

Object Arrangement Estimation using Color Edge Profile

Yuichi NAKAMURA, Naoaki SUMIDA, Yuichi OHTA

Institute of Information Sciences and Electronics, University of Tsukuba
Tsukuba 305-8573, JAPAN, (yuichi@is.tsukuba.ac.jp)

Abstract

In this paper, we propose a method for classifying image edges caused by different physical phenomena, i.e. reflectance change, shadow, occlusion, etc., by using color information around the edge. We assumed several simple models for object spatial arrangements. For each of them, typical locus of RGB values along the normal direction of each edge segment is modeled. Each locus is parameterized by several features. In the classification of actual edges, the most plausible phenomenon is selected by checking the consistency between the parameters from an actual edge profile and those from each model. For the improvement of accuracy, Dempster-Shafer probability model is employed to deal with the above parameters that are often weak and uncertain. Experiments showed good performances.

1 Introduction

Intensity discontinuity is one of the most important features in analyzing images. Especially color distributions have extremely useful information for delineating object boundaries and for discriminating physical phenomena around edges. Although there has been limited research on this topic, there are a few related works. Kanade proposed a method for using edge profiles in making a pair or group of edges [1]. Marik, et. al, proposed edge discrimination based on color distribution [3]. Maxwell, et. al, proposed a method to utilize an edge profile for image segmentation [4, 5]. They proved the utility of gray or color distribution around an edge.

We propose a novel method for using this information to discriminate physical phenomena around edges. It has the following advantages which enable more detailed

scene analysis:

- It utilizes several physical object arrangement models.
- It analyzes a locus in the RGB space.
- It effectively combines weak and ambiguous clues by Dempster-Shafer theory.

2 Modeling Color Edge Profile

2.1 Physical Edge Model

Intensity changes are usually caused by several typical phenomena. For these phenomena, we consider six physical models shown in 1.

Albedo: The albedo changes across edges with no substantial depth discontinuity.

Occlusion: Two regions belong to different objects meet at the edge. No substantial shadow exists between the two objects.

Shadow: Two regions belong to one homogeneous region on an object. One is a shadow region cast by another object, and the other is directly illuminated by a normal light source.

Ridge: Two regions border at a ridge or a valley of an object.

Compound: Two objects overlap, one of the objects casts a shadow on the other.

Touch: Two regions belong to two objects touching or close to each other. Direct illumination is attenuated in the crack or gap.

These phenomena are not always separable: for example, an intermediate phenomenon between *Touch* and *Compound* can be imagined. To cope with this problem,

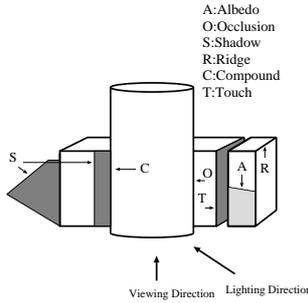


Figure 1: Physical phenomena

we use probability model in which a probability can be assigned to a set of phenomena. This avoids the risk of overspecifying the phenomenon without enough information.

2.2 Edge Profile in the RGB Space

In the above model, the most intrinsic characteristics can be found in the intensity profile along an edge's gradient. First, we assume that the RGB value becomes gradually stable as a sampling point is getting apart from the edge. Therefore, if we re-sample over enough length, we can expect at least two clusters at the start and the end of the sample sequence. Let C_1 and C_2 be the representative RGB values, *i.e.* the cluster centers, on both sides. For each model presented in the previous section, the ideal intensity changes can be modeled as follows:

Albedo: Rapid transition from one cluster to the other as shown in 2(a). The RGB value p of a locus point can be modeled as follows:

$$p = kC_1 + (1 - k)C_2, \quad 0 \leq k \leq 1$$

Occlusion: The same as the *Albedo* case except that the region on the occluding side often has smooth shading.

Shadow: Two color clusters and the origin of the RGB space are aligned, since we assume the ambient light can be modeled as the attenuation of the direct light as shown in 2(b).

$$C_1 \cong kC_2, \quad p = k'C_1, \quad k' \leq k \leq 1$$

Ridge: The same as *Shadow* except that the edge is often accompanied by highlights.

Compound: The locus is mainly composed of two parts. (a) The RGB value quickly changes from the cluster C_1 on the occluding side to kC_2 on the shaded part of the other side, *i.e.* the attenuated value of C_2 . (b) The

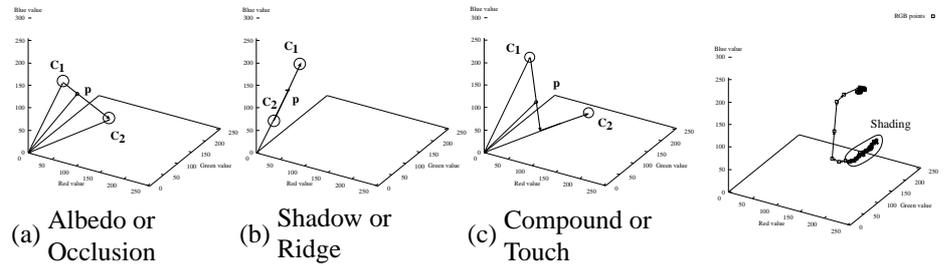


Figure 2: Edge profile models in the RGB space

Figure 3: Smooth Shading

RGB value recovers quickly from the attenuated value kC_2 to C_2 (2(c)).

$$p = \begin{cases} (a) & kC_2 - C_1 \\ (b) & k'C_2 \end{cases} \quad (1)$$

$$0 \leq k < k' \leq 1$$

Touch: The shaded portion of the *Compound* is small, and mutual reflection may be observed.

In some of the above cases, because precise modeling of ambient light is difficult, we modeled the ambient light as the attenuation of the direct light.

2.3 Partial Edge Profile

Since some of the above physical models have similar loci, we modeled additional partial characteristics which have enough power to discriminate all of the above physical models. Although those are powerful clues for discrimination, we cannot always observe them.

Smooth shading: A smooth intensity change with almost the same hue, as shown in 3, is often apparent. This can be used to distinguish *Occlusion* from characteristics such as *Albedo*.

Highlight/Shadow: A highlight portion can be observed as an overshoot portion on a locus. Similarly, if a small shadow portion is observed in the case of *Compound* or *Touch*, the corresponding portion of a locus is an overshoot toward the origin of the RGB space.

Smooth and slow transition: Shadow boundaries are usually unclear. This implies the smooth and slow transition between clusters. This feature is different from the above *Smooth shading* in the sense that hue could change along the transition.

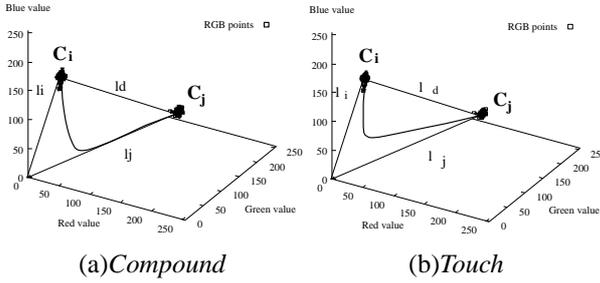


Figure 4: Comparison between *Compound* and *Touch*

Locus shift: *Compound* and *Touch* cannot be clearly discriminated with the models discussed in the previous section. In the case of *Compound*, however, the locus shifts to an imaginary line segment between one of the cluster and the origin of the RGB space. In the case of *Touch*, this shift is rarely observed because of mutual reflection.

3 Feature Extraction from Edges

To discriminate actual loci based on the above locus models, we use several features extracted from an edge point sequence. First, edge points are detected by an ordinary edge detection methods, *e.g.* Canny edge detector. The edge points are tracked and stored as edge point sequences. Then, RGB values on the original image are re-sampled along a *sampling probe*, which is a line segment located at every pre-determined interval on an edge point sequence, and is parallel to gradient direction at the point.

3.1 Features on the whole

We define the features for parameterizing the characteristics presented in 2.2. If an edge is one of the above models, the corresponding features are always expected.

Alignment of the clusters (F1): As shown in 5, consider two line segments l_i and l_j , which are between the origin and the clusters C_i and C_j , respectively. The orientation difference between l_i and l_j is considered as a parameter which shows the hue difference of the two clusters. If the angle is smaller than the pre-determined threshold θ_{th} , this feature is considered to be present.

Linearity of transition (F2): To parameterize linear transition between two clusters, the distance between

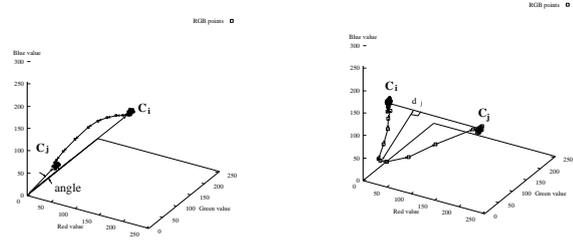


Figure 5: Collinearity of two clusters and the origin

Figure 6: Linearity of the transition between two clusters

each transitional point p and the line segment (l_d) between two clusters, *i.e.* C_i and C_j , are calculated as shown in 6. If the average distance \bar{d} is smaller than the pre-determined threshold d_{th} , this feature is considered to be present.

Bending to the origin 1 (F3): To detect the bending portion of a locus as shown in 7, we defined the closeness to three line segments l_i , l_j , and l_d defined above. First, consider transition points p_k and p_{k+1} . The line closest to the midpoint of these two points is considered to get the segment length $|p_{k+1} - p_k|$ as a score. Then, by summing up the score by each transition point, we get the score for each line. If the score of l_i or l_j is larger than the score of l_d , this feature is considered to be present.

Third cluster (F4): If the shadow portion in the *Compound* or *Touch* model is large enough, the bending portion in a locus is often observed as other intermediate cluster(s). If the following conditions are satisfied, the fourth feature is considered to be present.

- The number of clusters is three or more.
- Feature F2 is observed for the clusters on both ends of a locus.
- Feature F1 is observed between one of these clusters and each of intermediate clusters.

3.2 Features on portion

Similarly to the above feature definitions, we defined features on partial characteristics of a locus.

Distribution in a cluster (F5): If a cluster is of line-like shape, the coefficient of determination tends to be large for the first principal component axis. This feature is

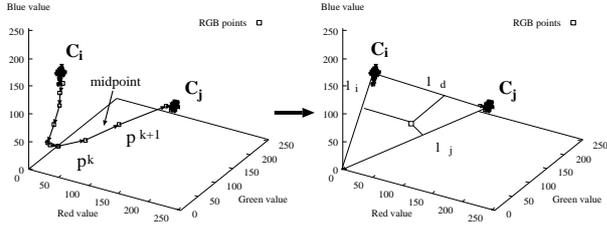


Figure 7: Bending to the origin

considered to be present if the first eigenvalue of the variance-covariance matrix is large, and the coefficient of determination is larger than the threshold.

Overshoot (F6): In order to detect Shadow/Highlight features, the RGB values are projected onto one dimensional profile in gray level, then the overshoots are extracted [6].

Slow and smooth transition (F7): The number of points in transition between clusters are counted. If the number is larger than the threshold, this feature is considered to be present.

Bending to the origin 2 (F8): If the scores defined for feature F4 have close values to each other, transition on the locus stay apart from all the imaginary line segment l_i , l_j , and l_d . This strongly suggests inter-reflection between two faces. If s_i/s_j is close to 1, this feature is treated as present.

4 Probabilistic Model

Any single feature defined above is not strong enough to classify an actual edge profile into a single phenomenon, since features are often blurred, degraded, or even eliminated by noise. We use DS theory to integrate weak clues.

4.1 Basic Probability Assignment

Each feature supports a set of phenomena and does not support the rest. For example, if F1 is observed, there is large possibility that *Ridge* or *Shadow* is present, and small possibility of the feature detection errors which cannot be of any help for discrimination. Considering this, large basic probability m_s is assigned to the former, and the rest $(1 - m_s)$ is assigned to *Uncertain*, that is a set of all phenomena.

Table 1: Basic probability on each feature

Phenomena	F1		F2		F3	
	1	0	1	0	1	0
Albedo		A_1/B_1	A_1/B_1			A_1/B_1
Occlusion		A_1/B_1	A_1/B_1			A_1/B_1
Shadow	A_1/B_1		A_1/B_1			
Ridge	A_1/B_1		A_1/B_1			
Touch		A_1/B_1		A_1/B_1		
Compound1		A_1/B_1		A_1/B_1		
Compound2			B_1			B_1
Uncertain	A_2/B_2	A_2/B_2	A_2/B_2	A_2/B_2	A_1/B_1	A_2/B_2
A_1	0.8	0.8	0.7	0.5	1.0	0.6
A_2	0.2	0.2	0.3	0.5		0.4
B_1	0.8	0.8	0.8	0.6	1.0	0.8
B_2	0.2	0.2	0.2	0.4		0.2

A_i : in the case of two clusters

B_i : in the case of three or more clusters

C_i : do not care the number of clusters

F6.1: Overshoot(upward) F6.2: Overshoot(downward)

Compound1: without shadow cluster, Compound2: with shadow cluster

1: a feature is observed 0: a feature is not observed

Since we need guidelines in determining the basic probability values, we counted the coincidences of observed features and the actual phenomena. Based on this probability, we determine the basic probabilities as shown in 1. A_i shows the basic probability assigned when only two clusters are found, while B_i is the probability when three or more clusters are found.

4.2 Combination

For the probability combination, we utilized the most basic scheme, shown in the following formula.

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j)} \quad (2)$$

After the basic probability combinations, we choose a candidate or a set of candidates. We tried two methods based on the following idea:

- A candidate or a set of candidates which has a large lower probability should be chosen.

- A phenomenon or a set of phenomena which have large upper probability should be given priority in searching for candidates.

Method1: The phenomenon or set of phenomena which has the largest lower probability is chosen. The sets of phenomena are pre-determined, and no further combinations of phenomena will be tried.

Method2: Until the lower probability gets larger than the threshold, the system tries to make larger sets of phenomena.

1. The system sets $n = 1$, and registers each single phenomenon as a candidate.
2. The system examines the newly registered sets of phenomena. From these, it searches for a set of phenomena which has the largest upper probability and has the lower probability larger than the threshold. If found, it return that set as the chosen candidate.
3. If nothing is found, then the system tries to make larger sets of phenomena. First, it finds the set of phenomena which has the largest upper probability. Then, it makes any combinations between the set of phenomena and any single phenomenon which is not already included in the set. It then registers the newly generated sets. If all sets of phenomena with n elements are checked, goto step 4, otherwise goto step 2.
4. The system sets $n = n + 1$, and goes to step 2.

In Method2, the results heavily depend on the threshold for the lower probability. More specifically, if the threshold is set smaller, the size of chosen set becomes smaller, while the accuracy, *i.e.* the rate that the correct phenomenon is included in the chosen set, is degraded. On the other hand, if the threshold is set larger, we obtain a larger set large possibility set. Since this trade-off should be examined, the performance is measured by changing the threshold.

5 Experiments

We applied our method to relatively simple indoor scenes, two of which are shown in 8. First, edge points are extracted, and RGB values are re-sampled as mentioned in 3. The interval between the sampling probes was set to 30 pixels, and each probe was 30 pixels in length on both side, totaling 60 pixels.

The performance was checked by the following four experiments.

Experiment1: Features: F1–F4, Method: Method1, and the candidates for the discrimination were chosen from the four sets, $\{\{Albedo, Occlusion\}, \{Shadow, Ridge\}, \{Touch, Compound1\}, \{Compound2\}\}$.

Experiment2: Features: F1–F4, Method: Method2, and the smallest sets for the discrimination were the same as Experiment1.

Experiment3: Features: F1–F8, Method: Method1, and the candidates for discrimination were individual phenomena.

Experiment4: Features: F1–F8, Method: Method2, and the smallest elements for discrimination were individual phenomena.

where *Compound1* is *Compound* without shadow clusters, and *Compound2* is that with shadow clusters.

In the above Experiment1 and Experiment2, the units for discrimination were not individual features, since features from F1 through F4 do not have enough information to narrow the candidates down to a single phenomenon. On the other hand, it is potentially possible to discriminate individual phenomena if we use all of the features F1 through F8.

The results are shown from 2 through 5. These are comparisons with the ideal, manually determined, answers for over 300 edge segments. More specifically, the rate that the correct manually selected phenomenon is included in the chosen set is shown. The performance changes according to threshold value changes are also shown in these tables. For 2-a, 2-b, and 2-c, the threshold value is set to 0.5, 0.7, and 0.9, respectively. The same can be said for Experiment4. The average number of elements which the chosen set has is shown in the bottom row of the tables. As we can see in 2, around 90% accuracy is obtained for Experiment1 and Experiment2. This is because the lower probability always becomes large for a simple scene in which most of the edges can be classified into the physical models defined in this research.

On the other hand, 4 shows the difficulties in discriminating individual phenomena for even a simple scene. This is mainly because features F5–F8 cannot always be observed, and therefore sufficient information is not always obtained. Even in those cases, we can see Method2 improves the result, as shown in the column of Experiment4 a–c. This shows that we can safely obtain a

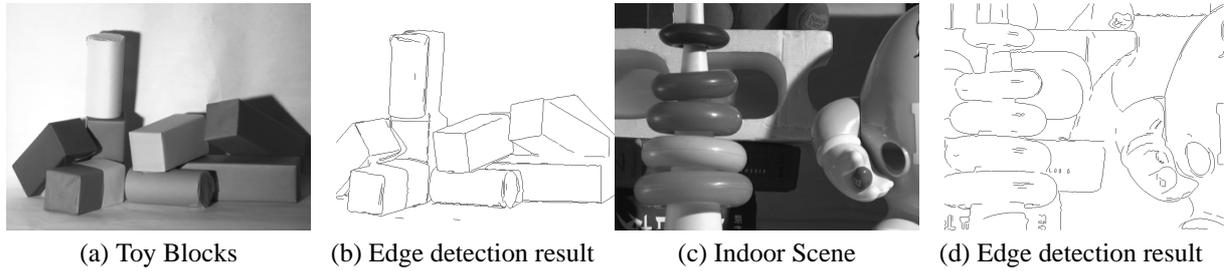


Figure 8: Images for the experiments

Table 2: Accuracy for image (a) (Experiment 1, 2)

Phenomena	Exp.1	Exp.2-a	Exp.2-b	Exp.2-c
Al/Occl	89%	94%	100%	100%
Sh/Ri	87%	87%	87%	91%
To/Co1	94%	94%	94%	94%
Co2	88%	88%	88%	100%
Average #	1	1.14	1.58	2.78

Table 4: Accuracy for image (a)(Experiment 3, 4)

Phenomena	Exp.3	Exp.4-a	Exp.4-b	Exp.4-c
Albedo	0%	100%	100%	100%
Occlusion	53%	69%	78%	95%
Shadow	50%	75%	90%	95%
Ridge	68%	79%	89%	89%
Touch	50%	50%	83%	100%
Compound1	100%	100%	100%	100%
Compound2	88%	88%	88%	88%
Average #	1	1.35	1.98	2.62

set which has high probability by considering the non-separable phenomena as a group. For a relatively complicated scene, such as (c), although the accuracy is lower, good results are achieved by Method2 as shown in 5.

6 Conclusion

In this paper, we proposed the color edge profile utilization for object arrangement discrimination. We first introduced several object arrangement models and locus models in the RGB space. To discriminate them, we defined several features, and used the evidence accumulation scheme of DS theory. In our experiments, our method showed good performances for relatively simple scenes. For complicated scenes, there is a certain trade-off concerning the accuracy and the number of candidates.

Table 3: Accuracy for image (c) (Experiment 1, 2)

Phenomena	Exp.1	Exp.2-a	Exp.2-b	Exp.2-c
Al/Oc	61%	70%	94%	100%
Sh/Ri	92%	92%	92%	92%
To/Co1	53%	53%	73%	73%
Co2	100%	100%	100%	100%
Average #	1	1.35	1.76	3.30

Table 5: Accuracy for image (c)(Experiment 3, 4)

Phenomena	Exp.3	Exp.4-a	Exp.4-b	Exp.4-c
Albedo	0%	100%	100%	100%
Occlusion	18%	45%	59%	79%
Shadow	11%	56%	70%	85%
Ridge	60%	80%	96%	100%
Touch	29%	29%	57%	86%
Compound1	50%	50%	75%	75%
Compound2	100%	100%	100%	100%
Average #	1	1.75	2.24	2.47

However, this can be improved by our candidate selection method.

References

- [1] T. Kanade. Recovery of the three-dimensional shape of an object from a single view. *Artificial Intelligence, Vol.17*, pages 409–460, 1981.
- [2] V. Lacroix. The primary raster: a multiresolution image description. *Proc. 10th ICPR*, pages 903–907, 1990.
- [3] R. Marik, J. Mates, and J. Kittler. Colour-based semantic line labelling. *Proc. ACCV'93*, pages 450–453, 1993.
- [4] B. Maxwell and S. Shafer. A framework for segmentation using physical models of image formation. *Proc. CVPR'94*, pages 361–368, 1994.
- [5] B. Maxwell and S. Shafer. Physics-based segmentation: Moving beyond color. *Proc. CVPR'96*, pages 742–749, 1996.

- [6] Y. Nakamura and M. Nagao. Edge profile extraction and its utilization for recognition (in Japanese). *Trans. Information Processing Society of Japan*, 34(10):2075–2084, 1995.
- [7] P. Perona and J. Malik. Detecting and localizing edges composed of steps, peaks and roofs. *Proc. 3rd ICCV*, pages 52–57, 1990.
- [8] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, 1976.