,$ $(i,j) \in \mathcal{I}$ where $\mathcal{I} : \{(i,j) : i = 0, \ldots, H-1; j = 0, \ldots, K-1\}$. We associate an urn $u_n^{(i,j)} : (B_{0,n}^{(i,j)}, B_{1,n}^{(i,j)}, \ldots, B_{L,n}^{(i,j)})$ with each pixel $p^{(i,j)}$ at time n, where $B_{l,n}^{(i,j)}$ is the number of balls of color l in the urn. With this representation we define a similarity function for each pixel as

$$m_F^l(p_n^{(i,j)}) = \frac{B_{l,n}^{(i,j)}}{\sum_{k=1}^{L} B_{k,n}^{(i,j)}}.$$

This can be interpreted as the probability that pixel $p^{(i,j)}$ belongs to class l.

The general class of algorithms for the contagion-based segmentation process will now be described. Initialization of the urn composition, provided by ML or NMC, is critical for the algorithm to converge to an appropriate segmentation. When the correlated Gaussian assumption is used for the ML estimate, the urn initialization proceeds in the following manner. The similarity of pixel $p^{(i,j)}$ to class l is determined by the Mahalanobis distance, $\text{Distance}_{(i,j)}(l)$. Next, the distances are converted to probabilities by

$$\mathrm{P}_{(i,j)}(l) = \frac{\text{Distance}_{(i,j)}(l)^{-1}}{\sum_{l=1}^{L} \text{Distance}_{(i,j)}(l)^{-1}}. \quad (1)$$

Finally, the probabilities are mapped directly to the urn composition of pixel (i,j) by

$$\mathrm{B}_{l,0}^{(i,j)} = T * \mathrm{P}_{(i,j)}(l), \quad (2)$$

where $\mathrm{B}_{l,0}^{(i,j)}$ is the number of balls of color l in pixel (i,j)'s urn at time 0 and T is the total number of balls initially in the urn.

Likewise, when applying NMC, the relative distances of the feature vectors of a pixel to the centroids in feature space are used to determine the similarity of the pixel to each class as shown in (1), and the urns are initialized as in (2).

Once the urns are initialized, the general modified Polya-Eggenberger urn sampling scheme proceeds as follows. For $n > 0$, the urn composition of each pixel (i, j) at time n is updated by sampling from a combination of the participating urns $\mathcal{V}_{n-1}^{(i,j)}$ with $\mathcal{V}_{n-1}^{(i,j)} : \{u_{n-1}^{(r,s)} : (r, s) \in N_q^k\}$, where N_q^k is the neighborhood system defined as in [6]: $N_q^k : \{q = (r, s) \in \mathcal{I} : (i - r)^2 + (j - s)^2 \leq k\}$. A simple, yet effective, sampling procedure is as follows: The urn $u_n^{(i,j)}$ for pixel $p^{(i,j)}$ is updated by first combining the balls of $u_{n-1}^{(i,j)}$ and the N neighboring urns:

$$\mathcal{C}_{n-1}^{(i,j)} = \text{ASSOCIATE}\,(\mathcal{V}_{n-1}^{(i,j)}). \qquad (3)$$

The ASSOCIATE function forms a collection of balls, $\mathcal{C}_{n-1}^{(i,j)}$, from the urns of the neighborhood. Examples of the ASSOCIATE function include grouping the urns of $\mathcal{V}_{n-1}^{(i,j)}$ into a "super" urn or sampling one ball from each urn to form the collection.

Next, a selection operation on the new collection of balls, $\mathcal{C}_{n-1}^{(i,j)}$, is performed, i.e.

$$Z_n^{(i,j)} = \text{SELECT}\,(\mathcal{C}_{n-1}^{(i,j)}). \qquad (4)$$

The SELECT function may determine the next state of the urns by sampling one ball from $\mathcal{C}_{n-1}^{(i,j)}$ or by taking the majority class of $\mathcal{C}_{n-1}^{(i,j)}$.

We denote by $Z_n^{(i,j)}$ the outcome of the SELECT function:

$$Z_n^{(i,j)} = l \text{ if the } n^{th} \text{ selected ball is color } l.$$

If $Z_n^{(i,j)} = l$, add Δ balls of color l to urn $u_n^{(i,j)}$. This yields a new urn composition for each pixel, as given above.

The above procedure is iterated until $n = N$. At time N, the final composition of each individual urn $u_N^{(i,j)}$, $(i, j) \in \mathcal{I}$ determines the final labeling of the image.

For this paper, we have developed two specific methods based on the general urn process. In method 1, the contents of the N urns in the neighborhood of pixel $p^{(i,j)}$ are collected into a "super" urn. One ball is sampled from the "super" urn, and Δ balls of that color are added to the urn of pixel $p^{(i,j)}$, $u_n^{(i,j)}$. In method 2, one ball is sampled from each of the N urns in the neighborhood to form the collection $\mathcal{C}_n^{(i,j)}$. Then Δ balls of the majority color in $\mathcal{C}_n^{(i,j)}$ are added to $u_n^{(i,j)}$.

4 Statistical Properties

The idea behind our urn sampling scheme is to promote spatial contagion of the pixel labels. At the end of the iterative process, homogeneous regions should be described by one label. It is in this sense that the urn process generates MRFs; the label of a pixel is determined by the urns in its neighborhood. In this section, we report asymptotic results to provide insight as to why the urn sampling scheme allows the initial majority color of a region to dominate the population of the urns in that region.

Consider the original, binary Polya sampling scheme. The asymptotic properties of the joint distribution can be characterized in the temporal case, i.e., when all spatial interactions are inhibited at each sampling step. In this case, it can be shown [11] that the proportion of white balls in each urn after the n^{th} trial $\rho_n^{(i,j)}$, where

$$\rho_n^{(i,j)} = \frac{\rho + \left(Z_1^{(i,j)} + Z_2^{(i,j)} + \cdots + Z_n^{(i,j)}\right)\delta}{1 + n\delta},$$

is a martingale [4] and admits a limit Y as the number of draws increases indefinitely. Indeed, $\rho_n^{(i,j)}$ (or equivalently the sample average $\frac{1}{n}\sum_{k=1}^{n} Z_k^{(i,j)}$) converges with probability 1 to Y [4]. This limiting proportion Y is a continuous random variable with support the interval $(0, 1)$ and beta probability density function with parameters $(\rho/\delta, (1 - \rho)/\delta)$:

$$f_Y(y) = \begin{cases} \frac{\Gamma(1/\delta)}{\Gamma(\rho/\delta)\Gamma((1-\rho)/\delta)} y^{\frac{\rho}{\delta}-1}(1 - y)^{\frac{1-\rho}{\delta}-1}, \\ if\ 0 < y < 1; \\ 0, \text{otherwise.} \end{cases}$$

The behavior of this pdf can be interpreted as follows: Assuming $\delta = 1$ for simplicity, if the original fraction of white balls in the urn is close to 1, then the limiting distribution of $W_n^{(i,j)}$ will be skewed towards 1. A similar behavior is obtained for the case when ρ is close to 0. Therefore, the limiting pattern will reflect the underlying probability

$$Pr\left(p_1^{(i,j)} = x\right) = \rho^x (1 - \rho)^{(1-x)}.$$

For the L-ary labeling case, the above observations generalize with convergence to the Dirichlet distribution [7].

We now examine the asymptotic behavior of the two specific sampling schemes given above. Consider sampling from the "super" urn. Restating the problem, suppose there are N urns in the neighborhood of pixel X_s, each *initially* with b_i black balls and w_i white balls, and $b_i + w_i = T$ for all i, $i = 1, 2, \ldots, N$. We put the contents of all N urns into a "super" urn, sample one ball, and add Δ balls of the same color into the urn of pixel X_s. The following properties are easily derived:

The probability of sampling exactly k black balls from n iterations of the "super" urn is

$$Pr(X = k) = \binom{n}{k} \frac{B(\alpha + k, \beta + n - k)}{B(\alpha, \beta)}, \quad (5)$$

where $\alpha = \sum_i \frac{b_i}{\Delta}$, $\beta = \sum_i \frac{w_i}{\Delta}$, and the beta function $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$ [3].

The above process can be regarded as being generated by a sequence of independent Bernoulli trials with parameter Z, where Z is random with beta distribution. In fact, it is identical with different parameters to the Polya-Eggenberger distribution in the single-urn case given above.

Remarkably, it can also be shown [3] that the average proportion of black balls in the "super" urn at any time equals the original proportion of black balls. This shows that the composition of the urn is highly dependent on the original proportion of the balls. Eventually, the majority class of the urns in a given neighborhood will spread and dominate the population of balls in that neighborhood. Therefore, we conclude that this urn sampling scheme will reinforce the majority class in a spatial neighborhood; it constitutes a positive-feedback system that yields limiting patterns of the self-reinforcing type [2]. The contagion effectively models the Markovian dependencies of the pixel labels.

The second method is described as follows: We sample one ball from each urn in $\mathcal{V}_{n-1}^{(i,j)}$. From this collection of balls, we compute the majority class, denoted by $Z_n^{(i,j)}$. We update urn $u_n^{(i,j)}$ in the same manner described in the previous section. Eventually, the initial majority class of each urn in the neighborhood will dominate its composition, thereby propagating the label throughout the neighborhood.

It is difficult to find a general closed-form expression for $P(Z_n^{(i,j)} = k)$, the probability that class k is the majority of the individual samples. The difficulty arises because we are trying to find the majority of a set of samples of a non-i.i.d. process. Hence, we resort to heuristic arguments.

The sequence of images generated by both methods exhibits both spatial and temporal dependencies represented by a Markovian relationship in terms of the urns $u_n^{(r,s)}$; more specifically:

$$Pr\{u_n^{(i,j)} | U_{n-1}, U_{n-2}, \ldots, U_0\} = Pr\{u_n^{(i,j)} | \mathcal{V}_{n-1}^{(i,j)}\},$$

where $U_n : [u_n^{(i,j)}]$ is the urn matrix associated with I_n, and $\mathcal{V}_{n-1}^{(i,j)}$ is the set of participating urns defined in the previous section.

5 Comparison of Methods

While there are many techniques for image segmentation, here we briefly examine the relationships between the urn sampling scheme and other methods with closely similar flavors, namely relaxation labeling and simulated annealing.

The idea behind the urn sampling scheme is that within a region, one label should be dominant. By repeatedly sampling with replacement from an urn, or a group of urns, the most frequently occurring color or label will asymptotically dominate the populations in the urns. Hence, contagion will promote the homogeneity of local regions.

Another interpretation of the urn process can be derived from the fact that when the urns are initialized by the ML estimate, the urn compositions represent the conditional probabilities of the image pixels. The subsequent iterative process updates these probabilities by adding balls of certain colors. The Polya-Eggenberger sampling process is such that it will naturally emphasize the majority labels.

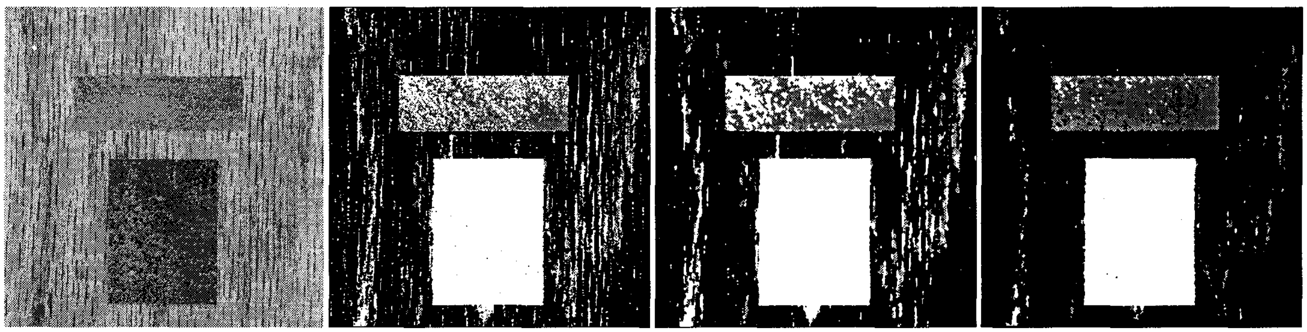
Relaxation labeling for pixel classification is an iterative procedure which assigns a best label to a pixel under certain pre-defined constraints. As noted by Kittler [9], RL's update mechanism is based on heuristic arguments so that the update and resulting segmentation is influenced by the pre-defined constraints between class labels. In the urn sampling scheme, improved segmentation is achieved without imposing such constraints. Indeed the population of each urn represents the support for the labels at each pixel. The changing of the urn compositions represents the changing of the supports; when one color dominates the population, the ambiguity of the label of a pixel is reduced.

Likewise, SA can also be viewed as a mechanism which updates the conditional probability of each pixel. The conditional probability serves as an initial estimate of the optimal point on a highly non-convex surface; an energy function for the prior distribution of the labels is adopted, and the simulated annealing procedure iteratively adds or subtracts from the conditional probability to arrive at the optimal point.

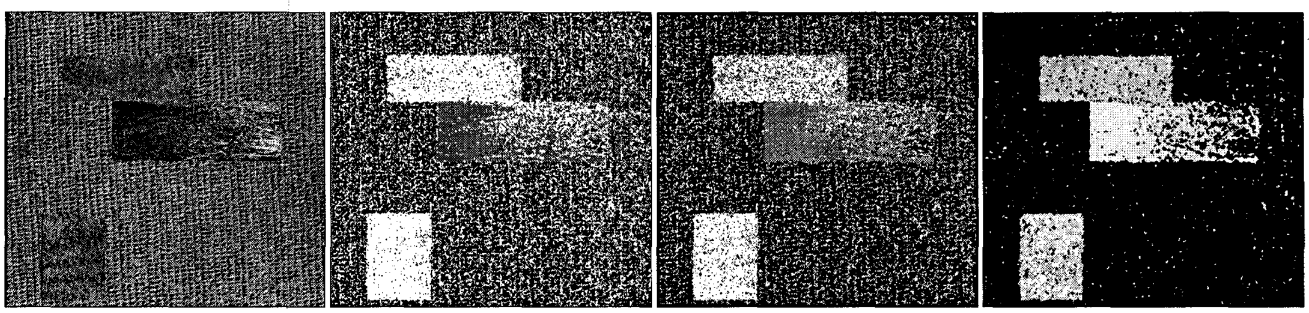
6 Experimental Results

In all of the examples given in this section, urn sampling method 1 is used; each urn is initialized with 100 balls, and Δ, the number of balls added at each iteration, is 10. Urn sampling method 2 gives similar results.

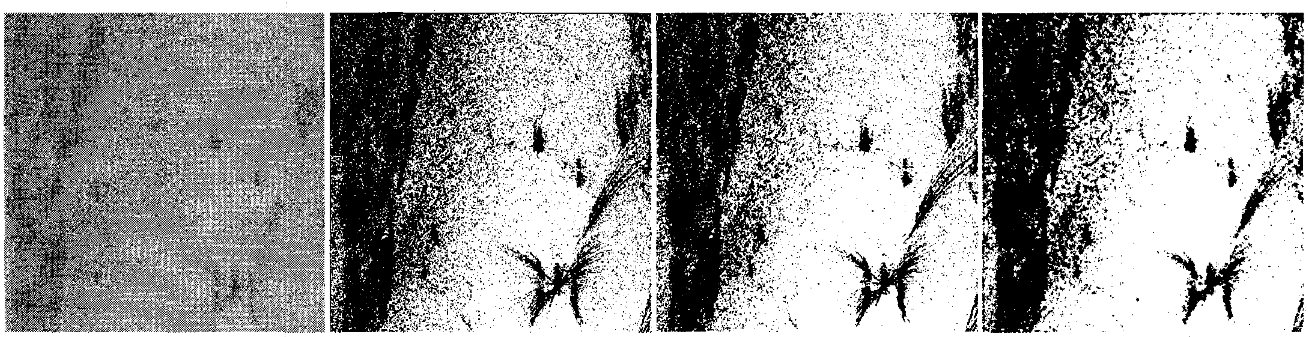
In Figures 1 and 2, the ability of the urn process to segment an image into regions of different textures is demonstrated. The initial ML estimate is found by



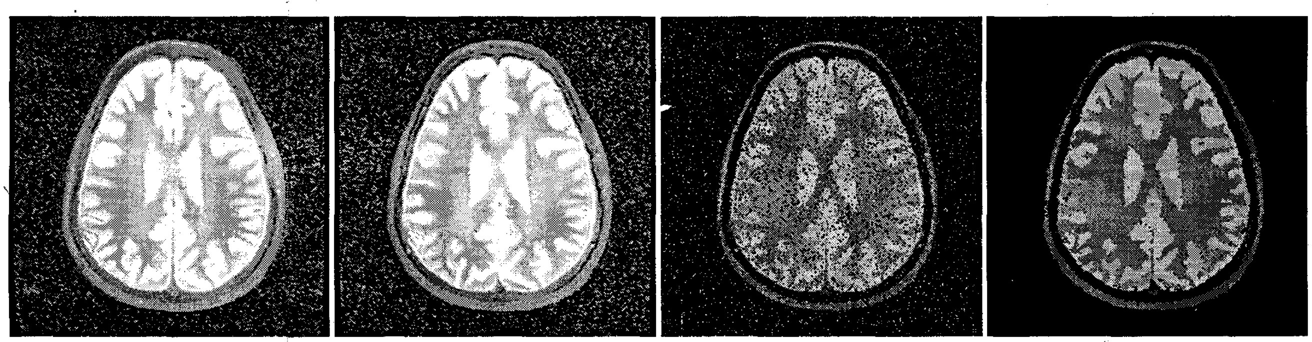
1a. Original texture image. 1b. ML segmentation. 1c. RL segmentation. 1d. Urn segmentation.



2a. Original texture image. 2b. ML segmentation. 2c. 1 iteration. 2d. 15 iterations.



3a. Original UWB SAR image. 3b. ML segmentation. 3c. Simulated Annealing. 3d. Urn Process.



4a. Proton density image. 4b. T2 image. 4c. NMC segmentation. 4d. Urn Process.

Figures 1-4: Experimental Results.

assuming that the textures can be described by a correlated Gaussian model. This model is unable to describe the grainy texture of the background, resulting in the inaccurate segmentation shown in Figures 1 and 2(b). The urn sampling process operates on the urns to produce a smoother segmentation. Figure 1 shows that the urn process provides a better segmentation than relaxation labeling. Note that the contagion-based segmentation usually preserves the edges of the texture regions better than the RL method. Figure 2 illustrates the iterative improvement resulting from the urn sampling scheme. The noisy background segmentation causes the RL algorithm to diverge to a nonsensical solution. However, the urn representation allows the urn sampling scheme to gradually adjust the urn compositions to achieve a smoother segmentation.

The UWB SAR images used in this study were obtained from the Army Research Laboratory. The image in Figure 3(a) shows the Aberdeen Proving Grounds in Aberdeen, MD during the summer of 1995. In segmenting this image, we start with ML segmentation. As shown in Figure 3(b), the resulting labeling is speckled, a characteristic of the ML segmentation technique. Application of 20 iterations of simulated annealing generates only a slightly smoother segmentation of the image (Figure 3(c)). For details on the implementation of SA for SAR image segmentation, the interested reader is referred to [12]. Note in Figure 3(d) that the urn process provides results that are comparable to SA in only 10 iterations.

To segment the MR image in Figure 4 we obtain an initial segmentation by NMC. The proton density and T2 relaxation times are the components of the two-dimensional feature vector used for NMC. Since NMC is based solely on the means of the vectors, the initial segmentation is especially sensitive to the inherent noise of the MR image modality. This leads to the speckled segmentation shown in Figure 4(c). The distances from a pixel's feature vector to the centroids in feature space determine the initial composition of the urns. The urn process then operates on the urns to produce a smoother segmentation. The output after ten iterations is shown in Figure 4(d).

7 Conclusion

In this paper, we have illustrated how modified Polya urn sampling schemes can be implemented for image segmentation. A modified sampling scheme promotes temporal and spatial contagion of pixel labels to produce a smooth segmentation. Theoretical and experimental comparisons between the urn sampling method, relaxation labeling, and simulated annealing are given.

Acknowledgments

We wish to thank Azriel Rosenfeld and Rama Chellappa for their guidance and many insightful discussions on relaxation labeling and image segmentation. The help of Lockheed Sanders and the Army Research Laboratory in Adelphi, MD for making available the UWB SAR imagery used in this work is gratefully acknowledged.

References

[1] F. Alajaji and P. Burlina, "Image Modeling and Restoration Through Contagion Urn Schemes," in *Proc. International Conference on Image Processing*, 1995.

[2] W. B. Arthur, Y. M. Ermoliev, and Y. M. Kaniovski, "Path-Dependent Processes and the Emergence of Macro-Structures," *European Journal of Operational Research*, pp. 294–303, 1987.

[3] A. Banerjee, P. Burlina, and F. Alajaji, "Image Segmentation and Labeling Using the Polya Urn Model," Tech. Rep. CS-TR-3801, Center for Automation Research, University of Maryland, May 1997.

[4] W. Feller, *An Introduction to Probability Theory and its Applications*, vol. 2, Wiley, 1971.

[5] K. Fukunaga, *Statistical Pattern Recognition*, Academic Press, 1990.

[6] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distribution, and Bayesian Restoration of Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 721–741, 1984.

[7] N. Johnson and S. Kotz, *Urn Models and Their Application*, Wiley, 1977.

[8] P. A. Kelly, H. Derin, and K. D. Hartt, "Adaptive Segmentation of Speckled Images Using a Hierarchical Random Field Model," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 1988.

[9] J. Kittler and J. Illingworth, "Relaxation Labelling Algorithms - A Review," *Image and Vision Computing*, Vol. 3, No. 4, pp. 206–216, 1985.

[10] Y. LeClerc, "Constructing Simple Stable Descriptions for Image Partitioning," *International Journal of Computer Vision*, Vol. 3, pp. 73–102, 1989.

[11] G. Polya and F. Eggenberger, "Uber die Statistik Verketteter Vorgänge," *Z. Angew. Math. Mech.*, pp. 279–289, 1923.

[12] E. Rignot and R. Chellappa, "Segmentation of Polarimetric Synthetic Apperture Radar Data," *IEEE Trans. on Image Processing*, Vol. 1, pp. 281–300, 1992.

[13] A. Rosenfeld, R. Hummel, and S. Zucker, "Scene Labeling by Relaxation Operations," *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 6, pp. 420–433, 1976.