A FACET MODEL FOR IMAGE DATA: REGIONS, EDGES, AND TEXTURE

Robert M. Haralick

Virginia Polytechnic Institute
and State University
Blacksburg Virginia 24061

ABSTRACT

In this paper we present a facet model for image data which motivates an image processing procedure that simultaneously permits image restoration as well as edge detection, region growing, and texture analysis. We give a mathematical discussion of the model, the associated iterative processing procedure, and illustrate it with processed image examples.

1. INTRODUCTION

The world recorded by imaging sensors has order. This order reflects itself in the regularity of the image data taken by imaging sensors. A model for image data describes how the order and regularity in the world manifests itself in the ideal image and how the real image differs from the ideal image. In this paper we describe a facet model for image data and suggest some procedures for image restoration, segmenting, and texture analysis using the facet model.

J. C. Simon and R. M. Haralick (eds.), Digital Image Processing, 337–356.

The facet model for image data assumes that the spatial domain of the image can be partitioned into regions having certain gray tone and shape properties. The gray tones in a region must all lie in the same simple surface. The shape of a region must not be too jagged or too narrow.

To assure regions which are not too jagged or narrow, the simplest facet model assumes that for each image there exists a $K > 1$ such that each region in the image can be expressed as the union of $K \times K$ blocks of pixels. The value of K associated with an image means that the narrowest part of each of its regions is at least as large as a $K \times K$ block of pixels. Hence, ideal images which have large values of K have very smoothly shaped regions.

To make these ideas precise, let Zr and Zc be the row and column index set for the spatial domain of an image. For any $(r, c) \in Zr \times Zc$, let $I(r, c)$ be the gray value of resolution cell (r, c) and let $B(r, c)$ be the $K \times K$ block of resolution cells centered around resolution cell (r, c). Let $P = \{P(1), \ldots, P(N)\}$ be a partition of $Zr \times Zc$ into its regions.

In the slope facet model, for every resolution cell $(r, c) \in P(n)$, there exists a resolution cell $(i, j) \in Zr \times Zc$ such that:

(1) shape region constraint
$(r, c) \in B(i, j) \subset P(n)$

(2) region gray tone constraint
$I(r, c) = a(n)r + b(n)c + g(n)$

The actual image J differs from the ideal image I by the addition of random stationary noise having zero mean and covariance matrix proportional to a specified one.

$$J(r, c) = I(r, c) + n(r, c)$$

where

$$E[n(r, c)] = 0$$

$$E[n(r, c)\, n(r', c')] = ks(r - r', c - c')$$

The flat model of Tomita and Tsuji (1977) and Nagao and Matsuyama (1978) differs from the slope facet model only in that the coefficients a(n) and b(n) are assumed to be zero. Nagao and Matsuyama also use elongated neighborhoods with a variety of orientation. This variety of neighborhoods, of course, leads to a more general and more complex facet model. A second way of generalizing the facet model is to have the facet surfaces be more complex than sloped planes. For example we could consider polynomial or trignometric polynomial surfaces. In the remainder of this paper, we consider only the flat facet and the sloped facet model.

To illustrate the validity of the facet model we consider the image shown in Figure 1a. Using a 2x2 window and iterating with the slope facet procedure described in the next section to a fixed point, there results the image shown in Figure 1b after 20 iterations. The logarithm of the absolute value of the difference between the original and the 2x2 slope facet image is shown in Figure 1c. To see the effect of the window size, Figure 1d shows the resulting fixed point image after 16 iterations of the slope facet procedure. The logarithm of the absolute value of the difference between the original and the 3x3 slope facet image is shown in Figure 1e. It is clear that as the window size increases, the error increases and becomes more spatially correlated. Notice that most of the error occurs around region edges. In Figure 1d one can begin to see region edges becoming just a little jagged due to the fact that the 3x3 window is too large of a window for the slope facet model for this image.

These kinds of experiments have been repeated with other kinds of images with similar results. More work needs to be done to determine the best compromise between window size and complexity of the function fitting each facet. We will be reporting on such studies in a future paper.

$\hat{J}(-1,-1)$

8	5	2
5	2	-1
2	-1	-4

$\hat{J}(-1,0)$

5	5	5
2	2	2
-1	-1	-1

$\hat{J}(-1,1)$

2	5	8
-1	2	5
-4	-1	2

$\hat{J}(0,-1)$

5	2	-1
5	2	-1
5	2	-1

$\hat{J}(0,0)$

2	2	2
2	2	2
2	2	2

$\hat{J}(0,1)$

-1	2	5
-1	2	5
-1	2	5

$\hat{J}(1,-1)$

2	-1	-4
5	2	-1
8	5	2

$\hat{J}(1,0)$

-1	-1	-1
2	2	2
5	5	5

$\hat{J}(1,1)$

-4	-1	2
-1	2	5
2	5	8

Figure 1 shows the filtering masks to be used for least squares estimation of the gray value for any position in a 3 x 3 block. Each mask must be normalized by dividing by 18.

Figure la shows the original.

Figure 1b shows the fixed point 2 x 2 slope facet (20 iterations).

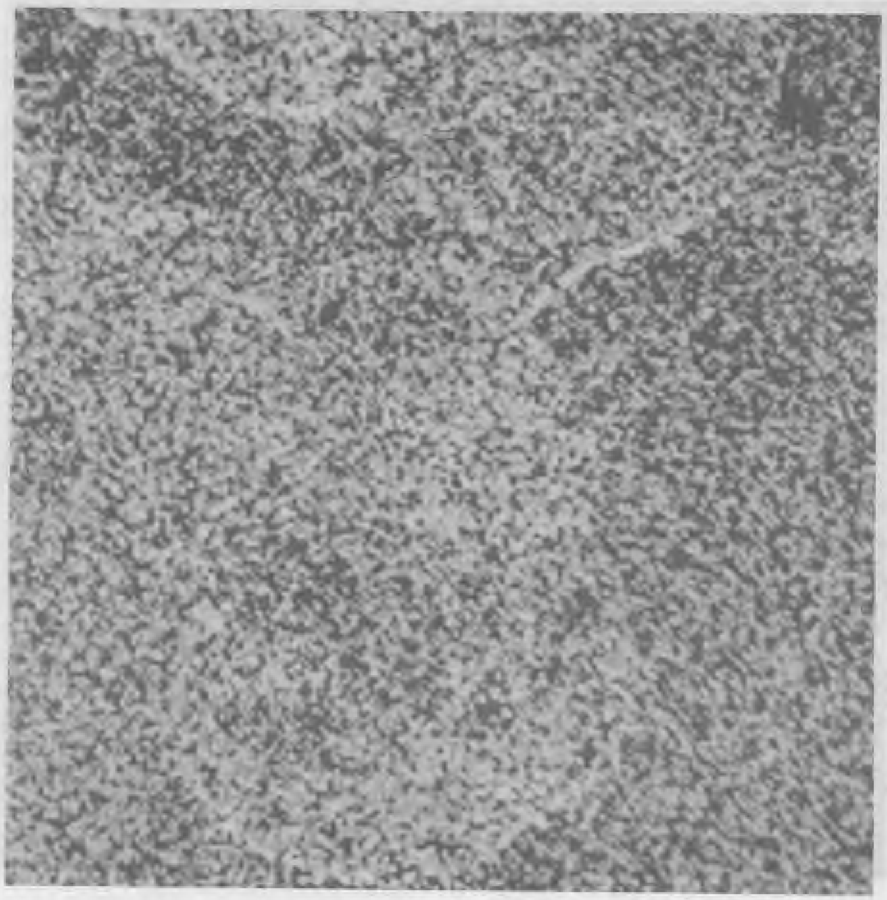

Figure 1c shows the log error at 2 x 2 fixed point (white high).

Figure 1d shows the fixed point 3 x 3
slope facet (16 iterations).

Figure le shows the log error at 3 x 3 fixed point (white high).

2. IMAGE RESTORATION UNDER THE FACET MODEL

Image restoration is a procedure by which a noisy image is operated on in a manner which produces an image which has less noise and is closer to the form of the ideal image than the observed image is. Iterative and relaxation techniques have become important techniques for achieving this (Rosenfeld, 1977; Rosenfeld, 1978; Overton and Weymouth, 1979; Lev et. al. 1977). The facet model suggests the following simple non-linear filtering procedure. Each resolution cell is contained in K^2 different K x K blocks. The gray tone distribution in each of these blocks can be fitted by either a flat horizontal plane or a sloped plane. One of the K x K blocks has smallest error of fit. Set the output gray value to be that gray value fitted by the block having smallest error. For the flat facet model this amounts to computing the variance for each K x K block a pixel participates in. The output gray value is then the mean value of the block having smalllest variance.

The filtering procedure for the sloped facet model is more complicated and we give a derivation here of the required equations. We assume that the block lengths are odd so that one of the block's pixels is its center. Let the block be (2L + 1) x (2L + 1) with the upper left-hand corner pixel having relative row column coordinates (-L, L), the center pixel having relative row column coordinates (0, 0), and the lower right-hand corner pixel having relative row column coordinates (L, L). Let J(r, c) be the gray value at row r column c. According to the sloped facet model, for any block entirely contained in one of the image regions.

$$J(r, c) = ar + bc + g + n(r, c)$$

where n(r, c) is the noise.

A least squares procedure may be used to determine the estimates for a, b, and g.

Let f(a, b, g) =

$$\sum_{r=-L}^{L} \sum_{c=-L}^{L} (ar + bc + g - J(r, c))^2.$$

The least squares estimate for a, b, and g are those which minimize f. To determine these values, we take the partial derivative of f with respect to a, and b, set these to zero and solve the resulting equations for a, b, and g. Doing this we obtain:

$$a = \frac{3}{L(L+1)(2L+1)^2} \sum_{r=-L}^{L} r \sum_{c=-L}^{L} J(r, c)$$

$$b = \frac{3}{L(L+1)(2L+1)^2} \sum_{c=-L}^{L} c \sum_{r=-L}^{L} J(r, c)$$

$$g = \frac{1}{(2L+1)^2} \sum_{r=-L}^{L} \sum_{c=-L}^{L} J(r, c)$$

The meaning of this result can be readily understood for the case when the block size is 3 x 3. Here L = 1 and

$$a = \frac{1}{6} [J(-1, *)-J(1,*)]$$

$$b = \frac{1}{6} [J(*,-1) - J(*, 1)]$$

$$g = \frac{1}{9} J(*, *)$$

where an argument of J taking the value dot means that J is summed from -L to L in that argument position. Hence, a is proportional to the slope down the row dimension, b is proportional to the slope across the column dimension, and g is the simple gray value average over the block. Figure 2 illustrates the masks that may be used to compute a, b, and g for a couple window sizes.

The fitted gray tone for any resolution cell (r, c) in the block is given by

$$H(r, c) = ar + bc + g$$

For the case where L=1,

$$H(r, c) = [J(1, *)-J(-1, *)]r/6$$

$$+ [J(*, -1) - J(*, 1)]c/6$$

$$+J(*, *)/9$$

Writing this expression out in full:

$$H(r, c) =\{J(-1, 1) (-3r - 3c + 2)$$

$$+ J(-1, 0))-3r +2)$$

$$+ J(-1, 1) (-3r + 3c + 2)$$

$$+ J(0, -1) (-3c +2)$$

$$+ J(0, 1) (3c + 2)$$

$$+ J(1, -1) (3r - 3c + 2)$$

$$+ J(1, 0) (3r + 2)$$

$$+ J(1, 1) (3r + 3c + 2)\}/18$$

This leads to the set of linear filter masks shown in Figure 2 for fitting each pixel position in the 3 x 3 block.

The sloped facet model noise filtering would examine each of the K x K blocks a pixel (r, c) belongs to. For each block, a block error can be computed by

$$e^2 = \sum_{r=-L}^{L} \sum_{c=-L}^{L} (J(r, c) - J(r, c))^2$$

One of their K^2 blocks will have lowest error. Let (r*, c*) be the coordinates of the pixel (r, c) in terms of the coordinate system of the block having smallest error. The output gray value at pixel (r, c) is then given by H(r*, c*) where H is the affine function estimating the gray values for the block having the smallest error of fit.

Haralick and Watson (1979) prove convergence of this iteration procedure for any size or set of neighborhood slopes.

3. REGION AND EDGE ANALYSIS

Edge detection and region growing are two areas of image analysis which are opposite in emphasis but identical in heart. Edges obviously occur at bordering locations of two adjacent regions which are significantly different. Regions are maximal areas having similar attributes. If we could do region analysis, then edges can be declared at the borders of all regions. If we could do edge detection, regions would be the areas surrounded by the edges. Unfortunately, we tend to have trouble doing either: edge detectors are undoubtedly noisy and region growers often grow too far.

The facet model permits an even handed treatment of both. Edges will not occur at locations of high differences. Rather, they will occur at the boundaries having high differences between the parameters of sufficiently homogeneous facets (Haralick, 1980). Regions will not be declared at just areas of similar value of gray tone. They will be the facets: connected areas whose resolution cells yield minimal differences between of region parameters. Here, minimal means smallest among a set of resolution cell groupings. In essence, edge detection and region analysis are identical problems that can be resolved with the same procedure and in this section we describe how.

Recall that the facet model iterations produce the parameters a and b. The fact that the parameters a and b determine the value of the slope in any direction is well known. For a planar surface of the form:

$$g(r, c) = ar + bc + g$$

the value of the slope at an angle t to the row axis is given by the directional derivative of g in the direction t. Since a is the partial derivative of g with respect to r and b is the partial derivative of g with respect to c, the value of the slope at angle t is [a cos t + b sin t]. Hence, the slope at any direction is an appropriate linear combination of the values for a and b. The angle t which maximizes this value satisfies

$$\cos t = \frac{a}{a^2 + b^2} \quad \text{and} \quad \sin t = \frac{b}{a^2 + b^2}$$

and the gradient which is the value of the slope in the steepest direction is

$$a^2 + b^2 .$$

The sloped-facet model is an appropriate one for either the flat world or sloped world assumption. In the flat world each ideal region is constant in gray tone. Hence, all edges are step edges. The observed image taken in an ideal flat world is a defocussed version of the ideal piecewise constant image with the addition of some random noise. The defocussing changes all step edges to sloped edges. The edge detection problem is one of determining whether the observed noisy slope has a gradient significantly higher than one which could have been caused by the noise alone. Edge boundaries are declared in the middle of all significantly sloped regions.

In the sloped facet world, each ideal region has a gray tone surface which is a sloped plane. Edges are places of either discontinuity in gray tone or derivative of gray tone. The observed image is the ideal image with noise added and no defocussing. To determine if there is an edge between two adjacent pixels, we first determine the best slope fitting neighborhood for each of the pixels by the iteration facet procedure. Edges are declared at locations having significantly different planes on either side of them.

This model does not take into account defocussing and, therefore, does not recognize whether or not highly sloped regions are edges. The determination of whether a sloped region is an edge region or not may depend on the significance and magnitude of the slope as well as the semantics of the image. One of the important research goals of the facet model is to work out this kind of problem.

In either the case of the noisy defocussed flat world, or the noisy sloped world we are faced with the problem of estimating the parameters of a sloped surface for a given neighborhood and then calculating the significance of the difference of the estimated slopes of two adjacent neighborhoods. To do this we proceed in a classical manner. We can use a least square procedure to estimate parameters and we can measure the strength of any difference by an appropriate F-statistic.

In summary, the facet image restoration iteration procedure produces more than just a restored gray tone. For each pixel, it also produces the a, b, and g parameters. Using these parameters we can determine whether or not neighboring pixels lie in the same connected facet. Because the parpmeters come not form the pixel's central neighborhood but from the pixel's best neighborhood, the determination of whether two pixels lie in the same facet requires that the parameters for each pixel be taken out of their individual relative coordinate systems and be placed in the same coordinate system. Linking together neighboring pixels with the same a, b, g paramenters, permits us to identify the facets which are characterized by the connected sets of pixels that constitute them. These facets become the regions and edges are the boundaries between regions.

4. TEXTURE ANALYSIS

Following Haralick (1979) textures can be classified as being weak textures or strong textures. Weak textures are those which have weak spatial-interaction between the texture primitives. To distinguish between them it may be sufficient to only determine, for each pair of primitives, the frequency with which the primitives co-occur in a specified spatial relationship. In this section we discuss a variety of ways in which primitives from the facet model can be defined and the ways in which relationships between primitives can be defined.

4.1 PRIMITIVES

A primitive is a connected set of resolution cells characterized by a list of attributes. The simplest primitive is the pixel with its gray tone attribute. Sometimes it is useful to work with primitives which are maximally connected sets of resolution cells having a particular property. An example of such a primitive is a maximally connected set of pixels all having the same gray tone or all having the same edge direction.

Gray tones and local properties are not the only attributes which primitives may have. Other attributes include measures of shape of a connected region and homogeneity of its local property. For example, a connected set of resolution cells can be associated with its length or elongation of its shape or the variance of its local property.

Attributes generated by the facet model include the a, b, and g parameters plus the average error of fit for the facet. These attributes can be used by themselves or used to generate derived attribute images such as that created from a^2+b^2. The relative extreme primitives can be defined in the following way:

Label all pixels in each maximally (minimally) connected relative maxima (minima) plateau with an unique label. Then label each pixel with the label of the relative maxima (minima) that can reach it by a monotonically decreasing (increasing) path. If more than one relative maxima (minima) can reach it by a monotonically decreasing (increasing) path, then label the pixel with a special "c" for common. We call the regions so formed the descending components of the image.

4.2 SPATIAL

Once the primitives have been constructed, we have available a list of primitives, their center coordinates, and their attributes. We might also have available some topological information about the primitives, such as which are adjacent to which. From this data, we can select a simple spatial relationship such as adjacency of primitives or nearness of primitives and count how many primitives of each kind occur in the specified spatial relationship.

More complex spatial relationships include closest distance or closest within an angular window. In this case, for each kind of primitive situated in the texture, we could lay expanding circles around it and locate the shortest distance between it and every other kind of primitive. In this case our co-occurrence frequency is three dimensional, two dimensions for primitive kind and one dimension for shortest distance. This can be dimensionally reduced to two dimensions by considering only the shortest distance between each pair of like primitives.

Co-occurrence between properties of the descending components can be based on the spatial relationship of adjacency. For example, if the property is size, the co-occurrence matrix could tell us how often a descending component of one size is next fo a descending component of another size.

To define the concept of generalized co-occurrence, it is necessary to first decompose an image into its primitives. Let Q be the set of all primitives on the image. Then we need to measure primitive properties such as parameter value, region, size shape, etc. Let T be the set of primitive properties and f be a function assigning to each primitive in Q a value for each property of T. Finally, we need to specify a spatial relation between primitives such as distance or adjacency. Let SC Q x Q be the binary relation pairing all primitives which satisfy the spatial relation. The generalized co-occurrence matrix P is defined by:

$$P(t1, t2) = \frac{\#\{(q1, q2) \in S \mid f(q1)=t1 , f(g2) = t2\}}{\#S}$$

P(t1, t2) is just the relative frequency with which two primitives occur with specified spatial relationship in the image, one primitive having property t1 and the other primitive having property t2.

Zucker (1974) suggests that some textures may be characterized by the frequency distribution of the number of primitives any primitive has related to it. This probalility p(k) is defined by:

$$p(k) = \frac{\#\{(q \in Q \mid \#S(q) = k\}}{\#Q}$$

Although this distribution is simpler than co-occurrence, no investigation appears to have used it in texture discrimination experiments.

5. CONCLUSION

In this paper we considered the gray tones of an image to represent the height of a surface above the row-column coordinates of the gray tones. The observed image is then the surface of the underlying ideal image plus random noise. The ideal image is composed of a patchwork of constrained surfaces sewed together.

We called each patch a facet and in the ideal image, the facets must satisfy the constraints of the facet model for image data: the facet model constrains the shape of each facet to be exactly composed as a union (possibly over-lapping) of a given set of neighborhood shapes and it constrains the surface to be a sloped plane surface (or some other more general polynomial surface).

The goal of image restoration is to recover the ideal gray tone surface which underlies the observed noisy gray tone surface. Although the noise prevents recovering the precise underlying ideal surface, we can recover that gray tone surface which is the "closest surface" to the observed noisy surface and which also satisfies the facet model constraints.

The procedure we suggest for recovering the underlying surface consists of iterating best neighborhood least squares fits.
Associated with each given pixel is a set of all the neighborhoods of given shapes that contain it. Each one of these neighborhoods can be fitted with the best fitting surface. One of these neighborhoods will have a best fitting surface whose residual error is smallest among all the neighborhoods tried. The parallel iterative procedure consists of replacing each pixel gray tone intensity with the height of the best fitting surface in its lowest residual error neighborhood. The procedure is guaranteed to converge and actually achieves essential convergence in a few iterations. The resulting image is an enhanced image having less noise, better contrast, and sharper boundaries.

Image restoration is not the only use of the facet model. The facet model processing provides us with additional important information. By collecting together all pixels participating in the same surface facet, we transformed the pixel as our processing and analysis unit into the surface facet as our processing and analysis unit. Now edge boundaries, for example, can be defined to occur at the shared boundary of all neighboring facets whose surface paramenters are significantly different. Homogeneous regions can be defined by linking together all those neighboring surface facets whose parameters are significantly the same. Texture can be characterized by the co-occurrence statistics of neighboring primitives which are not the pixel gray tones as in the usual occurrence approach but which are the facets characterized by their boundary, shape, size, and surface parameter attributes.

Our paper has been mainly theoretical laying out a variety of uses of the facet model in image processing. Future papers will describe our experimental results.

REFERENCES

Haralick, R. M. 'Statistical and Structural Approaches to Texture', Proceeding of the IEEE Vol 67, No. 5, May 1979, p 786-804.

Haralick, R. M. and L. Watson 'A Facet Model For Image Data', Proceeding IEEE Conference on Pattern Recognition and Image Processing Chicago, Ill., August 1979, p 489-497.

Haralick R. M., 'Edge and Region Analysis for Digital Image Data', Computer Graphics and Image Processing Vol. 12, No. 1, January 1980, p 60-73.

Lev, A., S. W. Zucker, and A. Rosenfeld, 'Iterative Enhancement of Noisy Images', IEEE Transactions on Systems, Man, and Cybernetics Vol. SMC-7, No. 6, June 1977, p 435,442.

Nagao, M. and T. Matsuyama, 'Edge Preserving Smoothing', Computer Graphics and Image Processing Vol 9, 1979, p 394-407.

Overton K. J., and T. E. Weymouth, 'A Noise Reducing Preprocessing Algorithm' Proceeding IEEE Conference on Pattern Recognition and Image Processing Chicago, Ill., August 1979, p 498-507.

Rosenfeld, A. 'Iterative Methods In Image Analysis', Proceedings IEEE Conference on Pattern Recognition and Image Processing Troy, N.Y., June 1977, p 14-18.

Rosenfeld, A. 'Relaxation Methods in Image Processing and Analysis', Proceedings of the 4th International Joint Conference on Pattern Recognition Kyoto, Japan, November 1978, p. 181-185.

Tomita, F. and S. Tsuji, 'Extraction of Multiple Regions by Smoothing in Selected Neighborhoods', IEEE Transactions on Systems Man, and, and Cybernetics Vol SMC-7, No. 2, February 1977 p. 107-109.