Physics-Based Segmentation: Moving Beyond Color

Bruce A. Maxwell
Robotics Institute
Carnegie Mellon University

Steven A. Shafer
Microsoft Corporation

Abstract

We previously presented a framework for segmentation of complex scenes using multiple physical hypotheses for simple image regions. A consequence of that framework was a proposal for a new approach to the segmentation of complex scenes into regions corresponding to coherent surfaces rather than merely regions of similar color. Herein we present an implementation of this new approach and show example segmentations for scenes containing multi-colored piece-wise uniform objects. Using our approach we are able to intelligently segment scenes with objects of greater complexity than previous physics-based segmentation algorithms. The results show that by using general physical models we obtain segmentations that correspond more closely to coherent surfaces in the scene than segmentations found using only color.

Figure 1: A picture of a piece-wise uniform multi-colored object under white illumination.

1. Introduction

Images containing multi-colored objects and multiple materials are difficult to understand and segment intelligently. Simpler scenes containing uniformly colored objects of known material type can be segmented into regions corresponding to objects using color and one or two physical models to account for color variations due to geometry and phenomena such as highlights [1] [4] [5]. Using these methods, a discontinuity in color between two image regions is assumed to imply discontinuities in other physical characteristics such as the shape and reflectance.

Multi-colored objects, like the mug in Figure 1, violate this assumption, and thus cannot be segmented into objects by previous physics-based segmentation methods. The change in color between two image regions does not imply a discontinuity in shape, illumination, or other characteristics. To correctly interpret scenes containing more complex objects such as the mug, multiple physical characteristics must be examined to determine whether two image regions of differing color are part of the same surface. The most successful physics-based segmentation methods to date do not attempt to solve this problem. Instead, they place strong restrictions on the imaging scenario they can address--especially material type and illumination--to permit the effective use of one or two easily distinguished models [1] [4] [5].

The difficulty inherent in segmenting images with multiple materials and multi-colored objects is that by expanding the space of physical models considered for the shape, illumination, and material optics, a given image region can be described by a subspace of the general models; each point within this subspace is a valid explanation for the image region. In Figure 1, for example, one of the squares on the mug, in isolation, could be explained as a red object under white light, a white object under red light, or a grey metal object reflecting a red environment. All of these are valid explanations for the image region.

Therefore, to segment an image with multi-colored objects and numerous possible materials, shapes, and types of illumination, we must select not only the model parameters, but also the models themselves. Furthermore, we have to realize that the image may be ambiguous; we cannot simply select a single hypothesis, but must entertain several possibilities. In other words, we can never expect to get *the* single correct interpretation of an image, only one or more *possible* correct interpretations.

* This research was partially supported by the Advanced Research Projects Agency under the Department of the Army, Army Research Office under grant number DAAH04-94-G-0006.

Model selection, or instantiation has only recently been introduced to physics-based vision. Breton *et al.* have presented a method for instantiating models for both the illumination and shape, however, they still consider only a single model for material type (Lambertian) [3]. In [8] we presented a framework for segmentation using multiple physical hypotheses for shape, illumination, and material properties. Our model space incorporated general parameterizations of the transfer function, illumination, and shape, and our framework was based upon the division of this model space into broad classes, or subspaces. By reasoning about these subspaces, we proposed a method for accepting or rejecting mergers between the hypotheses of adjacent regions.

This paper describes an initial implementation of that framework using a small set of hypotheses. With this set, images containing multi-colored piece-wise uniform dielectric objects can be segmented so that the final segmentation more closely corresponds to surfaces, or objects in the scene than segmentations found using only color.

2. Modeling scenes

Our model for a scene consists of three elements: surfaces, illumination, and the light transfer function or reflectance of a point or surface in 3-D space. These elements constitute the *intrinsic characteristics* of a scene, as opposed to *image features* such as pixel values, edges, or flow fields [13]. The combination of models for these three elements is a *hypothesis* of image formation. By attaching a hypothesis to an image region we get a *hypothesis region*: a set of pixels and the physical process which gave rise to them. When an image region has multiple hypotheses, we call the combination of the image region and the set of hypotheses a *hypothesis list*.

Without prior knowledge of image content, no matter how an image is divided there are numerous possible and plausible hypotheses for each region. Variation in the color of an image region can be caused by changes in the illumination, the transfer function, or both. Likewise, variation in intensity can be caused by changes in the shape, illumination, transfer function, or any combination of the three. Numerous algorithms that extract information from single images--e.g., shape-from-shading and illuminant direction estimation--work because they assume the image variation is due to changes in only one element of the hypothesis (shape) [2] [14].

In [8] we proposed a general parametric representation for each element of a hypothesis based upon the known physical parameters. Because of their generality, however, the raw parametric models do not provide any guide to segmentation. Unlike the method of Breton *et. al.*, there are too many parameters in our models to undertake a brute-force discretization of the space of possible models [3]. Instead, we divide the parameter space for each element into a set of classes, or subspaces. These subspaces are broad enough to allow coverage of a large portion of the general element models, and yet they provide enough specificity to allow reasoning about the relationships of adjacent hypothesis regions.

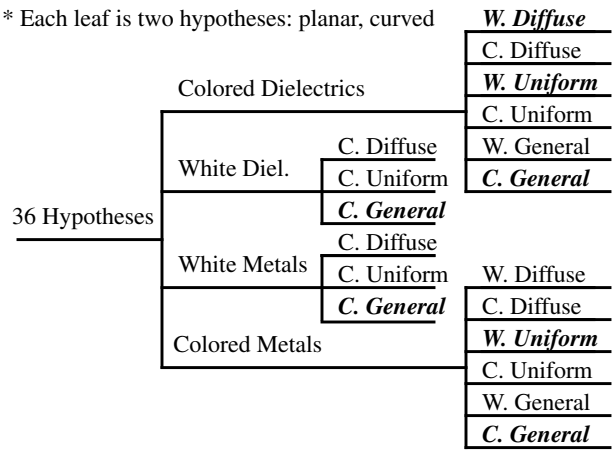


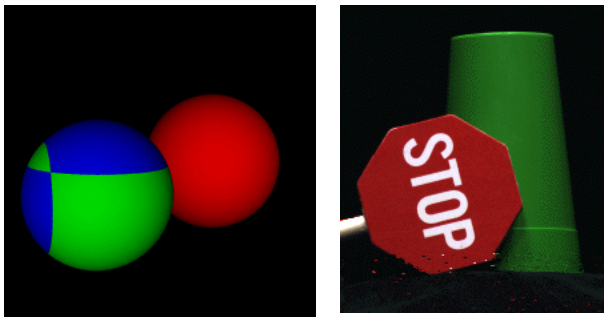
Figure 2: 36 fundamental hypotheses for a colored region. The 14 "common" hypotheses are highlighted.

The possible combinations of the broad classes we identify are shown in Figure 2 for a colored region. The first branching indicates the transfer function, the second the illumination. Each leaf of the tree has two hypotheses depending on whether the surface class is curved or planar. The set of 36 possible color-producing combinations of the broad classes we have defined as the set of *fundamental hypotheses* for a colored region. A similar set of 12 fundamental hypotheses exists for a white or grey region. Each of the fundamental hypotheses is a valid explanation for the appearance of a given region.

To denote a specific hypothesis we use the notation (<transfer function>, <illumination>, <shape>). The three elements of a hypothesis are defined as:

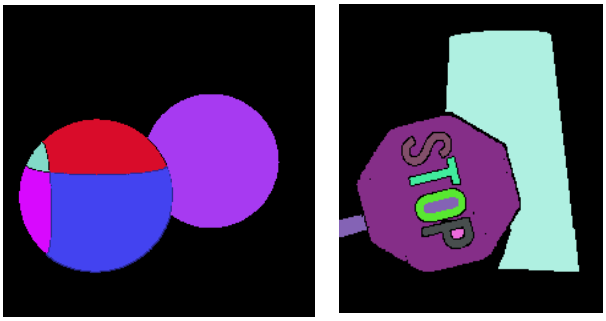
<transfer function> ∈ {Colored dielectric [CD], White dielectric [WD], Col. metal [CM], Grey metal [GM]},
<illumination> ∈ {Col. diffuse [CD], White diffuse [WD], Col. uniform [CU], White uniform [WU], Col. complex [CC], White complex [WC]}, and
<shape> ∈ {Curved [C], Planar [P]}.

By reasoning about the relative merits of the fundamental hypotheses, we are able to identify a useful subset containing 14 colored and 6 grey/white hypotheses. The hypotheses within this subset are highlighted in Figure 2. For a more extensive discussion of the generation and selection of hypotheses, see [8].



(a) (b)

Figure 3: (a) Synthetic image of two spheres, (b) image of a red and white painted wooden stop-sign and a green plastic cup taken in the Calibrated Imaging Laboratory, CMU.



(a) (b)

Figure 4: Initial segmentations of the two test images using chromaticity.

After identifying the fundamental hypotheses and further pruning them, we asked the question, which hypothesis pairs could be part of the same surface? Using rules based on physical constraints we created a table which specified all compatible pairs for the 14 colored and 6 grey/white hypotheses. A key characteristic of this table, discussed in [8], is that it is sparse, strongly constraining which hypotheses can be considered part of the same surface. Note that the table only specifies which hypotheses are definitely *not* compatible. All hypothesis pairs that can potentially be merged must undergo further analysis based on the physics of those hypotheses. In this sense, the table follows a principle of least commitment, only ruling out merges between hypotheses that are incompatible in all cases.

Another result of our framework is that it allows us to reason about different physical explanations without completely instantiating their representations. To obtain the merger table, for example, we did not need to specify a representation of shape or illumination, nor did we have to specify exact values for a given representation. The broad classes alone proved sufficient to reason about the hypothesis elements and create a table of potential merges.

The rest of the paper describes the initial implementation of a segmentation algorithm based on our framework. The algorithm proceeds as follows. First, we segment the image using region growing based on normalized color. Then the set of initial hypotheses are assigned to each region. The next step analyzes all possible pairs of adjacent hypotheses to test if they are potentially part of the same surface. Then, using the results of this step we create a hypothesis graph whose edges represent the compatibility of adjacent hypotheses. Finally, from the information in the hypothesis graph we extract the most likely final segmentations of the image.

. Initial partitioning algorithm

To test the segmentation method, we use pictures of multi-colored objects on a black background. Figure 3(a) and (b) are two example test images. Figure 3(a) is a synthetic image created using Rayshade (a public domain ray tracer). Figure 3(b) was taken in the Calibrated Imaging Laboratory at Carnegie Mellon University, as was the picture of the mug in Figure 1. While obtaining the real images, an attempt was made to include examples of only the broad hypothesis classes used in this implementation. The pole of the stop-sign is unpainted light wood, the stop-sign itself is painted red with white lettering, and the cup is green plastic. The lighting in the image comes from two fluorescent panel lights; one is above and to the right, the other above and left.

The initial segmentation is accomplished using a region growing method with normalized color, defined by

$$(c_{nr}, c_{ng}, c_{nb}) = \left(\frac{r}{r+g+b}, \frac{g}{r+g+b}, \frac{b}{r+g+b}\right) \qquad (1)$$

as the descriptive characteristic. The algorithm traverses the image in scanline order looking for seed regions where the current pixel and its 8-connected neighbors have similar normalized color and none of these pixels already belong to another region or are too dark. When it finds a seed region, it puts the current pixel on a stack and begins a region growing process based on normalized color.

When a region has finished growing, the search for another seed region continues until all pixels in the image have been checked. In the end, all pixels that are part of a region are marked with their region id in the region map. All other pixels are either too dark, or are part of a discontinuity or rapidly changing portion of the image. For now we simply ignore these pixels and concentrate on the homogeneous regions. At the end of the overall analysis, we can classify these pixels using k-nearest neighbor clus-

tering or by simply growing the homogeneous regions. For more details on the initial segmentation, see [9].

The overall goal of the initial segmentation algorithm is to find regions that can be considered part of the same surface or object. By locally growing the image regions some variation in color is allowed, but the regions generally do not grow through discontinuities caused by small variations in the transfer function or illumination.

Figure 4 shows the initial segmentations of the two test images in Figure 3. Once the initial segmentation is complete, the initial hypothesis list is assigned to each region and the analysis and merger process begins.

. Attaching hypotheses

For this implementation of the algorithm we use the hypothesis list H_c = {(Colored dielectric, White Uniform, Curved), (Colored dielectric, White uniform, Planar)} for colored regions and the hypothesis list H_w = {(White dielectric, White uniform, Curved), (White dielectric, White uniform, Planar)} for white/grey regions. These hypotheses are arguably the most important fundamental hypotheses as they represent colored and white/grey dielectric surfaces like plastic, paint, ceramics, and paper.

A region is labeled as white/grey if

$$(c_{nr} - 0.33)^2 + (c_{ng} - 0.33)^2 + (c_{nb} - 0.33)^2 < 0.0016 \quad (2)$$

where (c_{nr}, c_{ng}, c_{nb}) is the average normalized color of the region defined by (1). This threshold was set based upon the images in the test set. As the set of hypotheses considered in our current implementation all require white illumination, the exposure times for the different color bands were set so that a white board appeared white under the illumination used for the test images.

. Hypothesis analysis

Now we describe two methods for proceeding with the analysis portion of the algorithm. The more obvious and direct method we call *direct instantiation*. This involves finding estimates of and representations for the specific shape, illumination environment, and transfer function of each hypothesis. By directly comparing the representations of two adjacent hypotheses, we obtain an estimate of how similar they are. As this test compares the intrinsic characteristics, the result is both necessary and sufficient to decide whether to merge two hypotheses.

An alternative method of analysis, *weak compatibility testing*, does not directly model the hypothesis elements. Instead, it tests certain physical characteristics of adjacent hypotheses. The similarity of these characteristics are necessary but not sufficient tests of hypothesis compatibility. By using multiple tests, however, weak compatibility testing succeeds in finding most incompatible hypothesis pairs. We have explored both of these alternatives and found that this method, while less theoretically satisfying, is the more practical alternative.

.1. Direct instantiation

Direct instantiation was our first attempt at determining hypothesis compatibility. We tried to harness traditional methods of image analysis to obtain estimates of the shape and illumination of adjacent hypotheses.

While this approach is theoretically attractive, direct instantiation of hypotheses is difficult. We implemented the direct instantiation approach for the hypotheses (Colored plastic, White Uniform illumination, Curved) and (White plastic, White Uniform illumination, Curved) for which some tools of analysis do exist for finding both the shape and illumination of a scene. We implemented Zhang & Chellappa's illuminant direction estimator [14] and Bischel & Pentland's shape-from shading algorithm [2]. For a more extensive discussion of these experiments, see [9].

Our conclusion was that existing tools for analyzing the intrinsic characteristics of a scene cannot, in general, be used on small regions of an image because it violates basic assumptions necessary for the tools to function properly. The traditional literature generally assumes that an algorithm is applied to an entire object or image, not a small portion of one object. Furthermore, if we attempt to generalize direct instantiation to other hypotheses or more complex situations, we are currently limited by the lack of image analysis tools.

Another, perhaps even more important drawback of direct instantiation is that it forces you make commitments about hypothesis characteristics early in the segmentation process when you have the least information about the scene. Because of this, while direct comparison of the hypothesis elements is sufficient to make merge/not-merge determinations, it is more likely to be wrong.

.2. Weak compatibility testing

An alternative to direct instantiation is to use the knowledge constraints provided by the hypotheses to find physical characteristics that have a predictable relationship between hypothesis pairs that are part of the same object. If the characteristics do not match the prediction, we can rule out merging the adjacent hypotheses. These comparisons are necessary but not sufficient tests of compatibility. As these physical characteristics are generally local, however, they are more appropriate for region-based analysis than the direct-instantiation techniques. Furthermore, the weak methods are more robust and give good results.

Table 1: Results of the reflectance ratio test for Figure 3(b) with = 0.004.

Region 1	Region 2		σ^2	P(σ^2<^{2})
Sign	Letter S	.4463	.0004	1.0
Sign	Letter T	.4449	.0005	1.0
Sign	Letter O	.4503	.0004	1.0
Sign	Letter P	.4541	.0006	1.0
Sign	Cup	.2107	.0125	0.0
Sign	Pole	.1709	.0710	0.0
Letter O	O hole	-.4358	.0008	1.0
Letter P	P hole	-.4562	.0004	1.0

. . Reflectance ratio

One physical characteristic we use is the reflectance ratio for nearby pixels as defined by Nayar and Bolle [10].

Consider two adjacent hypotheses h_1 and h_2 that both specify colored dielectrics under white illumination. If h_1 and h_2 are part of the same piece-wise uniform object and have a different color, then the discontinuity at the border must be due to a change in the transfer function, and this change must be constant along the border between the two regions. Furthermore, along the border the two regions must share similar shape and illumination. If h_1 and h_2 belong to different objects, than the shape and illumination do not have to be the same.

The reflectance ratio is a measure of the change in transfer function between two pixels that is invariant to illumination and shape if latter two elements are similar at each pixel. If the shape and illumination of two pixels p_1 and p_2 are similar, then the reflectance ratio, defined in equation (3), where I_1 and I_2 are the intensity values of pixels p_1 and p_2, reflects the change in albedo between the two pixels [10].

$$r = \left(\frac{I_1 - I_2}{I_1 + I_2}\right) \tag{3}$$

For each border pixel p_{1i} in h_1 that borders on h_2 we find the nearest pixel p_{2i} in h_2. If the regions belong to the same object, the reflectance ratio should be the same for all pixel pairs (p_{1i},p_{2i}) along the h_1,h_2 border. A simple measure of constancy is the variance of the reflectance ratio. If h_1 and h_2 are part of the same object, this variance should be small, due mostly to the quantization of pixels, noise in the image, and small-scale texture in the scene.

If, however, h_1 and h_2 are not part of the same object, then the illumination and shape are not guaranteed to be similar for each pixel pair, violating the specified conditions for the characteristic. Differing shape and illumination should result in a larger variance in the reflectance ratio. We can select an expected variance based upon the noise, variance in object's transfer functions, and quantization effects and use this expected variance to differentiate between these two cases. Table 1 shows the variances in the border reflectance ratios of the region pairs for the test image of the stop-sign and cup.

We then use a chi-squared test to compare the variance for a particular region pair to the expected variance. The result of the chi-squared test is a likelihood that the variance in the reflectance ratio along the border is not caused by differing shape and illumination. While this test does not directly compare the shape and illumination of the two regions, the variance of the reflectance ratio along the border does implicitly measure their similarity.

The reflectance ratio can be used to compare all of the hypothesis pairs in our current implementation.

. . Gradient direction

The direction of the gradient of image intensity can be used in a similar manner to the reflectance ratio. The direction of the gradient is invariant to the transfer function for piece-wise uniform dielectric objects--except due to border effects at region boundaries. Therefore, by comparing the gradient direction of border pixel pairs for adjacent regions we obtain an estimate of the similarity of the shape and illumination. To avoid border effects, the algorithm first calculates the gradient direction of non-border pixels and then grows the results outwards.

As with the reflectance ratio, we sum the squared difference of the gradient directions of adjacent border pixels to determine a sample variance for each hypothesis pair. We then use the chi-squared test to compare the sample variance to an expected variance chosen based on numerous test images. Because of the conditions required for the gradient directions of adjacent borders to be similar, we interpret the result as a probability that the illumination and shape are similar along the border of the two regions.

Not surprisingly, the effectiveness of this characteristic is limited to regions with well-defined gradient directions. For planar or almost uniform surfaces with small gradients the angle of the gradient is very sensitive to noise and quantization errors. To avoid using noisy direction values, we only use border pixel pairs whose gradient intensities are above a selected threshold. Furthermore, for each hypothesis pair we require that a sufficient percentage of the applicable border pixels are above the gradient intensity threshold or the results are considered invalid.

Like the reflectance ratio, the gradient direction can be used to compare all of the current hypothesis pairs.

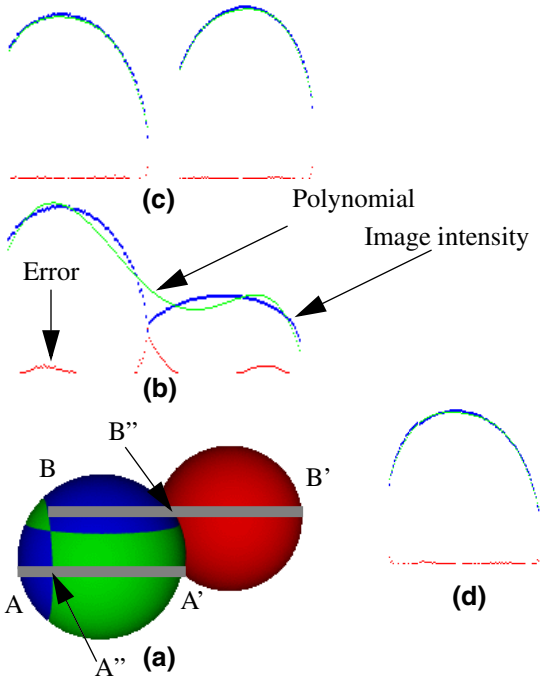


Figure 5: (a) Test image, (b) fitting one polynomial to B-B', (c) fitting two polynomials to B-B', (d) fitting one polynomial to A-A'. In (b), (c), and (d), the bottom line shows the error in the fit, while the top two lines show the image intensity overlaid by the least-squares polynomial.

. . Intensity profile analysis

So far, we have examined only calculated characteristics of the image, not the actual image intensities. The intensity profiles contain a significant amount of information, however, which we attempt to exploit with the following assertion: if two hypotheses are part of the same object and the illumination and shape match at the boundary of the hypotheses, then, if the scale change due to the albedo difference is taken into account the intensity profile along a scanline crossing both hypotheses should be continuous. Furthermore, we should be able to effectively represent the intensity profile across both regions with a single model. If two hypotheses are not part of the same object, however, then the intensity profile along a scanline containing both hypotheses is more likely to be discontinuous and representing the profile with two models should be more appropriate.

To demonstrate this property, consider Figure 5(d), which shows the intensity profile for the scanline from A to A'. We can calculate the average reflectance ratio along the border to obtain the change in albedo between the two image regions. By multiplying the intensities from A” to A' by the average reflectance ratio we adjust for the difference in albedo. As a result, for this particular case the intensity profile becomes smooth and a single model is a good representation. On the other hand, for the scanline B to B', the scaled curves are disjoint, and two models are a better representation as shown in Figure 5(b) and (c).

Rather than use the first or second derivatives of the image intensities to find discontinuities, we take a more general approach which maximizes the amount of information used and is not as sensitive to noise and small-scale texture in the image. Our method is based upon the idea that if two hypotheses are part of the same object then it should require less information to represent the intensity profile for both regions with a single model than to represent the regions individually. We use the Minimum Description Length [MDL], as defined by Rissanen [12], to measure complexity, and we use polynomials of up to order 5 to approximate the intensity profiles. The formula we use to calculate the description length of a polynomial model is given in equation (4), where x^n is the data, θ is the set of model parameters, k is the number of model parameters, and n is the number of data points [12].

$$DL = -\log P(x^n|\theta) + \frac{k}{2}\log n \tag{4}$$

Our method for a single scanline s_0 is as follows.

1. Model the intensity profile on scanline s_0 for hypothesis h_1 as a polynomial. Use the MDL principle to find the best order polynomial (we stop looking after order 5). Assign M_a the minimum description length.
2. Model the intensity profile on scanline s_0 for hypothesis h_2 as a polynomial. Again, use the MDL principle to find the best order and assign its MDL to M_b.
3. Model the scaled intensity profile of scanline s_0 for both h1 and h2 as a polynomial, find the best order using MDL, and assign the smallest MDL to M_c.
4. Compare (M_a + M_b) to M_c. To normalize the results of this test to the range [0,1], we use the measure of merit given by (5),

$$L_m = 1 - \frac{M_c - (M_a + M_b)}{M_c + (M_a + M_b)} \tag{5}$$

and any result >1.0 gets set to 1.0.

To obtain a robust measure for a region pair, we average the result of this procedure over all scanlines containing a border pixel, looking either vertically or horizontally depending upon the local border tangent. We then compare this average to the median likelihood and take the more extreme value (towards 0 or 1 depending on whether the average is less than or greater than 0.5, respectively). For more discussion of the profile analysis, see[9].

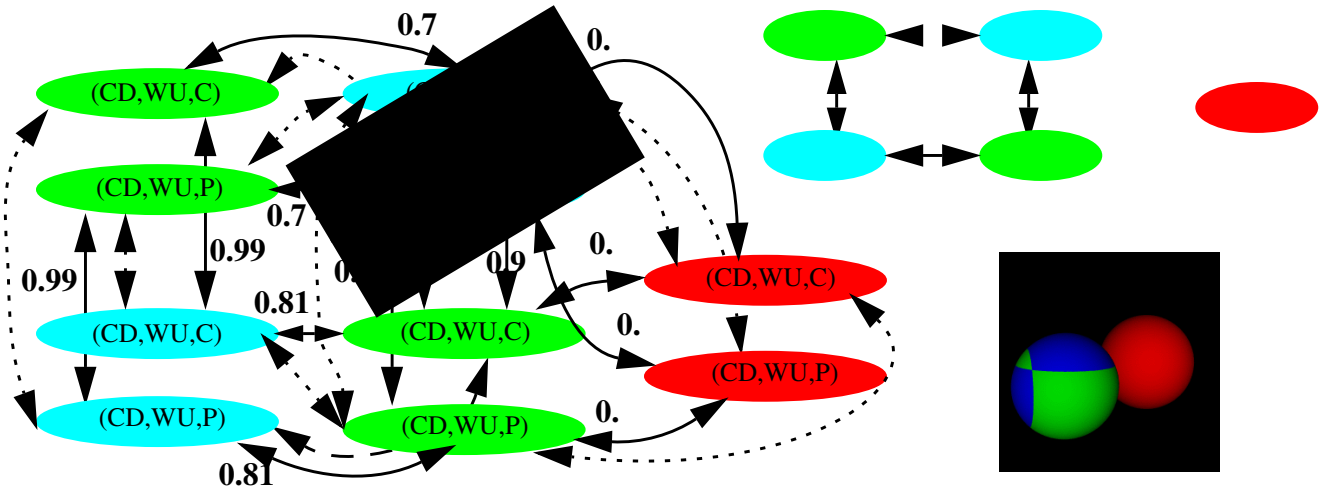


Figure 6: (a) Hypothesis graph for the image of two spheres. The numbered solid edges indicate the likelihoods of merging adjacent regions. Unlabeled dashed edges indicate incompatible hypotheses with a likelihood of 0. All adjacent hypotheses have a "not-merge" edge (not shown) with likelihood 0.5. Note, as more hypotheses are included, the hypothesis graph simply gets more levels as shown by the ellipses. (b) Best segmentation graphs extracted from (a). (c) Image of two spheres.

. Creating the hypothesis graph

For the hypotheses used in our initial implementation we can apply our three tests to obtain an estimate of whether region pairs are part of the same object.

Once all possible hypothesis pairs are analyzed we generate a hypothesis graph in which each node is a hypothesis and edges connect all hypotheses that are adjacent in the image. We then assign to each edge the likelihood that the two hypotheses it connects are part of the same object. These likelihoods are a weighted average of the results of the three compatibility tests, with the profile analysis weighted more heavily than the other two tests.

Note, however, that each edge actually has two weights associated with it. The weight assigned to the edge is a likelihood that the two hypotheses are part of the same object and should be merged in a segmentation. However, there always exists the alternative that the two hypotheses are not part of the same object and should not be merged in a segmentation. In order to find "good" segmentations, we must somehow assign a weight to the not-merge alternative. Because of the [0..1] range of the compatibility tests, we select a value of 0.5 as the cost of not merging two hypotheses. For edges such as those between the two spheres with a weight of 0.35, this means that the not-merge edge will be preferred in the final segmentation.

The hypothesis graph for Figure 3(a) is shown in Figure 6. The set of possible segmentations of the image given the complete hypothesis graph is the set of subgraphs such that each subgraph includes exactly one hypothesis from each region.

7. Extracting segmentations

A step-wise optimal algorithm exists for extracting segmentations from a single-layer graph of nodes and probabilities. Both LeValle & Hutchinson, and Panjwani & Healey have used it to segment images containing texture [7][11]. At each step the algorithm merges the most likely two nodes until it reaches a threshold based on the number of regions or the likelihood of the best merge. We have modified this algorithm to work on the multi-layer hypothesis graph.

The addition of more layers to the graph, however, means that there will almost always be more than one "best" segmentation--defined as the segmentation with the maximum sum of the likelihoods of all of the edges it contains. We would like to be able to identify all of these best segmentations, and guarantee that each hypothesis is included in at least one segmentation.

Our solution is to run the algorithm on N different graphs, where N is the number of hypotheses in the image. For each hypothesis $h \in H$, we set up a complete graph and then remove all other hypotheses from the region containing h. This forces each hypothesis to be included in at least one segmentation. The most likely segmentation from this group we specify as the best grouping of image regions. Because all discontinuous region pairs have a likelihood of 0.5, there will almost always be multiple equally likely segmentations with the same grouping of regions but different hypotheses for the individual groups. For example, in Figure 6(b), there are four possible segmentations of the two-sphere image. In all four cases each

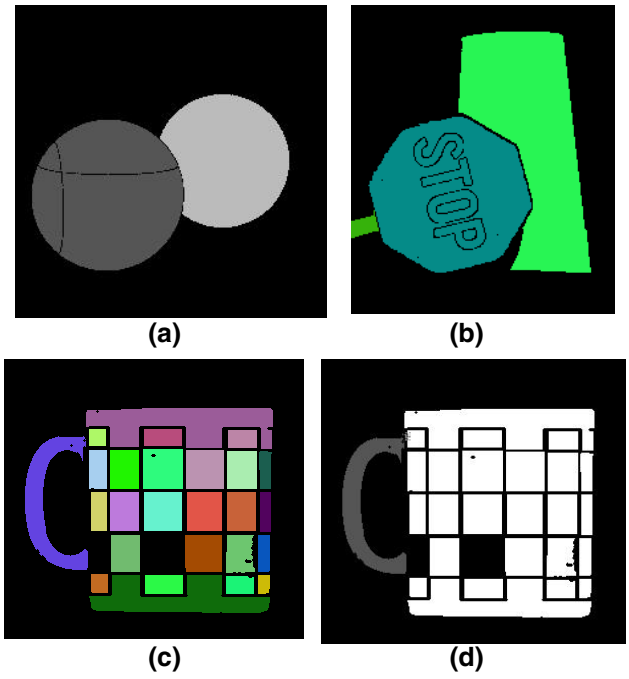


Figure 7: (a) visualization of the best region grouping for the sphere image, (b) best region grouping for the stop-sign and cup, (c) initial segmentation of, and (d) best grouping for the mug image (missing areas are too dark for analysis).

sphere is identified as a coherent surface as shown in Figure 7(a). The four segmentations are the possible interpretations of each surface being a disc or a sphere in the scene. Figure 7(b) shows the best region grouping for the stop-sign and cup image, and Figure 7(c) and (d) show, respectively, the initial segmentation and final region grouping of the mug image from Figure 1.

Note that in this implementation we do not filter or prefer hypotheses based on the image data--e.g. we might prefer planar hypotheses for regions of uniform intensity. Therefore, the planar interpretation of the two spheres image is as likely as the curved in the final hypothesis graph shown in Figure 6(b). One line of research we are pursuing is to adjust the edge weights based on the compatibility of the hypotheses with the image data. The extraction algorithm would then prefer segmentations containing hypotheses that better explain the appearance of the scene. Incorporating this information does not change the overall methodology and, in fact, is motivated by the analysis of "weirdness" presented in [8].

8. Conclusions

We have successfully implemented a system based upon the framework outlined in [8]. With only two hypotheses implemented we are able to segment images containing more complex objects than previous physics-based algorithms. Furthermore, our approach is unique among physics-based methods in that we attempt to find regions corresponding to coherent surfaces rather than regions of similar color. Future research includes expanding the number of hypotheses per region, in particular to incorporate specular highlights, and filtering hypotheses based upon their compatibility with the image.

References

[1] R. Bajcsy, S. W. Lee, and A. Leonardis, "Color image segmentation with detection of highlights and local illumination induced by inter-reflection," in *Proc. International Conference on Pattern Recognition*, Atlantic City, NJ, pp.785-790, 1990.

[2] M. Bichsel and A. P. Pentland, "A Simple Algorithm for Shape from Shading," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1992, pp459-465.

[3] P. Breton, L. A. Iverson, M. S. Langer, S. W. Zucker, "Shading flows and scenel bundles: A new approach to shape from shading," in Computer Vision - European Conference on Computer Vision, May 1992, pp.135-150.

[4] G. Healey, "Using color for geometry-insensitive segmentation," *Journal of the Optical Society of America* A 6(6), pp.920-937, June 1989.

[5] G. J. Klinker, S. A. Shafer and T. Kanade, "A Physical approach to color image understanding," *International Journal of Computer Vision*, 4(1), pp.7-38, 1990.

[6] L. Lapin, *Probability and Statistics for Modern Engineering*, Boston, PWS Engineering, 1983.

[7] S. M. LaValle, S. A. Hutchinson, "A Bayesian Segmentation Methodology for Parametric Image Models," Technical Report, University of Illinois at Urbana-Champaign Robotics/Computer Vision Series, UIUC-BI-AI-RCV-93-06, 1993.

[8] B. A. Maxwell and S. A. Shafer, "A Framework for Segmentation Using Physical Models of Image Formation," in *Proceedings of Conference on Computer Vision and Pattern Recognition*, IEEE, pp361-368, 1994.

[9] B. A. Maxwell and S. A. Shafer, "Physics-Based Segmentation: Looking Beyond Color," Technical Report, Robotics Institute, Carnegie Mellon University, CMU-RI-TR-95-37, 1995.

[10] S. K. Nayar and R. M. Bolle, "Reflectance Based Object Recognition," to appear in the *International Journal of Computer Vision*, 1995.

[11] D. Panjwani and G. Healey, "Results Using Random Field Models for the Segmentation of Color Images of Natural Scenes," in Proceedings of International Conference on Computer Vision, June 1995, pp.714-719.

[12] J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, Singapore, World Scientific Publishing Co. Pte. Ltd., 1989.

[13] J. M. Tenenbaum, M. A. Fischler, and H. G. Barrow, "Scene Modeling: A Structural Basis for Image Description," in *Image Modeling*, ed. Azriel Rosenfeld, New York, Academic Press, 1981.

[14] Q. Zheng and R. Chellappa, "Estimation of Illuminant Direction, Albedo, and Shape form Shading," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, July 1991, pp.680-702.