The Effect of Stable Points on the Convergence of Markov Random Fields

Greg Budzban

Department of Mathematics
Southern Illinois University
Carbondale, IL 62901

William Casey

Department of Mathematics
University of Missouri
Columbia, MO 65211

June 25, 1998

Abstract

The convergence rate of processes governed by simulated annealing (SA) continues to be a topic of intense research. It has often been reported how the final labeling pattern in image segmentation using Markov random fields is sensitive to the temperature schedule used. In this paper, we describe the effect of introducing stable points into the initial labeling array. The stable points are seen to eliminate the dependence of the terminal point of the sample path on the temperature schedule, in addition to improving the stability of the phase transition temperature. A theorem is proved concerning the effect of the stable points on the class structure of the SA Markov chain and a conjecture is stated on the existence of an optimal set of stable points. Finally some directions for future research are discussed.

1 Introduction

A Markov Random Field (MRF) is an elegant mathematical representation for a digital image. Since their introduction by Geman and Geman [1], they and the stochastic optimization process called simulated annealing which controls the convergence of the MRF, and which together produce what is known as a simulated annealing Markov chain (SAMC) have been the objects of a great deal of study [see 2, 3 for more recent work]. Part of their attraction is the MRFs ability to scale, and they have been suggested as an excellent choice for a hierarchical image processing system. Each level of the hierarchy would simply change the structure of the neighborhood system of the MRF, thus incorporating more and more contextual information as the complexity of the decision process increases [4].

However, the practical implementation of Markov random fields in image processing is hampered by a number of technical issues. The convergence of the sample paths of the (SAMC) to a fixed image is slow. This convergence rate problem is exacerbated if one is modeling a hierarchical process, since usually at least three levels of processing are needed, if one follows the low, mid, and high level hierarchy suggested by Marr [5]. In addition, the terminal point of the sample path of the SAMC has been shown to be extremely dependent on the temperature schedule. This sensitivity of the SAMC to temperature schedule is due, at least in part, to the extreme complexity of the energy landscape for even a modest sized real image, which is littered with local minima. Thus the emphasis on searching for an optimal temperature schedule for a given class of problems is understandable [6,7].

2 Main Results

In this paper, we report the effect of introducing stable points into the initial labeling array of the SAMC. The stable points, chosen in a spatially uniform way, are fixed to a given label at the beginning of the evolution of the process. In the context of a segmentation problem, we chose no more than 1% of the pixels to be fixed. Since we were interested, at this point in the investigation, on how the stable points would effect the convergence of the SAMC, the labels were then selected by hand to conform to a correct segmentation of the image. In the future, this process would be done automatically, using a deterministic pixel classifier with a confidence measure to weed out labels of low certainty. Next the SAMC was allowed to evolve according to a linear temperature schedule with a fixed number of iterations. The initial temperature was changed to produce different schedules (see Figure 1 for details). The stable points are never visited and, thus, remain fixed throughout the evolution of the process.

0-8186-8821-1/98 $10.00 © 1998 IEEE

A number of interesting phenomena were observed. The terminal point of the sample path was virtually identical for each temperature schedule. In addition, the so-called phase transition temperature, that temperature at which the image begins to "freeze" into the terminal point, was stabilized at a temperature of 1 (see Figure 1 again). This is in direct contradiction to earlier work that attempted to eliminate the problem of schedule dependence by finding this critical value and doing a fixed temperature MRF evolution [6].

The effect of the stable points on the class structure of the SAMC is summarized by the following theorem.

Theorem 2.1 *Let $X(k) = (X_s(k))$ be a SAMC indexed spatially. Let $\{s_1, s_2, \cdots, s_j\}$ be the set of stable sites. Then $X(k)$ is a reducible Markov chain with h^j ergodic components, where h is the number of possible labels at each sites.*

PROOF. Let $X(k)$ be a SAMC where the configuration at time k_0 is given by $X(k_0) = (X_s(k_0))$, $s \in S$, where S is the set of pixels of the image. Suppose we have chosen $L = \{l_1, l_2, \cdots, l_h\}$ as the set of labels for the image, and thus the set of possible configurations for $X(k)$ is $\Omega = L^S$. Recall that two states $w_0, w_1 \in \Omega$ communicate if there is a positive probability that the Markov chain can transform each one into the other in a finite number of steps. That is, for some $m > 0$

$$P(X(k+m) = w_1 \mid X(k) = w_0) > 0,$$

and similarly for transforming w_1 into w_0. Now it is clear that there are h^j possible labelings of the stable sites. Each of these possible labelings defines a cylinder set of configurations, $C_n = L^{S-\{s_1,\cdots s_j\}} \times \{l_{n_{s_1}}\} \times \cdots \times \{l_{n_{s_j}}\}$. Any two configurations in C can eventually be transformed into each other.

However, if $w_0 \in C_m = L^{S-\{s_1,\cdots s_j\}} \times \{l_{m_{s_1}}\} \times \cdots \times \{l_{m_{s_j}}\}$ and $w_1 \in C_n$, $n \neq m$, then since the stable sites are never updated,

$$P(X(k+m) = w_1 \mid X(k) = w_0) = 0$$

for all $m \geq 0$. Thus, $X(k)$ is a reducible Markov chain with h^j ergodic components. □

Finally, due to this result, we see why the stable points solve many of the aforementioned problems. After fixing the labels at the stable points, the sample path of the process is trapped inside one of the h^j ergodic components. The energy landscape is far less complex within these smaller components. In fact, we have the following conjecture.

Conjecture 2.2 There exists an optimal set, S, of stable points, where optimality means

(1) Each global energy minimum is contained within an ergodic component for which it is the unique minimum within the component.

(2) No set of cardinality less than S satisfies property (1).

3 Future Research

The technique of stable points provides many interesting possibilities for future research. These include,

(1) What spatial distribution of stable points is optimal? More precisely, given that the conjecture is true, from what spatial distribution must one sample to produce an optimal set of points? From empirical work we have seen that stable points near edges are the most useful. Thus we believe that a good first attempt at an optimal spatial distribution is one that is locally Gaussian along an axis perpendicular to the nearest edge. It is symmetric with respect to the edge and its standard deviation is inversely proportional to the strength of the edge.

(2) Is there a way to slowly introduce stable points and maintain their benefit? By having a sequence of introductions one can reduce the probability of initial labeling error. However, the best mechanism for introducing the stable points is not immediately clear.

References

[1] Geman, D. and Geman, S., "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans. PAMI*, Vol. 6, pp. 721–741, 1984.

[2] Modestino, J. and Zhang, J., "A Markov Random Field approach to Image Interpretation," *IEEE Trans. PAMI*, Vol. 14, pp. 606–615, 1992.

[3] Niemiro, W., "Limit Distributions of Simulated Annealing Markov Chains," *Discussiones Mathematicae: Algebra and Stochastic Methods*, Vol. 15, pp. 241–269, 1997.

[4] Budzban, G. and Stirewalt, T., "Hierarchical Markov Random Fields with Stable Points," *Proc. SPIE*, Vol. 2304, pp. 105–111, 1994.

[5] Marr, D., *Vision*, Freeman Press, 1982.

[6] Johnson, D. Aragon, C., McGeough, and Schevon, "Optimization by Simulated Annealing: an Experimental Evaluation-Graph Coloring and Number Partitioning," *Operations Research*, Vol. 39, pp. 378–406, 1991.

[7] Connolly, D., "An Improved Annealing Scheme for the QAP," *European Journal of Operational Research*, Vol. 46, pp. 93–100, 1990.

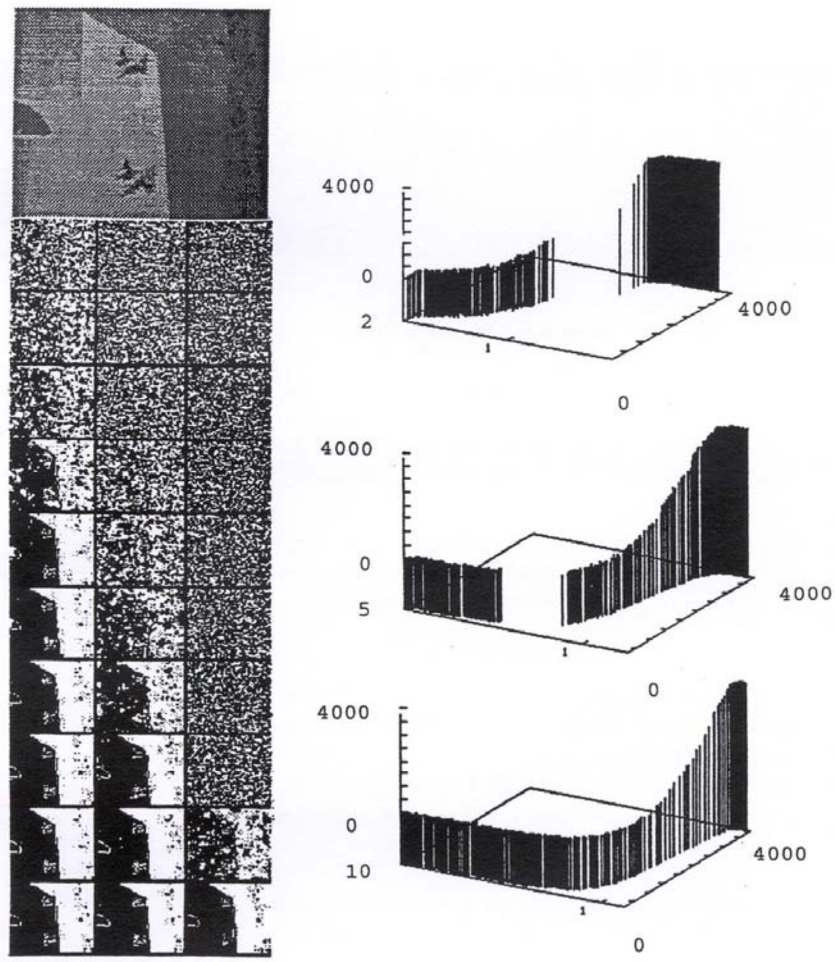

FIGURE 1. Upper Left. Original Image.

Lower Left. Three "film strips" showing different temperature schedules. Each begins at some initial temperature and converges linearly to 0, using 500 iterations. For the left film strip, the initial temperature was 2, the center was 5, the right was 10.

Right. Graphs showing the formation of 3×3 and 7×7 "crystals" within the image. The z-axis counts the number of 7×7 crystals, while the y-axis counts 3×3. Upper, middle, and lower correspond to left, center and right for the film strips.