A MARKOV RANDOM FIELDS MODEL FOR HYBRID EDGE- AND REGION-BASED COLOR IMAGE SEGMENTATION

Slawo Wesolkowski and Paul Fieguth

Systems Design Engineering
University of Waterloo
Waterloo, Ontario, Canada, N2L-3G1
email: {swesolko,pfieguth}@uwaterloo.ca

ABSTRACT

In this paper, a framework based on a Markov Random Field approach for color image segmentation enhanced by edge detection is presented. We use a previously developed methodology to transform the image into an R'G'B' space to remove any highlight components preserving the vector-angle component, representing color hue but not intensity, to remove shading effects. To improve the segmentation process we describe the idea of a line process. This allows for the integration of region segmentation with edge detection in a Markov Random Field framework. We discuss the advantages of this new model with respect to the previously developed image segmentation model.

1. INTRODUCTION

The perception of objects in the real world without illumination effects (known as color constancy) has been a major research subject in the image science and technology communities. In spite of shading and highlight effects, humans are quite able to perceive object surfaces in a scene, a difficult task for computer systems. A Markov Random Field-based algorithm for color image segmentation invariant to shading and highlight effects developed in the context of the Dichromatic Reflection Model of Shafer [11] has been introduced [5].

In [5], the authors describe a Markov Random Field framework for color image segmentation using a pixel-based highlight invariant transformation and the inner vector product or vector angle as a similarity measure between two transformed color pixels. Both ideas were first jointly used in [18]. This consisted in applying a principal component analysis and vector angle clustering-based approach for color image segmentation in which the algorithm chooses the most optimal (in the Mean Squared Error-sense) multivector inner vector product fit to the data [3].

In [5], the vector angle was found to provide a sound color similarity framework with respect color theory. Most methods for color image segmentation presented in the literature use the Euclidean distance [13, 14]. However, the Euclidean distance is a particularly poor measure of color similarity since the RGB space is *an*-isotropic, especially when specular reflection and shading are present in the image.

The use of an MRF color model was warranted as it was shown that the Dichromatic Reflection Model was inadequate to solve the problem that sufficiently dark shades of any color all look alike (i.e., black), and similarly specular reflections or highlights converge to the same color (the color of the illuminating light, normally white) [5]. Consequently, a *spatial* Markov Random Field model was deemed *essential* in order to perform the segmentation by assigning a highlight pixel to a colored group based on its surrounding context.

For color image segmentation, it would be useful to help the segmentation algorithm with edge information. Hybrid methods combining region segmentation and edge detection have been developed which are based on heuristics and the use of the Euclidean distance such as in [4]. The focus of the present paper is to extend the color image processing and segmentation formulation based on the Dichromatic Reflection Model [11, 15], and Markov Random Fields introduced in [5] to one taking into account edge information. We propose to define the spatial context using a Gibbs/Markov approach, as outlined in Section 3, and introduce edge information using a hybrid edge/segmentation model. In this way, we propose a much more rigorous model based on stochastic optimization theory and illumination invariance properties derived from color science.

Certainly others have used Markov random fields for image segmentation [1, 9, 19]; however, normally these methods involve Gauss-Markov random fields, where the GMRF defines a spatial texture for the R, G, B components, from which segmentation can proceed as a separate hypothesis-testing procedure applied to the GMRF likelihood [10]. Our approach builds on the vector angle-based models introduced in [5]: we wish to find the segmented

Proceedings of the 2002 IEEE Canadian Conference on Electrical & Computer Engineering

Fig. 1. Original RGB color scene image, showing highlights and shading, captured using white light.

image directly as the result of energy minimization of some appropriately-defined Gibbs random field. Furthermore, the regions are not distinguished on the basis of texture, rather on shading and highlight invariant color, as well as edge information. That said, textured surfaces where the pixel variations are due to local shading effects (such as the surface of an orange) will be segmented correctly, since the normalized color is similar for all such pixels; whereas textures with intrinsically different colors (such as marble or paisley) are not the focus of our approach.

The formulation of our Gibbs model will be similar to others used for segmentation [6, 9] except for a number of variations due to the peculiarities of our transformed space and the addition of edge information. We demonstrate the advantages of constructing an energy function for Markov Random Field-driven image segmentation using a measure related to the inner vector product and integrating edge information.

This paper first describes the color illumination invariances and the development of an optimization criterion for segmentation. Next, the advantages of the new framework are outlined. Finally, conclusions and directions for future work are given.

2. COLOR THEORY

The Dichromatic Reflection Model [11, 15] assumes that light reflected from objects has two components:

1. specular reflection or highlight: this effect is characterized visually by a mirror-like reflection of the illuminating light from a surface;
2. diffuse or body reflection: this is light reflected from an object surface in all directions and is characterized by the wavelength of the color of the illuminated object

The focus of this paper will be on inhomogeneous dielectric materials such as plastics. We assume the illumination light is white or the image has been white balanced [18].

To remove the effects of highlights it is necessary to transform the pixel coordinates according to the following transformation [12, 18]:

$$\vec{c}'_{i,j} = \vec{c}_{i,j} - \frac{1}{n}\sum_{k=1}^{n} c_{i,j,k} \quad (1)$$

$$\begin{bmatrix} R'_{i,j} \\ G'_{i,j} \\ B'_{i,j} \end{bmatrix} = \begin{bmatrix} R_{i,j} \\ G_{i,j} \\ B_{i,j} \end{bmatrix} - \frac{1}{3}(R_{i,j} + G_{i,j} + B_{i,j}) \quad (2)$$

where $R_{i,j}$, $G_{i,j}$ and $B_{i,j}$ represent the color values and $\vec{c}_{i,j}$ represents the color pixel vector at location (i, j). In this transformation, the reflectance variation caused by interface reflection is removed by projecting the observed reflectance in an n-dimensional vector space along the illumination vector onto an (n-1)-dimensional subspace that is perpendicular to the illumination vector [12]. From a practical point of view, a histogram of the RGB pixels making up a homogeneously-colored region containing a highlight patch would show two connected clusters (one for the homogenous color and one for the highlight). In the transformed $R'G'B'$ space, the pixels making up that region would appear as a single linear cluster [18, 5]. Therefore, given that the RGB components are assumed to be white balanced, the application of (2) eliminates the interface reflection term.

The simplest way to obtain a shading invariant representation is to normalize the new color vectors to unit length [12, 16]. This operation puts all vectors on the unit hypersphere, except for the null vector (0,0,0) for which this operation is undefined. Since the Euclidean distance between two normalized transformed color vectors does not reflect accurately the perceptual difference between the two vectors [17], the invariance operation was factored directly into the similarity measure calculation by using one minus the cosine of the vector angle $\theta_{\vec{c}',\vec{d}'}$ between two highlight invariant color vectors $\vec{c}', \vec{d}'$; the similarity measure then becomes

$$\Theta(\vec{c}', \vec{d}') = 1 - \frac{<\vec{c}', \vec{d}'>}{|\vec{c}'||\vec{d}'|} \quad (3)$$

So if $\vec{c}'$ and $\vec{d}'$ are similar in orientation then (3) will be close to zero. Both vectors will be deemed close irrespective of the shading factors $\alpha_c(i, j)$ and $\alpha_d(i, j)$ associated with them. Therefore, this method is also shading invariant. In practice, the color vectors R', G', B' are normalized, which reduces (3) to a simple dot product calculation for each pixel comparison.

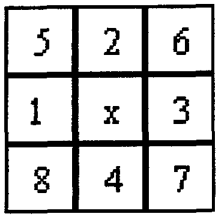

Fig. 2. Neighborhood on a lattice of regular sites. Numbers 1-4 depict a 4-pixel or first order neighborhood and 1-8 describe an 8-pixel or second order neighborhood.

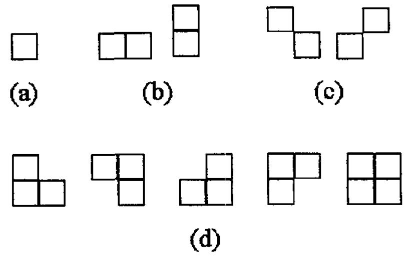

Fig. 3. Cliques for first and second order neighborhoods.

3. MARKOV RANDOM FIELDS

To accurately model the problems in this paper we will have two primary concerns: how to define an objective function for the optimal solution of the image segmentation problem, and how to find its optimal solution. Given the various uncertainties in the imaging process (e.g., quantization noise, interreflections, impurities on the camera lens and transmission noise), it is reasonable to define the desired solution in an optimization sense, such that the "perfect" or "exact" solution to our segmentation problem is interpreted as the optimum solution to the optimization objective.

Gibbs Random Fields (GRFs) [6, 19] provide a natural way of modeling context dependencies between, for example, image pixels of correlated local features [9]. Thanks to the improved insights and understanding provided by the Hammersley Clifford theorem [9], which allows Markov random field (MRF) modeling to be reinterpreted as an energy function minimization, GRF models can be used in practice. Second, the availability of methods for Gibbs sampling and Simulated Annealing [6] allow for the stochastic optimization of GRF/MRF models.

The MRF-based segmentation model is defined by the contextual relationships within the local neighborhood structure. Since our goal is the assertion of local constraints, rather than an accurate modeling of spatial textures, as in other GMRF color-segmentation research [10], we shall only be concerned with first or second order random fields, both simplifying the model and limiting the computational complexity. First and second order neighborhoods are shown in Figure 2.

Suppose we are given a color image on a pixel lattice $\mathcal{L} = \{i, j\}$. As just discussed in Section 2, each pixel $\{RGB\}_{i,j}$ is transformed to its normalized representation $\vec{c}_{i,j}$.

We can precompute the adjacent-pixel vector-angle criteria

$$\psi_{i,j} = \Theta(\vec{c}_{i,j}, \vec{c}_{i+1,j}) \quad \phi_{i,j} = \Theta(\vec{c}_{i,j}, \vec{c}_{i,j+1}) \tag{4}$$

However, not all of the vector angles are computed with the same accuracy. Even a small amount of pixel noise on a dark or highlight region results in nearly totally random vector angles, which would be separated into single-pixel regions. Given the covariance of the vector angle difference, computed by analytic or Monte-Carlo means, we introduce weights

$$w_{i,j} = \frac{1}{\mathrm{var}(\Theta(\vec{c'}_{i,j}, \vec{c'}_{i+1,j}))} \quad v_{i,j} = \frac{1}{\mathrm{var}(\Theta(\vec{c'}_{i,j}, \vec{c'}_{i,j+1}))} \tag{5}$$

Then the Gibbs energy U for segmentation using a first order model can be formulated as follows [5]:

$$\begin{aligned} &U[\{l(i,j)\}] = \\ &\sum_{i,j}\{\alpha(w_{i,j}\psi_{i,j}^2\delta_{l(i,j),l(i+1,j)} + v_{i,j}\phi_{i,j}^2\delta_{l(i,j),l(i,j+1)}) + \\ &\qquad \beta\left[(1-\delta_{l(i,j),l(i,j+1)}) + (1-\delta_{l(i,j),l(i,j+1)})\right]\} \end{aligned} \tag{6}$$

where each pixel (i, j) is assigned an integer label $0 \leq l(i, j) < N$, and where α, β control the relative constraints on the homogeneity of a single region and the degree of region fragmentation, respectively.

Model (6) is a very credible segmentation criterion, representing a considerable advance beyond standard vector-angle methods, and yet (6) is little more complicated than a standard Ising/Potts model [19] and so is well-understood and easily implemented.

Contextual relationships used in MRF modelling are usually based on neighborhood structures where the center pixel is related to the adjoining pixels using a clique potential. A clique potential is the energy enclosed within the neighborhood structure or clique. For a first order MRF model, a 4-pixel neighborhood is considered. Cliques for this neighborhood structure are shown in Figure 3.

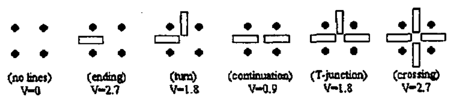

Fig. 4. Cliques with line processes for neighborhoods of size four.

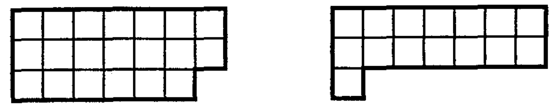

Fig. 5. Boundary length problem: both regions have the same boundary length, although very different volumes.

Edges can be modelled within the MRF context in two different ways: using line processes [6] and as a two-label (edge,non-edge) process. A line process model assumes that there is a possible "imaginary line" in between the center pixel and a neighborhood pixel. Examples of cliques with line processes are shown in Figure 4. In this paper, we will only consider models with line processes.

To implement a line process, first an edge detection algorithm in the form of two 1-D 2-pixel differences (one horizontal and one vertical) should be applied followed by an appropriate threshold. 1-D operators are required to find line sites *between* pixels whereas most edge detectors such as Sobel find edges on top of the pixel lattice since they operate on a 3-pixel window (i.e., compute the distance between the neighbors of the central pixel and placing the result at the central pixel site) [7]. In the case of a color or multispectral image, a distance measure should be used [17]. In this research since the transformed $R'G'B'$ space is used, the vector angle measure will be used as the distance measure.

Furthermore, to effectively implement line processes, a second order model will be needed and several changes to model (6) will be required. For example, if there is an edge present at a line site d, then the energy term associated with the two pixels bordering that site should be set to zero since the bond is "broken". However, if there is no edge, then the energy term remains as before. Furthermore, the boundary length constraint, which is composed of cliques of size two in (6), would be replaced by cliques of size four in the new model (these would be the only non-zero cliques [6]). Cliques of size four are shown in Figure 4 with their energy values. So while model (6) implicitly models edges using the boundary length constraint, the new model would do this explicitly using line processes.

Model (6) can then be reformulated to account for line processes in the following manner:

$$U[\{l(i,j)\}] = \sum_{i,j}\{\alpha(w_{i,j}\psi^2_{i,j}\delta_{l(i,j),l(i+1,j)} + v_{i,j}\phi^2_{i,j}\delta_{l(i,j),l(i,j+1)}) + \beta V\} \quad (7)$$

where V represents the edge pattern value of the clique as shown in Figure 4.

The primary drawback with (6) is that it is strictly a local, pixel-neighbor model and suffers from the same problems as other region-growing approaches: two vastly differently colored pixels may be grouped into a single region if they are linked by noisy or intermediately-colored pixels. A second undesired effect is that N constrains only the number of region labels, not the number of regions; that is, in regions of noise or color-gradients, (6) can generate a proliferation of small regions. Finally, the label criterion, controlled by β, measures boundary length, rather than region volume (see Figure 5). Therefore, in regions where the vector-angle criterion is vague (that is, in saturated or dark regions), a large number of pixels may have to be flipped to see *any* change in the energy, implying that only the slowest of annealing schedules will successfully converge.

There would be similar drawbacks with model (7). However, this model has the advantage of using explicit edge information whereas model (6) uses edge information implicitly. Both models could be improved by a region-merging step. Therefore, after the simulated annealing algorithm has converged, similar adjacent regions could be merged as long as the merging would lower the overall energy U. This would be followed by applying the simulated annealing algorithm and iterating until convergence. This process would be carried out until no more regions could be merged. The region merging would effectively carry out a large pixel "flip" simultaneously, thereby, considerably accelerating the convergence process. Furthermore, small regions would be merged into adjacent larger regions through merging eliminating the second disadvantage. With a region-merging step, the first drawback would remain.

Experiments are currently being carried out to verify these models and results will be presented at the conference.

4. CONCLUSIONS

We have presented a Markov Random Field-based model with a line process for shading and highlight invariant color image segmentation. The model's invariance properties have been verified using the Dichromatic Reflection Model. Furthermore, the model is based on a vector angle difference measure between color vectors and includes weights to take into account the reliability of calculating angles between various vector pairs. The advantage of using a line process as a constraint over the boundary length is evident in its explicit rather than implicit edge representation. The disadvantage would be inaccuracies in the edge detection algorithm (especially the thresholding step).

There are two immediate considerations for future work.

Furthermore, the limitation, as illustrated in Figure 5, of using the boundary length as an energy metric for each segmented region, should be revisited. The most obvious choice would be to prefer *larger* regions, where region size is measured by the number of pixels in the region. Although much more robust than boundary length, the number of pixels is a non-local criterion, and is therefore computationally much less convenient.

Finally, parameter estimation to obtain proper convergence of the MRF models is essential. In this paper, parameter estimation was ad-hoc. A formalized parameter estimation technique needs to be applied to fully evaluate the advantages of the MRF models over vector quantization and region growing-based methods when applied to real scene images.

5. REFERENCES

[1] S. A. Barker, and P. J. W. Rayner, "Unsupervised Image Segmentation Using Markov Random Fields," in M. Pelillo and E. R. Hancock (ed), *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 165-178, Springer-Verlag: 1997.

[2] R. Chellappa, and A. Jain, *Markov Random Fields: Theory and Application*. Academic Press, New York, 1993.

[3] R. D. Dony, and S. Haykin, "Image segmentation using a mixture of principal components representation," *IEE Proc. VISP*, vol. 144, pp. 73-80, April 1997.

[4] J. Fan, D.K.Y. Yau, A.K. Elmagarmid and W.G. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," IEEE Transactions on Image Processing, Vol. 10, No. 10 , Oct. 2001, pp. 1454-1466.

[5] P. Fieguth and S. Wesolkowski, "Highlight and Shading Invariant Color Image Segmentation Using Simulated Annealing," *Energy Minimization Methods in Computer Vision and Pattern Recognition III*, Sophia-Antipolis, France, September 2001, pp. 314-327.

[6] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Trans-PAMI*, Vol. 6, No. 6, 1984.

[7] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Addison-Wesley, 1993.

[8] G.J. Klinker, S.A. Shafer and T. Kanade, "A Physical Approach to Color Image Understanding," *Inter'l J. of Computer Vision*, Vol. 4, No. 1, pp. 7-38, 1990.

[9] S. Z. Li, "Modeling Image Analysis Problems Using Markov Random Fields," in C.R. Rao and D.N. Shanbhag (ed), *Stochastic Processes: Modeling and Simulation, Vol. 20 of Handbook of Statistics*. Elsevier Science, 2000, pp. 1-43.

[10] D. K. Panjwani, and G. Healey, "Markov Random Field Models for Unsupervised Segmentation of Textured Color Images," *IEEE Trans-PAMI*, Vol. 17, No. 10, 1995.

[11] S.A. Shafer, "Using color to separate reflection components," TR-136, Computer Sciences Dept., University of Rochester, NY, April 1984.

[12] S. Tominaga, "Surface Identification Using the Dichromatic Reflection Model," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 658-670, July 1991.

[13] S. Tominaga, "Color Classification of Natural Color Images," *Color Research and Application*, Vol. 17, No. 4, pp. 230-239, 1992.

[14] A. Tremeau, and N. Borel, "A Region Growing and Merging Algorithm to Color Segmentation," *Pattern Recognition*, vol. 30, no. 7, pp. 1191-1203, 1997.

[15] B.A. Wandell. *Foundations of Vision*, Sinauer Associates, Inc. Publishers, Sunderland, MA, 1995.

[16] S. Wesolkowski, M.E. Jernigan, R.D. Dony, "Global Color Image Segmentation Strategies: Euclidean Distance vs. Vector Angle," in Y.-H. Hu, J. Larsen, E. Wilson and S. Douglas (eds.), *Neural Networks for Signal Processing IX*, IEEE Press, Piscataway, NJ, 1999, pp. 419-428.

[17] S. Wesolkowski, *Color Image Edge Detection and Segmentation: A Comparison of the Vector Angle and the Euclidean Distance Color Similarity Measures*, Master's thesis, Systems Design Engineering, University of Waterloo, Canada, 1999.

[18] S. Wesolkowski, S. Tominaga, and R.D. Dony, "Shading and Highlight Invariant Color Image Segmentation Using the MPC Algorithm," *SPIE Color Imaging: Device-Independent Color, Color Hardcopy, and Graphic Arts VI*, San Jose, USA, January 2001, pp. 229-240.

[19] G. Winkler, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer-Verlag, Berlin, Germany, 1995.