COLOR IMAGE SEGMENTATION USING MARKOV RANDOM FIELDS*

Michael J. Daily

Hughes Research Laboratories
3011 Malibu Canyon Road
Malibu, CA. 90265

Abstract: We discuss the use of Markov Random Fields (MRFs) in color image segmentation of natural, outdoor scenes. MRFs provide an elegant means of specifying a local energy function which embodies the expected dependencies of neighboring pixels and includes both the prior and posterior probabilistic distributions. This local, neighborhood-based specification of dependencies avoids *ad hoc*, brittle methods using global image knowledge. We present a brief analysis of ongoing research in color differencing methods since they are central to the problem of color segmentation. We develop and compare the use of three different lattice structures for coupled MRFs with line and color processes based on squares, hexagons, and triangles and also discuss current efforts in MRF parameter understanding.

1. INTRODUCTION

The use of color information can significantly improve discrimination and recognition capability over purely intensity-based methods. Methods for low-level segmentation of color imagery are numerous. Techniques using recursive region splitting with histogram analysis [18],[19] and edge and boundary formation [5] have been applied to natural color imagery with some success. These methods typically suffer from a lack of important spatial knowledge in histograms and over-dependence on global thresholds and image assumptions. Statistical methods, such as classical Bayes decision theory, which are based on previous observation have also been quite popular [2],[21]. However, these methods depend on global *a priori* knowledge about the image content and organization. Until recently, very little work had used underlying physical models of the color image formation process in developing color difference metrics. Physically-based algorithms have produced excellent segmentations for color imagery obtained under controlled conditions [7],[14],[15]. In this paper, we discuss the use of Markov Random Fields (MRFs) in color image segmentation of natural, outdoor scenes. MRFs provide an elegant means of specifying a local energy function which embodies the expected dependencies of neighboring pixels and includes both the prior and posterior probabilistic distributions. This local, neighborhood-based specification of dependencies avoids *ad hoc*, brittle methods using global image knowledge. We present a brief analysis of ongoing research in color differencing methods since they are central to the problem of color segmentation. We develop and compare the use of three different lattice structures for MRFs based on squares, hexagons, and triangles and discuss current efforts in MRF parameter understanding.

* This research was supported in part by the Defense Advanced Research Projects Agency under contract DACA76-85-C-0007.

2. BACKGROUND FOR MARKOV RANDOM FIELDS

Markov Random Fields possess several characteristics which make them useful in color image segmentation. Properties such as smoothness and continuity of color regions over an entire image can be enforced using only dependencies among local neighbors. Discontinuities which separate regions of constant color may be computed while smooth regions are being found. In addition, the inclusion of both the prior and posterior distributions (through Bayes' rule) establishes a relationship between noisy observed imagery and the color segmentation results.

Details of MRF theory can be found in [4] and [17]. Briefly, a Markov Random Field is a lattice of sites; for example, an image of pixels. Since MRFs are stochastic processes, the pixels in an image may take on any of their allowed values, which means that all images can be generated. In addition, the conditional probability of a particular pixel having a certain value is only a function of the neighboring pixels, not of the entire image. The Hammersley-Clifford theorem establishes the equivalence between the conditional probabilities of the local characteristics in the MRF and local energy potentials in a Gibbs distribution. Therefore, the *a priori* probability that the MRF is in a particular state can be calculated by summing the local energies over the entire image. We are interested

in obtaining the MRF state that maximizes the *a posteriori* probability of the final segmentation given the observed data. From additional theorems, the *a priori* energies can be added to an *a posteriori* energy term involving the difference between the actual observed data and the current MRF state (or predicted image). Therefore, the Gibbs distribution energy function consists of two parts, one describing the interaction potential between neighbors, and the other associated with the difference between the predicted image and the actual observed data. Several methods of minimizing the energy function over the image (i.e. maximizing the probability) can be used, among them simulated annealing, deterministic procedures, and network solutions.

MRFs have been used for a variety of vision tasks from image and surface reconstruction to fusion of multiple low-level vision modules. Geman and Geman [4] have used the MRF approach on simple synthetic intensity images for image reconstruction. They added a useful twist to the standard approach by coupling two MRFs, one for the intensity process, and one for a binary line process. The binary line process marks the location of discontinuities in the intensity surface, making the energy non-convex and highly non-linear. Using a method they call the Gibbs Sampler, they perform the non-convex energy minimization using a Metropolis-like, simulated annealing algorithm. In the theoretical case, they prove the Gibbs Sampler will produce an annealing schedule which guarantees the global minimum; however, due to practical limitations, the schedule is sub-optimal. Marroquin [17] was the first to apply coupled MRFs with line processes to the problem of surface reconstruction from sparse depth data. More recently, Gamble and Poggio [3],[20] have used coupled MRFs for fusing low-level visual information. The formation of line process discontinuities and smoothing of processes in depth and motion data, color, and texture are coupled through separate lattices of MRFs to the intensity edges, which guide their formation. Following Marroquin, they chose sub-optimal deterministic procedures to minimize the energy functionals. Another method for surface reconstruction using sparse synthetic depth and intensity data is due to Chou and Brown [1]. Using MRFs and a technique called Highest Confidence First (HCF) to minimize the energy, they chose to update sites with the least stability (highest confidence for changing states) with respect to the current state before updating sites which were more stable. Koch, Marroquin, and Yuille [16] have used the coupled MRF approach to perform surface reconstruction from sparse depth data as well. However, they minimize the energy using linear graded neurons, as in the work of Hopfield and Tank[9]. This method lends itself more readily to implementation in analog VLSI [6],[16].

Our approach using MRFs to segment color imagery is mathematically similar to the above surface reconstruction problems, with the exception that there are multiple surfaces representing each spectral component of the color imagery.

3. METHODS FOR COMPUTING COLOR DIFFERENCES

Of primary importance to the process of color segmentation is the use of color differencing methods. The manner in which significant changes in color are detected must play an important role in the formation of color change boundaries. The addition of color information holds promise for improving segmentation results since the color of imaged surfaces is more stable under geometric changes than the image irradiance [8]. Other studies have shown the usefulness of chromatic information in improving stereo matching algorithms [11].

Typically, a color image is composed of three spectral components obtained using filters with different sensitivities in the red, green, and blue wavelengths. In many cases, it is desirable to transform from the resulting red, green, and blue (RGB) images to other color spaces which separate intensity from color (hue). Kender's [12] discussion of color transformations points out that linear transformations from RGB data (e.g. YIQ) are preferable. Non-linear transformations such as intensity, hue, and saturation (IHS) and normalized color suffer from non-removable singularities and spurious gaps in the color distribution. However, with these constraints in mind, non-linear transforms often provide more useful information by separating intensity and hue information, as in the IHS system. For additional information on color transforms, see [19] and [23].

By comparing color differences computed using several transformations, we have found that measuring differences in hue consistently produces larger relative changes than other corresponding color and intensity measures normalized to the same range. The fact that hue tends to produce the difference with highest magnitude suggests the usefulness of color as a means of segmenting. However, in many instances, hue is also more susceptible to variation from noise, and, in practice, may not yield significantly better results (see Section 5.0). Figures 1A and 1B show an experimental analysis of the usefulness of each measure in computing color differences. The graphs were obtained by normalizing and comparing the magnitude of the difference of each measure from the following set: *Euclidean RGB, intensity, hue, saturation, Maximum IHS, CIELAB, CIELAB hue, CIELAB*

Figure 1. (A) Original color image.

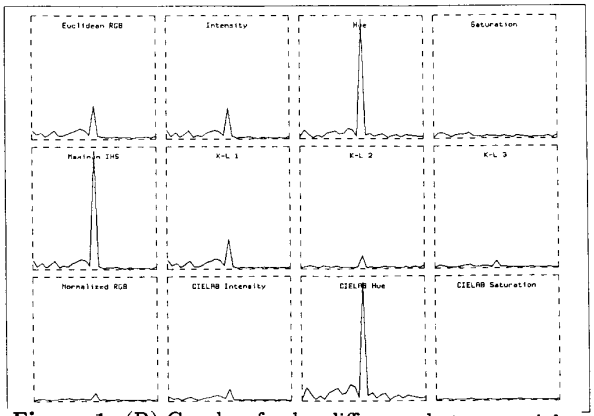

Figure 1. (B) Graphs of color difference between neighbors along the dark line in Figure 1A. The horizontal axis corresponds to distance along the dark line, while the vertical axis signifies the magnitude of the color difference.

chroma, CIELAB intensity, Karhunen-Loeve components, and normalized color. Differences were computed between neighbors along the line in Figure 1A. For the Maximum IHS technique, the maximum difference in each component in IHS space is used as long as the intensity is above 10% (25 where 255 is the maximum). When the intensity is below 10%, only the intensity difference is considered.The CIELAB color transformation attempts to produce a uniform color space where color differences a human perceives as equal correspond to equal Euclidean distances [23]. The Karhunen-Loeve components are derived using the orthogonal Karhunen-Loeve expansion which minimizes the mean-square error in basis functions as well as a measure of dispersion (entropy function) for the three RGB components [22].

More recent work attempting to quantify the significance of a color change in an image is based on physical models of the sensors and environment [7],[8]. Briefly, the spectral response of a color sensor $F(\lambda)$ may be described as the product of a filter transmission function $f_r(\lambda)$ and the camera response characteristics $f_c(\lambda)$, as shown in Figure 2. The output of the sensor D is then the integrated product of the image irradiance $I(\lambda)$ and the sensor response $F(\lambda)$. Any number of color filters may be used to improve color discrimination capability. The image irradiance at a point, $I(\lambda)$, may be approximated by using an orthogonal set of basis functions and the sensor output D, as in [8]. Color differences between points are then computed by normalizing and calculating a distance between basis functions integrated over the entire visible spectrum, which gives more reliable results than measures operating at only one spectral wavelength. We have implemented this color differencing method, and our preliminary results indicate the use of a physical model for the sensor will greatly improve color differencing capability. Our future plans include making careful comparisons of the various transformations and differencing techniques and their effect on color segmentations using MRFs.

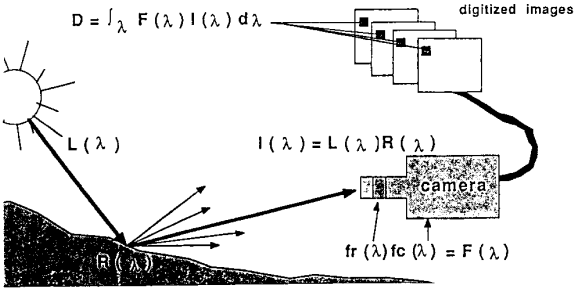

Figure 2. Physically-based sensor model for color differencing metric. $L(\lambda)$ is the spectral power distribution of the light. $R(\lambda)$ is the reflectance function of the surface. $F(\lambda)$ is the sensor response, while $I(\lambda)$ is the image irradiance. $D(\lambda)$ is the sensor output.

4. MRF LATTICES

4.1 Rectangular Lattices

The underlying lattice structure for the color and line processes plays an important role in the quality and structure of the segmentation. The simplest lattice structure is a rectangular grid, and the line processes are defined over some neighborhood of a pixel. By adjusting the energy function, larger cliques (neighbors of a site) encompassing additional directional and topographical structures can be heuristically added [3],[4]. For example, line process energies which penalize the formation of adjacent parallel discontinuities may be added. For the rectangular case, the line process can be broken into two components, horizontal and vertical, corresponding to the nearest neighbor, 4-connected system. Each of the four neighbors is of equal distance from the central pixel and is weighted identically.

In this paper, we have chosen to use the Hopfield net approach to minimization of the non-convex energy functionals, as in [16]. The sum of the following four energy terms, E_i, E_d, E_l, and E_g, must be minimized.

$$E_i = \lambda \sum_{i,j} ((\mathbf{f}_{i,j+1} \ominus \mathbf{f}_{i,j})^2 (1 - v_{i,j}) + (\mathbf{f}_{i+1,j} \ominus \mathbf{f}_{i,j})^2 (1 - h_{i,j}))$$

where E_i is the interpolation or smoothing term and λ controls the degree of smoothness in the interpolation term. The vector $\mathbf{f}$ represents the continuous-valued color process and is composed of red, green, and blue color components, and v and h correspond to the vertical and horizontal line processes respectively and vary continuously between 0 and 1. The symbol '$\ominus$' represents a color differencing scheme, such as described in Section 3. E_i is active when the line process variables v and h are less than one (i.e. the presence of a discontinuity is not yet certain). The energy data term is defined as

$$E_d = \alpha \sum_{i,j} ||\mathbf{f}_{i,j} - \mathbf{d}_{i,j}||^2$$

where the α parameter weights the importance of the input data, and $\mathbf{d}$ is a color process vector representing the observed color input image. This term ties the resulting segmentation to the original input. The third term, E_l, is the cost for introducing a new line process discontinuity and is defined as

$$E_l = \beta \sum_{i,j} (v_{i,j} + h_{i,j})$$

where β represents the penalty for formation of a line process discontinuity. When the difference terms in E_i become larger than the line process penalty β, it becomes cheaper to add a discontinuity than to continue smoothing. The final term forces the line processes to on (1) or off (0) states and is defined as follows:

$$E_g = \gamma \sum_{i,j} \left(\int_0^{v_{i,j}} g^{-1}(v)\, dv + \int_0^{h_{i,j}} g^{-1}(h)\, dh \right)$$

where g^{-1} is a standard sigmoid function representing the gain function for the line processes, and γ alters the gain term. The initial gain is nearly linear from 0 to 1 and, as the gain is increased, eventually becomes a step edge (see Figure 3), at which time the discontinuities are driven to their final states.

As in Hopfield's work [9], we chose the following update rule for the iterative minimization:

$$\frac{d\mathbf{f}_{i,j}}{dt} = \frac{-\partial E}{\partial \mathbf{f}_{i,j}}$$

$$\frac{dm_{i,j}}{dt} = \frac{-\partial E}{\partial v_{i,j}}$$

$$\frac{dn_{i,j}}{dt} = \frac{-\partial E}{\partial h_{i,j}}$$

where $E = E_i + E_d + E_l + E_g$, and m and n are internal state variables corresponding to the vertical and horizontal line processes respectively (i.e. $v_{i,j} = g(m_{i,j})$, and $h_{i,j} = g(n_{i,j})$). Finally, solving the above equations for the values of $m_{i,j}$, $n_{i,j}$, and $\mathbf{f}_{i,j}$, we obtain

$$m_{i,j} = \frac{\lambda(\mathbf{f}_{i,j+1} \ominus \mathbf{f}_{i,j})^2 - \beta}{\gamma}$$

$$n_{i,j} = \frac{\lambda(\mathbf{f}_{i+1,j} \ominus \mathbf{f}_{i,j})^2 - \beta}{\gamma}$$

$$\mathbf{f}_{i,j} = \frac{\lambda(L^v_{i,j}\mathbf{f}_{i,j+1} + L^v_{i,j-1}\mathbf{f}_{i,j-1} + L^h_{i,j}\mathbf{f}_{i+1,j} + L^h_{i-1,j}\mathbf{f}_{i-1,j}) + \alpha \mathbf{d}_{i,j}}{\lambda(L^v_{i,j} + L^v_{i,j-1} + L^h_{i,j} + L^h_{i-1,j}) + \alpha}$$

where $L^v_{i,j} = (1 - v_{i,j})$ and $L^h_{i,j} = (1 - h_{i,j})$. In order to find the minimum energy solution, these coupled systems of equations must be iterated until convergence. In practice, we have found that updating the line process once every 10 color process updates is adequate.

Figure 3B shows a sequence of steps in the energy minimization for the image of natural terrain in Figure 3A. The top row of Figure 3B is the current state of the segmented image. The second and third rows show the composite horizontal and vertical line processes and the gain function. From left to right, the gain is increasing, forcing the choice of discontinuities. The final segmentation is obtained after no change in the line or color process occurs for several increases in the gain function. For this example, we used a simple automatic method for setting the line process penalty β to the average neighborhood difference in color over the whole image. If fewer discontinuities and therefore fewer color regions are desired, β may be increased relative to the mean and standard deviation of the differences, while leaving the data weighting parameter α constant. The color differencing method used the maximum difference in IHS space, with hue and saturation values corresponding to low intensities being discarded. Figures 4 and 5 show results for different images on 128x128 MRFs using the same methods for parameter setting and a Euclidean RGB color differencing scheme. In Section 5, we address the difficult problems of choosing a color difference function and of determining parameter values.

4.2 Hexagonal Lattices

Vertical and horizontal line processes on a rectangular lattice suffer from a bias toward rectilinear structures, as is evident in Figures 3, 4, and 5. For natural terrain imagery where 90° angles are rare, the effects are especially notice-

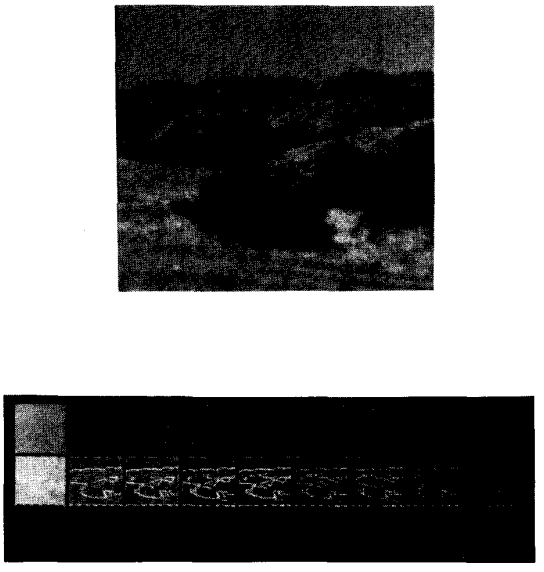

Figure 3. (A) Original color image. (B) Sequence of MRF states from left to right showing: (top row) segmented image, (middle row) horizontal and vertical line processes, (bottom row) gain function.

able. In his book, *Robot Vision*, Horn gives three important advantages of hexagonal grids for computer vision [10]. Hexagonal grids allow improved sampling and quantization during the image formation process, markedly improve the understanding of connectivity (since all neighbors touch a center pixel), and allow easy use of edge detection masks (e.g. Laplacian masks). For our purposes, hexagonal lattices in an MRF allow increased resolution of the line process and replace the unnatural rectilinear bias with a hexagonal and triangular bias much more suitable to natural imagery. A given site in the hexagonal lattice has six equally distant neighbors and three different line process directions, each at 120° angles, as shown in Figure 6. In addition, modifying the energy terms used in the rectangular case is straightforward, as discussed below. The major disadvantage to using a hexagonal lattice is the reduction in resolution required to simulate hexagonal sampling from a rectangularly sampled image.* Increased accuracy in the hexagonal sampling can be achieved by increasing the size of the hexagons in the lattice and thus decreasing the effect of the digitization bias of the rectangular grid (see Figure 7).

* This problem is purely an artifact of the rectangular grid CCD cameras used. A camera properly designed for computer vision research would use a hexagonal tesselation of photoreceptors.

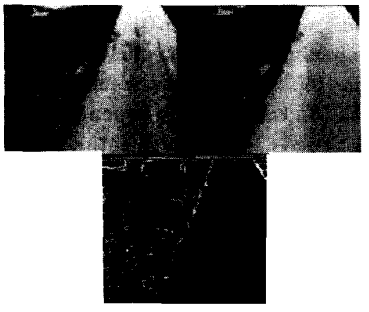

Figure 4. MRF segmentation on 128x128 rectangular lattice using Euclidean RGB color difference metric. (Top left) original image, (top right) final segmented image, (bottom) final line process discontinuities.

Figure 5. Same as above.

The energy terms for the hexagonal case are similar to those for the rectangular lattice discussed earlier, with the exception that three line processes must be used, and special indexing for even and odd rows must be included. The new energy terms are:

$$E_l = \lambda \sum_{i,j} ((\mathbf{f}_{i-1,k} \ominus \mathbf{f}_{i,j})^2 (1 - x_{i,j}) + (\mathbf{f}_{i,j+1} \ominus \mathbf{f}_{i,j})^2 (1 - y_{i,j}) + (\mathbf{f}_{i+1,k} \ominus \mathbf{f}_{i,j})^2 (1 - z_{i,j}))$$

$$E_d = \alpha \sum_{i,j} ||\mathbf{f}_{i,j} - \mathbf{d}_{i,j}||^2$$

$$E_l = \beta \sum_{i,j} (x_{i,j} + y_{i,j} + z_{i,j})$$

$$E_g = \gamma \sum_{i,j} \left(\int_0^{x_{i,j}} g^{-1}(x)\,dx + \int_0^{y_{i,j}} g^{-1}(y)\,dy + \int_0^{z_{i,j}} g^{-1}(z)\,dz \right)$$

where x, y, and z are the line processes labelled in Figure 6. The subscript k is set based on whether the current site is on an even or odd row as follows:

$$k = j + 1 \quad (i \text{ even})$$

$$k = j \quad (i \text{ odd})$$

As in the rectangular case, using Hopfield's update rule $d\mathbf{f}_{i,j}/dt = -\partial E/\partial \mathbf{f}_{i,j}$, $da_{i,j}/dt = -\partial E/\partial x_{i,j}$, $db_{i,j}/dt = -\partial E/\partial y_{i,j}$, and $dc_{i,j}/dt = -\partial E/\partial z_{i,j}$, we can solve for the values of $a_{i,j}$, $b_{i,j}$, $c_{i,j}$, and $\mathbf{f}_{i,j}$, which correspond to the three line processes, $x_{i,j}$, $y_{i,j}$, and $z_{i,j}$ and the color process, respectively. Let $L^x_{i,j} = (1 - x_{i,j})$, $L^y_{i,j} = (1 - y_{i,j})$, and $L^z_{i,j} = (1 - z_{i,j})$. Then the update rule for $\mathbf{f}_{i,j}$ is

$$\mathbf{f}_{i,j} = \frac{\lambda(L^x_{i,j}\mathbf{f}_{i-1,j} + L^x_{i+1,j}\mathbf{f}_{i+1,j} + L^y_{i,j}\mathbf{f}_{i,j+1} + L^y_{i,j-1}\mathbf{f}_{i,j-1} + L^z_{i,j}\mathbf{f}_{i+1,j} + L^z_{i-1,j}\mathbf{f}_{i-1,j}) + \mathbf{odd}_{i,j}}{\lambda(L^x_{i,j} + L^x_{i+1,j} + L^y_{i,j} + L^y_{i,j-1} + L^z_{i,j} + L^z_{i-1,j}) + \alpha}$$

$$a_{i,j} = \frac{\lambda(\mathbf{f}_{i-1,k} \ominus \mathbf{f}_{i,j})^2 - \beta}{\gamma}$$

$$b_{i,j} = \frac{\lambda(\mathbf{f}_{i,j+1} \ominus \mathbf{f}_{i,j})^2 - \beta}{\gamma}$$

$$c_{i,j} = \frac{\lambda(\mathbf{f}_{i+1,k} \ominus \mathbf{f}_{i,j})^2 - \beta}{\gamma}$$

where s1 and s2 are defined as

$$s1 = j + 1,\ s2 = j \quad (i \text{ even})$$

$$s1 = j,\ s2 = j - 1 \quad (i \text{ odd})$$

and k is defined as above. Again, the symbol '$\ominus$' represents a color differencing scheme. Figure 8 shows a comparison of the results for the rectangular and hexagonal grids applied to the same image with identical line process

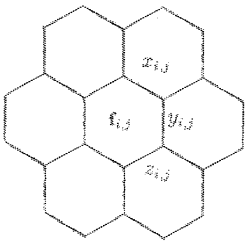

Figure 6. Structure of hexagonal lattice for MRF. The three line processes are labelled x, y, and z, and fall between the hexagonal sites. The color process vectors, $\mathbf{f}$, are located at the center of each hexagonal site and are composed of red, green, and blue components.

Figure 7. Hexagonally sampled image from rectangular image of Figure 1A using hexagons of width 8 pixels (creates a 64x64 hexagonal lattice).

penalty values and using the maximum IHS difference, as mentioned in Section 3. Both MRFs were computed on 35x40 lattices (so hexagons have approximately 14 pixel width) sub-sampled from the original 512x512 image and, using a Symbolics 3650, required approximately 15 and 20 minutes respectively to converge. The discontinuities for the hexagonal method fit the natural terrain much better than the rectangular method. Figure 9 shows the results for a 64x64 hexagonal MRF applied on a different natural terrain image, with similar parameters and using the Euclidean RGB metric.

4.3 Other Lattices

There are only three possible tesselations of the image plane using a regular polygon: a square, a hexagon, and a triangle. In addition to the square and hexagonal cases, we have investigated the use of triangular lattices for image segmentation. A triangular lattice can be used to define the color and line processes in a similar fashion to the hexagonal method. In fact, the triangular lattice actually forms a hexagonal lattice of lower resolution. The triangular lattice requires three line processes, and each site has only three equidistant neighbors along its sides. A potential advantage over the hexagonal and rectangular methods is that at each vertex in the triangular lattice,

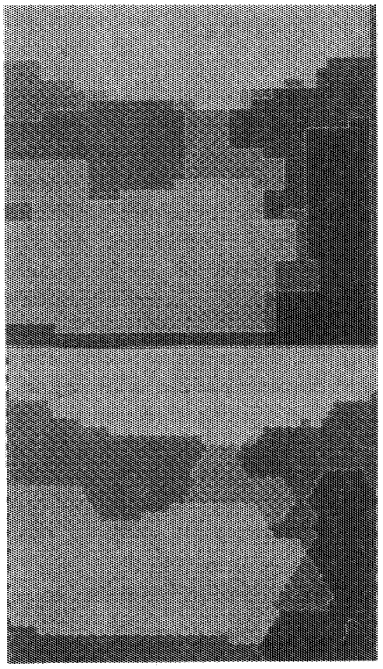

Figure 8. Comparison of rectangular and hexagonal MRFs using a 35x40 lattice (14 pixel width hexagons). The rectangular MRF above displays a marked bias toward unnatural right angles.

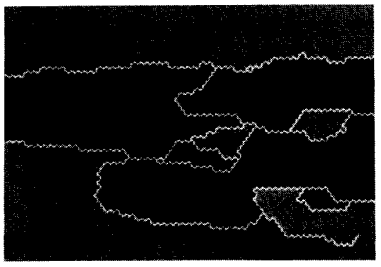

Figure 9. Segmentation results for color image of Figure 3A using a 64x64 hexagonal lattice and the Euclidean RGB color difference metric.

six possible directions for the discontinuity exist, while only three and four exist for the hexagonal and rectangular lattices, respectively. However, to date, results for the triangular lattice do not appear significantly different than those for the hexagonal technique and require additional complexity to implement. Properly indexing neighbors and each line process requires even and odd rows *and* columns to be addressed differently. In addition, convergence is painfully slow and tends to oscillate. We are currently investigating a formulation for continuous line processes that may allow more freedom in the choice of line process direction and location.

5. MRF PARAMETERS

Efficient methods of automatically setting parameters in the MRF segmentation algorithm will increase the overall usefulness of the method. There are four parameters of importance mentioned in Section 4. λ controls the degree of smoothing in the interpolation term. Higher values of λ increase the amount of smoothing. In our experiments with parameters, the value of λ is kept constant at one. The line process penalty, β, penalizes the formation of color discontinuities. Higher values of β will restrict the formation of line process discontinuities and thereby allow increased smoothing of the color process. The data weighting term, α, allows control over the importance of the input data. Higher values of α place more confidence in the validity of the input color image, while lower values are desirable for noisy data. The gain energy term is altered using the parameter γ, which we set equal to one for all of our parameter experiments. Larger values of γ favor the current state of the line processes.

Of particular importance in parameter estimation is the ratio of α to β. By increasing this ratio, a higher density of color discontinuities will form, while lowering the ratio produces fewer discontinuities and increases smoothing. We chose to test the effect of setting $\alpha = 0.1$ and varying β over the entire range of possible values. Figure 10 shows the effect of the line process penalty β on the percentage or density of discontinuities for the color image of Figure 7 using a 64x64 hexagonal lattice in the MRF. The color differencing method was Euclidean RGB normalized to the range from 0 to 255. The percentage of discontinuities was computed for each integer value of β after the network had reached a final solution. This required computation of 256 separate MRF segmentations, one for each value of β. To speed up the process of testing the parameter values, we developed a parallel implementation using 8 Lisp machines on an Ethernet. The results still required over 60 hours of computation. Fortunately, MRF algorithms are easily parallelized and are ideal for analog VLSI implementations or fine grained parallel architectures like the Connection Machine.

As shown in Figure 10, when β was zero, the final state of the line process consisted of 100% discontinuities, and no smoothing of the input data occured (i.e. the original image was the result). When β was equal to the maximum possible color difference of 255, no discontinuities formed and the result was merely a smoothed version of the input data. We have computed similar graphs for several different color images and found that the inverse logarithmic relationship of the curve between β and the percentage of discontinuities holds, with minor shifts in the location of the "elbow". Subjectively, the best segmentations appear to occur near this levelling in the density of discontinuities.

Of interest in Figure 10 is a peculiar stabilization and sudden drop in the density of the line process discontinuities at several intervals. These sudden drops in the percentage of discontinuities do not appear to be an artifact of either

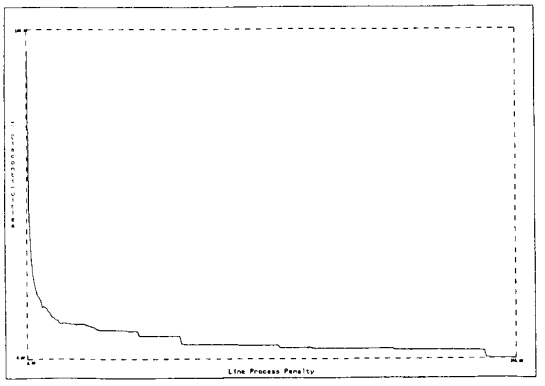

Figure 10. The effect of the line process penalty β (horizontal axis) on the percentage of discontinuities (vertical axis). For each integral value of β, the final state of the line process is computed.

the algorithms or the accuracy of the computer, but rather a result of the interaction between the underlying lattice structure and the image content. At extremely high resolution of the line process penalty, the sudden drops are equally as abrupt with no evidence of levelling out. We are currently exploring possible explanations. Perhaps as the line process penalty changes, lower energy states will correspond to shapes that best match the underlying lattice structure at discontinuous locations.

In fact, the line process penalty may be used as a kind of "temperature" to perform annealing instead of increasing the gain of the sigmoid function which is set to a constant value. The value of β is lowered from the maximum penalty to zero. When the change in the current state for a given value of β falls below a small threshold, β is decreased by some increment (one in this case). Each time the value of β is lowered, the density of discontinuities is plotted, resulting in a similar graph shown in Figure 11. Viewed from the perspective of statistical physics, the unchanging portions of the curve may correspond to "equilibrium states", while at critical temperatures the sudden addition of new line discontinuities may be similar to "phase transitions" in metals, ferromagnets, and ideal gases (see [4] and [13]).

We have also experimented with changing the parameters locally rather than globally over the image. For instance, the weighting term for the input data, α, may be given values which vary over the image corresponding to the expectation of noise in localized regions. Higher values of α place more confidence in the reliability of color information, whereas lower values indicate a low signal to noise ratio. While implementation of this scheme is simple, the major drawback to allowing parameters to vary locally is the incredible increase in difficulty of parameter understanding. The interrelationships of variable parameters make it difficult to predict the effects of changing parameters on the global segmentation result.

Another important issue in understanding and controlling segmentation results depends upon the choice of color difference functions, as mentioned in Section 3. We have only begun to analyze the effects of various differencing methods on the final segmentations. A comparison of the results for the Euclidean RGB metric and the Maximum IHS method described in Section 3 is shown in Figure 12. All other parameters are equal for the two results. The IHS method is clearly more sensitive to color changes for equal line process penalty values and caused excessive fragmentation. In the near future, we plan to use the color metric based on the sensor response discussed briefly in Section 3 and compare its results to existing techniques.

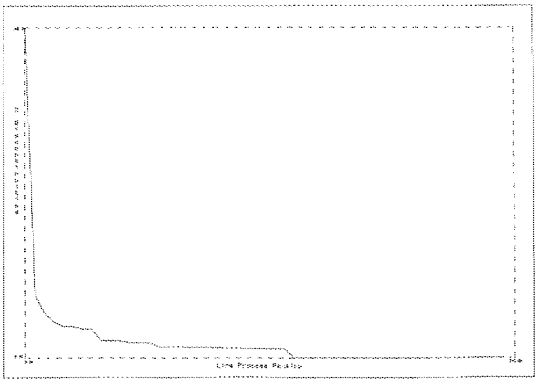

Figure 11. Intermediate states of the percentage of line process discontinuities at successively lower values of β for a single MRF.

6. CONCLUSIONS

We have shown results for color segmentation of natural terrain imagery using MRFs. Comparisons of the results for both square and hexagonal lattices using color differ-

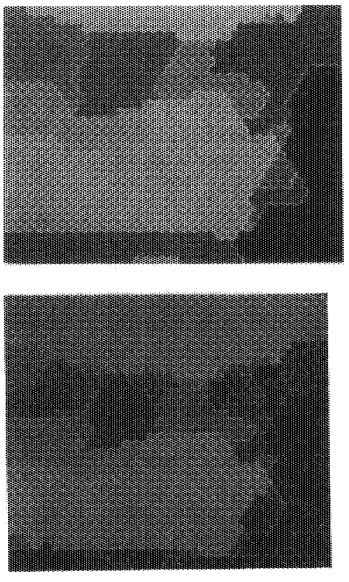

Figure 12. A comparison of two color difference methods computed using the same parameter values. Above, Euclidean RGB segmentation. Below, Maximum IHS segmentation. The IHS method produces excessive fragmentation since hue is a more sensitive difference measure at equal line process penalty values than RGB.

encing methods have been presented. We have also discussed ongoing research in the use of color metrics and in understanding MRF parameters.

We intend that the methods used in this paper for color segmentation be included in a larger perception system capable of recognizing various objects in natural terrain such as rocks, trees, bushes, gullies, and other types of vegetation. By producing an accurate description of the color of image regions which is both region-based (color process) and boundary-based (line process) and integrating additional visual cues from laser range data and other sensors (perhaps by coupled MRFs, as in [3]), we expect to produce more robust identification and therefore improve corresponding autonomous navigation capabilities. However, implementations of MRF algorithms will require specialized hardware to be practical for online use in robotic perception systems.

Acknowledgements

The author thanks John Harris for stimulating thought on the subject and the AI Center staff for their continued support.

References

[1] Chou, P. and C. Brown, "Multimodal Reconstruction and Segmentation with Markov Random Fields and HCF Optimization," *Proc. of the DARPA IU Workshop*, Cambridge, MA, pp. 214-221, April, 1988.

[2] Daily, M.J., J.G. Harris, K.E. Olin, K. Reiser, D.Y. Tseng, and F.M. Vilnrotter, "Knowledge-based Vision Techniques Annual Technical Report," U.S. Army ETL, Fort Belvoir, VA, October, 1987.

[3] Gamble, E. and T. Poggio, "Visual Integration and the Detection of Discontinuities: the Key Role of Intensity Edges," MIT AI Lab, A.I. Memo 970, October, 1987.

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

[5] Hanson, A.R. and E.M. Riseman, "Segmentation of Natural Scenes," in *Computer Vision Systems*, (A.Hanson and E.Riseman, Eds.), Academic Press, Orlando, FL, pp. 129-163, 1978.

[6] Harris, J.G. and C. Koch, "Resistive Fuses: Circuit Implementations of Line Discontinuities in Vision," to be submitted to the 1989 Snowbird Neural Network Meeting, 1989.

[7] Healey, G. and T. Binford, "The Role and Use of Color in a General Vision System," *Proc. of the DARPA IU Workshop*, Los Angeles, CA, pp. 599-613, February, 1987.

[8] Healey, G. and T. Binford, "A Color Metric for Computer Vision," *Proc. of the DARPA IU Workshop*, Cambridge, MA, pp. 854-861, April, 1988.

[9] Hopfield, J.J. and D.W. Tank, "Neural Computation of Decisions in Optimization Problems," *Biological Cybernetics*, Vol. 52, No. 14, pp. 152, 1985.

[10] Horn, B.K.P., *Robot Vision*, MIT Press, Cambridge, MA, 1986.

[11] Jordan, J.R. and A.C. Bovik, "Computational Stereo Vision Using Color," *IEEE Control Systems Magazine*, pp. 31-36, June, 1988.

[12] Kender, J., "Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use," Technical Report, Department of Computer Science, Carnegie-Mellon University, 1976.

[13] Kirkpatrick, S. and R.H. Swendsen, "Statistical Mechanics and Disordered Systems," *Comm. of the ACM*, Vol. 28, No. 4, pp. 363-373, April 1985.

[14] Klinker, G.J., S.A. Shafer, and T. Kanade, "The Measurement of Highlights in Color Images," *International Journal of Computer Vision*, Vol. 2, No. 1, pp. 7-32, June, 1988.

[15] Klinker, G.J., S.A. Shafer, and T. Kanade, "Image Segmentation and Reflection Analysis Through Color," *Proc. of the DARPA IU Workshop*, Cambridge, MA, pp. 838-853, April, 1988.

[16] Koch, C., J. Marroquin, and A. Yuille, "Analog 'Neuronal' Networks in Early Vision," MIT AI Lab, AI Memo 751, June, 1985.

[17] Marroquin, J. L., "Probabilistic Solutions of Inverse Problems," AI-TR 860, MIT AI Lab, September, 1985.

[18] Ohlander, R., K. Price, and D.R. Reddy, "Picture Segmentation using a Recursive Region Splitting Method," *Computer Graphics and Image Processing*, Vol. 8, pp. 313-333, 1978.

[19] Ohta, Y., T. Kanade, and T. Sakai, "Color Information for Region Segmentation," *Computer Graphics and Image Processing*, Vol. 13, pp. 222-231, 1980.

[20] Poggio, T., E. Gamble, and J. Little, "Parallel Integration of Vision Modules," *Science*, Vol 242, pp. 436-438, 1988.

[21] Thorpe, C., S. Shafer, and T. Kanade, "Vision and Navigation for the Carneige Mellon Navlab," *Proc. of the DARPA IU Workshop*, Los Angeles, CA, pp. 143-152, February, 1987.

[22] Tou, J.T. and R.C. Gonzalez, *Pattern Recognition Principles*, Addison-Wesley, Reading, MA, 1981.

[23] Wyszecki, G. and W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, John Wiley & Sons, New York, NY, 1982.