Representing Solar Active Regions with Triangulations

Michael J. Turmon and Saleem Mukhtar

Machine Learning Systems Group, Jet Propulsion Laboratory, Pasadena, CA 91109, USA

Abstract. The solar chromosphere consists of three classes which contribute differently to ultraviolet radiation reaching the earth. We describe a data set of solar images, means of segmenting the images into the constituent classes, and a novel high-level representation for compact objects based on a triangulated spatial 'membership function'. Such representations are fitted in a variable-dimension Markov chain Monte Carlo scheme.

Keywords. Triangulation, chromosphere, Markov chain Monte Carlo, image segmentation

1 Introduction

The solar atmosphere is comprised of several features having various characteristics, and distinct physical origin. The most apparent are active regions, which are associated with sunspots in the photosphere and plages in the chromosphere. These plages can be quite large in extent, and show themselves by their strong magnetic field and altered light emissions. To a lesser extent, so does the chromospheric network, which is an evenly-distributed cell-patterned web of enhanced activity. The remainder of the surface shows only ordinary fluctuations and is termed quiet sun. See Figure 1 and Zirin (1988) for more on chromospheric features.

The three classes contribute differently to the ultraviolet radiation reaching Earth's upper atmosphere, with the plages and magnetic network giving the largest contribution. This radiation cannot be sensed directly from the ground but the features giving rise to it can be; they are used as proxy inputs to models of solar irradiance. These models are crucial to understanding phenomena such as global warming and photochemical decomposition processes in the upper atmosphere; see Withbroe (1994).

Further, much can be learned about solar irradiance by relating irradiance changes observed via satellite to region evolution identified in spatially-resolved images. Current understanding of these effects, and of plage evolution in general, is of a qualitative sort and a more refined description of anticipated plage shapes and the evolution of plage regions would be of value.

The features studied here, as well as related photospheric phenomena, are observed by many instruments on Earth and a few in space. The primary source of data for this study is the set of CaII K full-disk spectroheliograms that has been collected daily at Sacramento Peak National Solar Observatory from the mid-sixties onward. The images are recorded on photographic film, an interval of which (from the mid-eighties forward) has been digitized to 2K×2K pixels.

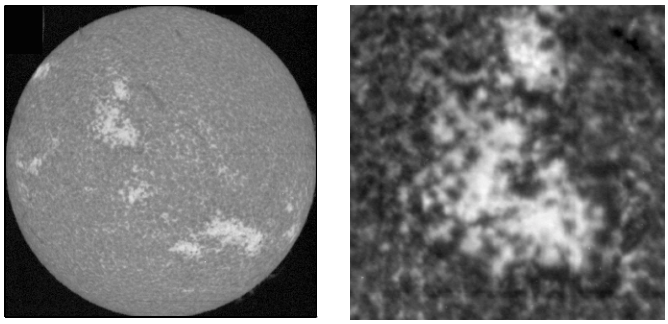


Fig. 1. A chromospheric image from 15 July 1992 shows both a decayed plage pair in the lower-right quadrant of the sun and a younger, more concentrated plage at upper left; a detail image of the latter is in the second panel

Currently, scientists typically either apply a threshold across the flattened image to determine plage areas, or manually surround the plages with polygons. The first method, while simple and objective, ignores all spatial information that is available. The second method clearly uses a large amount of side information possessed by the scientists, but is also highly subjective, difficult to even describe, and hard to repeat. We will describe a more objective and automatic procedure based on a hierarchical model of image features and formation.

We introduce our paper with an overview of the model and method we have used. While the Bayesian framework is not universally appropriate for inference problems, in the situation at hand the prior information is so apparent that approximating it is preferred to neglecting it. So, following Grenander (1991, e.g.), we establish a Bayesian formalism for a hierarchical representation of plages in three levels.

With each pixel of the observed image $\mathbf{y}$ we associate a small-integer label determining its class; these labels are $\mathbf{x}$. The labelling in $\mathbf{x}$ captures the information needed to, for example, determine how much of the chromosphere is plage. The plages themselves are large-scale phenomena which are not well-captured by pixel-level rules, so their representation should bind nearby plage sites into a cluster of heightened activity. Furthermore, even experts have uncertainty in precisely delineating plage regions, so the plage description should express this equivocation. Accordingly, the plage is represented by a membership function h across the image space, with large values indicating increased confidence that a site is plage. To combine these quantities, let there be a Markov relationship between the three levels of the stochastic model so that

$$P(h, \mathbf{x}, \mathbf{y}) = P(h)P(\mathbf{x} \,|\, h)P(\mathbf{y} \,|\, \mathbf{x}) \tag{1}$$

The interpretation is that an underlying, large-scale activity pattern h occurs, giving rise to a fine-scale pattern $\mathbf{x}$. The latter is then responsible for the observed image $\mathbf{y}$. In the next section we detail the model; then we describe the scheme for inference and provide some representative results.

2 Image representation and modelling

Denote a generic spatial position by $s = [s_1\, s_2] \in \bar{N} = [0,1]^2$. Observations are made at a lattice of sites $N \subset \bar{N}$. The class labels $\mathbf{x} = \{x_s\}_{s \in N}$ take values in the set $\{\mathtt{P}, \mathtt{N}, \mathtt{B}\}$, while entries in the corresponding observation $\mathbf{y}$ are each real-valued. We work from the data backward in defining the factors of (1).

Conditioned on the labelling, the likelihood factors as

$$P(\mathbf{y} \,|\, \mathbf{x}) = \prod P(y_s \,|\, x_s) \quad . \tag{2}$$

Labelled images supplied by scientists suggest the three densities $P(y \,|\, x)$ are lognormal, so, up to an additive universal constant,

$$-\log P(\mathbf{y} \,|\, \mathbf{x}) = \sum_{s \in N} \left(\frac{(\log y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2} + \log \sigma_{x_s} + \log y_s \right) \quad . \tag{3}$$

For $P(\mathbf{x} \,|\, h)$, we use the "Potts model", an ordinary Markov random field smoothness prior (Besag, 1974), modified so that the membership function $h(s) \in [0,1]$ favours the event $\{x_s = \mathtt{P}\}$:

$$-\log P(\mathbf{x} \,|\, h) = K_h + \beta \sum_{s \sim s'} 1(x_s \neq x_{s'}) + \alpha \sum_{s \in N} |1(x_s = \mathtt{P}) - h(s)| \quad . \tag{4}$$

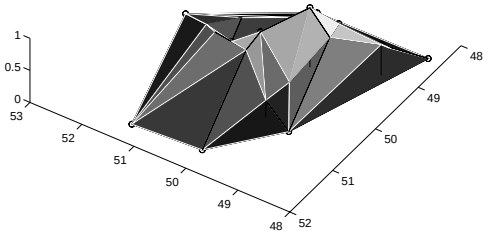
The relation $s \sim s'$ is true for "neighbouring" sites in N. On our rectangular lattice, sites are neighbours if they adjoin vertically, horizontally, or diagonally. Here $\alpha \geq 0$ indicates the influence of h, $\beta \geq 0$ favours agreement among labels, and K_h is an appropriate normalizing constant.

To represent a plage, or a cluster of related plages, we propose a tent-like structure defined by a triangulated planar graph

$$G = (V,\, E,\, h) \tag{5}$$

$V \subset \bar{N}$	a vertex set
$E \subset \bar{N}^2$	an edge relation
$h \colon V \to [0,1]$	a height function

The height function extends to all of $\bar{N}$ by linear interpolation across the faces of the pyramids. This structure models the "degree of membership" of a given pixel in the plage class and allows the binding of nearby plage regions into one coherent object. We note that, if the height function is thresholded at a given level, the resulting shape is a cluster of regions bounded by polygons — the same way scientists currently delimit plage regions manually. See the diagram below.



To define a probability distribution on membership functions, we generate each as the interpolated version of the Delaunay triangulation of iid points, uniform in $\bar{N}$. These points comprise V, and E is generated mechanically as the triangulation of V. Heights in $[0,1]$ are then assigned independently to the members of V to form tie-points. The probability density of such a membership function is induced by the one on V:

$$P(h) = Z^{-1} e^{-\gamma \operatorname{card}(V_h)} \quad . \tag{6}$$

A computational advantage of this scheme is that additions, deletions, and adjustments of one vertex have a local effect on the triangulation. Also, the penalty in log-probability paid by joining two separated graphs is the sum of component penalties, so that separated plages co-exist independently.

3 Inference

This describes the "synthesis problem"; the complementary "analysis problem" focuses on the posterior

$$P(h, \mathbf{x} \,|\, \mathbf{y}) = P(h, \mathbf{x}, \mathbf{y}) / P(y) \propto P(h, \mathbf{x}, \mathbf{y}) \quad .$$

Sampling from this distribution is a sufficient basis for any other inference scheme; we have in mind principally MAP. Adopting the well-known Markov chain Monte Carlo outlook (Besag *et al.* , 1995, for example), we sample alternately from $\mathbf{x}$ and h. The former is easily accomplished via the well-known Gibbs sampler, so we concentrate on updates to h with $\mathbf{x}$ held fixed.

One technical difficulty is the normalizing constant K_h which figures in the posterior. Existing Monte Carlo techniques for estimating K_h (Potamianos & Goutsias, 1997) simply involve sampling from $P(\mathbf{x} \,|\, h)$, but the computation involved for this is too large to justify the effort. At present we have assumed that the variation of K_h with respect to h is negligible compared to the designed variation in $P(h, \mathbf{x}, \mathbf{y})$, leading to an approximate posterior $\pi(h, \mathbf{x})$ with negative log-probability (excluding constant terms)

$$\beta \sum_{s \sim s'} 1(x_s \neq x_{s'}) + \alpha \sum_{s \in N} |1(x_s = \mathtt{P}) - h(s)| + \sum_{s \in N} \left(\frac{(\log y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2} + \log \sigma_{x_s} \right) + \gamma \operatorname{card}(V_h) \tag{7}$$

having a minimum at $(\hat{h}, \hat{\mathbf{x}})$. MAP inference proceeds, as noted, by alternately varying $\mathbf{x}$ and h while decreasing a temperature parameter.

Updates of h correspond to altering the vertex list, and are done with simple Metropolis-Hastings steps. Such a step proposes a new state h', computes $\rho(h, h') := \pi(h', \mathbf{x}) / \pi(h, \mathbf{x})$, and probabilistically accepts or rejects h' largely on this basis; this results in a Markov transition kernel $Q(v, dv')$ on the composite vertex-list set $\mathcal{V} = \cup_k \mathcal{V}_k$. If Q is designed properly, it has the posterior π as its stationary distribution. Beyond the obvious restrictions that Q be aperiodic and irreducible, it is sufficient that Q maintains detailed balance: under π, the mass moving directly from $A \subseteq \mathcal{V}$ to B equals that moving in the reverse direction.

First we describe a set of operators complete enough to ensure irreducibility. A *vertex move* operator M chooses a vertex at random and displaces it

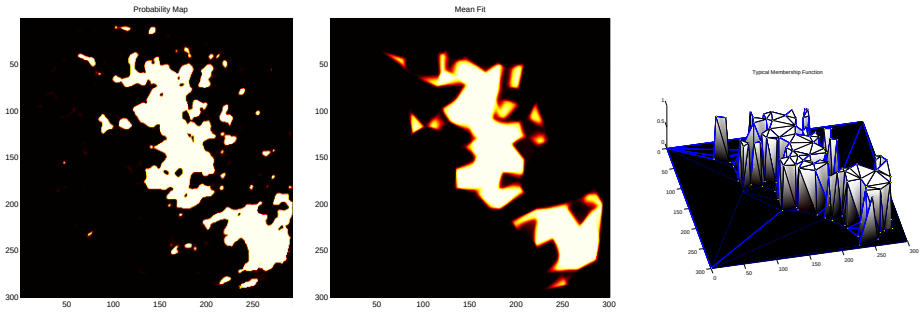


Fig. 2. Plage probability; mean inferred membership function; sample membership function

randomly. A *vertex raise* operator R raises or lowers a vertex at random. To allow movement between the constituent spaces of $\mathcal{V}$, we have *add* operators A_k, and corresponding *kill* operators A'_k, which move back and forth between $\mathcal{V}_k$ and $\mathcal{V}_{k+1}$.

Next, we define a transition kernel Q on the basis of these operators; this kernel is a "hybrid sampler" composed of each of the three move-types (M, R, A/A'). In each epoch in the simulation, one such move-type is chosen at random. Ensuring detailed balance within each move-type yields detailed balance in the superposition. Obtaining detailed balance in types M and R is trivial provided the distribution of the additive displacement is symmetric. (Modular addition will eliminate edge conditions.) Operators M and R are accepted with probability $\min(1, \rho(h, h'))$.

Obtaining detailed balance of A_k, A'_k is more complex because the flow between two different Euclidean spaces must be equalized. Following recent work of P. Green (1995), we find the chance of accepting a proposed deletion of v^* via A'_k should be the lesser of unity and

$$\rho(h, h') \times \frac{P(\text{select } A_k)}{P(\text{select } A'_k)} \times \frac{p_v(v^*)}{1/(k+1)} \quad . \tag{8}$$

(Here p_v is a density used to choose a new point for an add operation; in practice it is used to focus attention on interesting parts of the image.) The intuition is simple: the more likely it is to attempt deletion, the less likely we must be to accept it. The more likely it is to add v^* back in, the more willing we are to delete it. The factor of $k+1$ comes from the random choice of which vertex to delete: when v^* is added via A_k, there is one chance in $k+1$ that a subsequent application of A'_k will consider v^* for deletion.

4 Computational aspects and results

Initialization is important since a small feature may become hidden in a large triangle so that that π is not increased by any single vertex addition. The initialization procedure should therefore ensure locality of the effects of changes. A procedure that has proven effective is to initially replace the term of π enforcing agreement between h and the plage probability with one

penalizing per-triangle inhomogeneity:

$$\sum_T |T|\, q_T(1-q_T) \quad \text{, with} \quad q_T := |T|^{-1} \sum_{s\in T} 1(x_s = \mathtt{P})$$

and $|T|$ the number of pixels in triangle T. The modified criterion subdivides the image during an initial phase of 1000 epochs; then it is gradually replaced by the final criterion in a secondary stage twice this length. By the end of the second stage, a satisfactory basin of $\pi(h, \mathbf{x})$ has been found and the Metropolis iteration proceeds as described above.

Finally, to speed the sampling process the indicator $1(x_s = \mathtt{P})$ above is replaced with its expectation $P(x_s = \mathtt{P} \,|\, y_s)$. This is analogous to the use of conditional expectation in the ICE algorithm of A. Owen (1986) and allows the sampler to directly access the uncertainty in the label, instead of reacting to its probabilistic fluctuations as Gibbs iterations proceed.

Sample results for fitting a rather complex plage pair are shown in Figure 2. Fits with $\gamma = 2$, $\alpha = 0.4$ were obtained from a total of 30 000 Metropolis proposals taking 170 seconds of computation time on a Sun Ultrasparc. Roughly 175 proposals/sec are made by exploiting the significant cancellation in the quotient $\rho(h, h')$: only the changed triangles need be reconsidered. As desired, the membership function has suppressed the small-scale features and identified the two main objects and their principal outliers.

Acknowledgments

This work was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Thanks to Eric Mjolsness of JPL for helpful suggestions on the image representations used here, and to Judit Pap of the UCLA Department of Physics and Astronomy for help in understanding the scientific context of this problem. Delaunay triangulations were computed with the Triangle package authored by J. R. Shewchuck.

References

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. R. Statist. Soc. Ser. B*, **36**, 192–236.

Besag, J., Green, P., Higdon, D., & Mengersen, K. (1995). Bayesian Computation and Stochastic Systems. *Statistical Science*, **10**, 3–66.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.

Grenander, U., Chow, Y., & Keenan, D. (1991). *Hands: A Pattern-Theoretic Study of Biological Shapes*. New York: Springer.

Owen, A. (1986). Discussion of Ripley, "Statistics, images, and pattern recognition". *Canad. Jour. Statist.*, **14**, 106–110. Article covers pp. 83–111.

Potamianos, G., & Goutsias, J. (1997). Stochastic Approximation Algorithms for Partition Function Estimation of Gibbs Random Fields. *IEEE Trans. Inform. Theory*, **43**(6), 1948–1966.

Withbroe, G. L., & Kalkofen, W. (1994). Solar Variability and its Terrestrial Effects. *Pages 11–19 of: The Sun as a Variable Star: Solar and Stellar Irradiance Variations*. Cambridge: Cambridge Univ.

Zirin, H. (1988). *Astrophysics of the Sun*. Cambridge: Cambridge Univ.