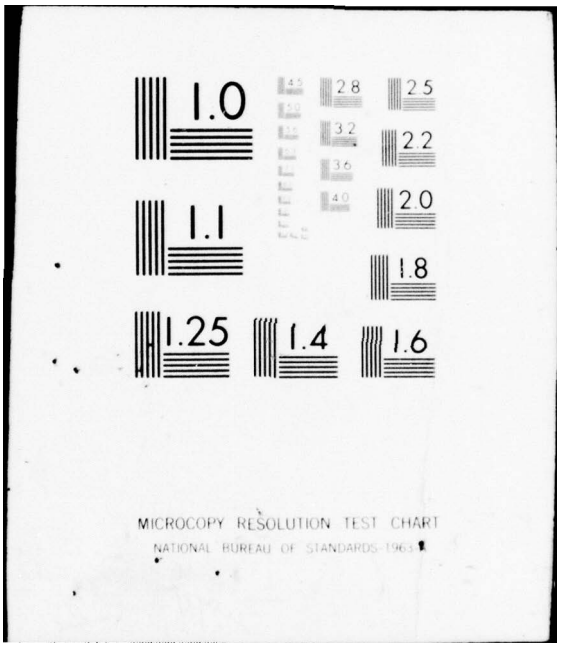
AD-A037 772 CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF COMPUTER --ETC F/G 14/2
SATURATION, HEU, AND NORMALIZED COLOR: CALCULATION, DIGITIZATIO--ETC(U)
NOV 76 J R KENDER F44620-73-C-0074

UNCLASSIFIED AFOSR-TR-77-0328 NL

1 OF 1
AD A037772

END

DATE FILMED
4-77


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Saturation, Hue, and Normalized Color: Calculation, Digitization Effects, and Use

John R. Kender

November 1976

DEPARTMENT
of
COMPUTER SCIENCE

Carnegie-Mellon University

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DDC
This technical report has been reviewed and is approved for public release IAW AFR 190-12 (7b).
Distribution is unlimited.
A. D. BLOSE
Technical Information Officer

Saturation, Hue, and Normalized Color:

Calculation, Digitization Effects, and Use

John R. Kender

November 1976

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pa. 15213

This work was done while the author was supported by the Defense Advanced Research Projects Agency under contract F44620-73-C-0074, monitored by the Air Force Office of Scientific Research. Partial support was provided by an IBM Fellowship.

Abstract

Three transformations used in the analysis of tricolor natural scenes are analyzed. All have nonremovable singularities, near which they are highly unstable. Given digital input, the distribution of their transformed values is highly nonuniform, characterized by spurious modes and gaps. These effects are quantified and illustrated. In addition, a significantly faster algorithm for hue is derived. Image segmentation techniques of edge detection, region growing, clustering, and region splitting are affected arbitrarily badly by such problems. Some stratagems are illustrated that help minimize the bad behavior. Linear transformations are presented as a generally favorable alternative to these three nonlinear ones.

1. Introduction

Among the several approaches to the understanding of natural scenes are those which are based upon the analysis of multispectral inputs. Generally these methods, taking a cue from the human visual system, assume a three-dimensional spectral vector for each spatial image coordinate (pixel), with red, green, and blue filtration of the scene providing the values of the respective vector coordinates. Several of these methods further attempt to borrow from human color perception studies by creating at each pixel an additional two-vector through the application of nonlinear transformations loosely called "saturation" and "hue", after their psychological analogues [2, 4, 5, 7, 11]. Occasionally, an additional two-vector of normalized red and normalized green is also generated at each pixel, using the nonlinear transformations upon which saturation and hue are based [1, 2, 4, 10]. Typically, analysis of the scene proceeds by the segmentation of the image into regions of "uniformity" along one of the above measures. Regions are found directly by the techniques of clustering, region growing, or region splitting; indirectly, they are described by their boundaries using the technique of edge detection.

This paper explores some of the problems which arise when these segmentation techniques are applied to the data that results from the above transforms. It is seen that the above four transformations are singular at some points and are unstable at many others, causing the generated values to respond arbitrarily poorly to small input variations. In addition, if the input data is digitized, the transformations introduce anomalies into the frequency distribution of their generated values, even given smooth input; no quantization of this output is sufficiently coarse to guarantee the removal of the resulting spurious modes and gaps. As all the above segmentation techniques are sensitive to the uniformity of a measure, the use of these transforms with any one of them can result in arbitrarily bad segmentations.

Some stratagems are suggested which lessen these intrinsic problems, and both problems and solutions are illustrated with examples. In addition, a significantly faster algorithm for calculating hue is derived. It is shown that the usual requirement that hue values be quantized can be turned to even further computational advantage.

"Color" can be defined and computed in many ways. One such method, which uses linear transformations, is briefly explored as an alternative to the above measures. Although this latter method has several advantages, it is seen that some cautions should be observed in its use, also.

2. Assumptions and Notations

It is assumed that the red, green, and blue data have been digitized only after any necessary signal-normalizing preprocessing has been performed on the raw quantized data. (For example, corrections for unequal filter densities or inverse-square law effects. These operations, often not required, usually generate floating point results). Thus, the red, green, and blue data will be assumed to be integers in the range of 0 (least reflectance) to M (most reflectance); typically $M = 2^m-1$, with $4 \le m \le 8$. Although no assumptions are made about the distribution of values along this range, it is normally the case that the digitization is performed so as to utilize as much of the available span of values as possible. It is also assumed that all the transformed values are also digitized in order to save space and subsequent processing time; this also appears to be the norm. Digitization is accomplished in the natural way. Values are multiplied by a value S (for Scale factor) and rounded; usually $S = 2^s-1$, for some s. Note that this scale factor usually depends on the transformation.

In the following sections, the following conventions will be used. The red, green, and blue coordinates of a pixel (the "tristimulus values") will be represented by their initial letter in upper case. Normalized coordinates ("chromaticity coordinates": the tristimulus values divided by the sum of tristimulus values at a pixel) are represented by their initial letter in lower case. Thus, if $R = G = B = 30$, then $r = g = b = 1/3$. "Brightness" is defined to be mean coordinate value, that is, $(R+G+B)/3$. The values of saturation, hue and normalized color at a pixel are distinguished from their scaled and digitized counterparts by always specifying the latter quantities with some reference to their digital nature. (Most of the discussion, however, is in terms of their actual real values). Multiplication is usually denoted in the text by juxtaposition; for clarity, however, it will occasionally be represented by the programming symbol "*". Program segments are written in an Algol-like language.

3. Saturation

The "saturation" transformation is so called since its calculation of excitation purity is analogous to the psychological phenomenon of the perception of saturation. The formula [Tenenbaum] is:

```
saturation := 1 - 3 min(r, g, b)
```

In terms of actual pixel values instead of normalized coordinates, the formula becomes:

```
saturation := 1 - 3 min(R, G, B)/sum(R, G, B)
```

As given, the value of saturation has a range of the closed unit interval. If saturation is to be digitized, then, where S is 2^s-1, for some as yet unspecified s (output byte size):

```
digitalsaturation := round(S * saturation)
```

The digital saturation has a range of from 0 (for pixels whose tristimulus values are equal and non-zero) to S (for pixels with at least one zero coordinate and at least one non-zero coordinate). It is not clear, however, despite the transformation's elegance, what problems may arise in the use of saturation in natural scene analysis.

3.1. Essential Singularity

It should first be noted that this transformation is ill-conditioned. It has an essential singularity at R = G = B = 0. The formula is undefined at this point in tristimulus space; further, this singularity is not removable, as the following shows.

Let sat(R, B, G) be the saturation of a pixel with coordinates (R, B, G). Consider sat(x, x, rx) where $0 \le r \le 1$, that is, pixels along the line through the white point and pure yellow. Then, as the third coordinate is always the minimum, it is seen that:

$$\lim_{x \to 0} \text{sat}(x, x, rx) = 2(1-r)/(2+r)$$

But, as $0 \le r \le 1$, this limit is 0 if r is 1, 1 if r is 0, and, since the expression is continuous in the unit interval, it attains every other value in the range of saturation for appropriate values of r. Thus, sat(0, 0, 0) cannot be uniquely defined to remove the singularity.

The psychological analogue to this problem occurs when a pure red stimulus fades in luminance to pure black: saturation remains total until nothing is seen. If continuity is to be preserved, black should then be considered totally saturated. But the same problem occurs when any other stimulus fades to pure black, even an achromatic one (which would imply a saturation of zero for black).

The black points, then, have the disturbing property that they cannot contribute to any analysis of the image based upon the use of saturation as a metric; their saturation is undefined. Any decision procedure based on the concept of feature value distance will be incapable of uniquely assigning these points to any specific region. Black points are no further in saturation from any one pixel type as they are from any other, including other black points. In particular, saturation edges are undefined at these points; clustering, growing, or splitting techniques (statistical techniques) have no distance values on which to make discriminations at these points, either.

Further, because of the nature of the singularity, any attempt at ad-hoc definition of the saturation of black is bound to fail. For example, consider:

```
saturation :=  if max(R, B, G) = 0 then
                   black
               else
                   (standard computation, . .)
```

No matter what value is arbitrarily chosen for "black" inside the range 0 to S, there will be instances where any decision procedure based on saturation values arbitrarily includes or excludes these points from a given figure or ground. (See Section 6.)

Black must be handled by assigning it a unique value unattainable by all other input triples; that is, it is a special value treated separately throughout. One way to do this would be to restrict S to 2^s-2, and let "black" be 2^s-1, for some s. Decisions about such black points must then be made differently from those involving valid saturation values; some use of information about neighboring pixels appears necessary, if at all possible.

It should be noted that in practice black points might not be a problem. Red, green, and blue may have sufficient dynamic range, and the scene enough illumination, so that black points are negligibly few. However, a narrow dynamic range of values, or an image with high contrast (perhaps as a result of preprocessing) may have many pixels whose tristimulus values are all zeros, and will then suffer these difficulties.

3.2. Effects of Input Perturbations

The transformation's essential singularity causes other problems; the transformation is intrinsically unstable near the singularity, as expected. The minimal change of -1 to any coordinate can cause the saturation to jump from a total saturation of 1 to a saturation that is undefined; this occurs in the case of (1, 0, 0) and its permutations. In addition, the minimum possible perturbation of ± 1 in any one of the coordinates can cause the output, while remaining well-defined, to change from extreme to extreme. This is seen at the pixel (R, B, G) = (0, 1, 1), which has the well defined total saturation of 1. However, (R, B, G) = (1, 1, 1) is totally unsaturated and its saturation is zero. Further, the change from (R, B, G) = (0, x, x) to (1, x, x), $x \geq 1$, changes the saturation value from 1 to (2x-2)/(2x+1), an expression whose values fall throughout the entire range of saturation values.

As the transformation is invariant with respect to permutation of its coordinates, the effects can be obtained in any coordinate. As sensor noise or minor variations in a region's reflectance can easily cause such perturbations (as well as much larger ones), such instabilities are annoyingly significant, and not easily avoidable.

3.3. Spurious Modes

There are other difficulties with this transformation which relate more directly to its effect on the distribution of its computed values, given a smooth input distribution. These problems partially arise from the requirement that, if the resulting saturation values are to be digitized, the choice of S is usually fixed before analysis begins. It is not clear which value for S is optimal. In fact, the above problems, being intrinsic to the transform, suggest that given any choice, certain pathologies will remain.

The essential problem with the distribution of these transformed values arises from the use of a division in their calculation. Because the input is digitized, certain output quotients are "favored" over others, creating false modes in the digitized saturation distribution. Certain other output values are, in fact, impossible to attain; this creates false gaps. This effect is to be distinguished from the comb-like distribution that occurs when digitized data is overscaled; the effect occurs even when the output scale is the same as the input scale. The basic cause is that fractions are equivalence classes of pairs of integers, and, given the restrictions on the input, certain of these equivalence classes can have more representatives than others. For example, the saturation value of 1/2 (min/sum = 1/6), can be attained, in various fractions, by many more input triples than can a saturation value of 508/511 (min/sum = 1/511).

A complete statistical analysis of the distribution of output values is very difficult given the complex nature of the distributions of the red, green, and blue reflectances that arise in natural scenes. Nonetheless, some of the troublesome features of the transformation can be pointed out using a limiting assumption on the input. (Other aspects will later be shown independently of any such assumption).

Suppose that the distribution of red, green, and blue values is uniform in the tristimulus space. That is, for all $0 \leq x, y, z, \leq M$:

$$\text{Prob}\{(R, G, B) = (x, y, z)\} = 1/(M+1)^3$$

This distribution is found in a minimum entropy picture; such images almost assuredly do not exist in nature, even given contrast enhancement. But it does seem to predict the behavior of the spurious digital saturation modes encountered in the natural scenes examined so far. As will be seen, this is principally because the more prominent of the modes generated from the uniform distribution are sharply delimited on their two sides by substantial gaps. (In fact, the larger the mode, the wider the gaps). As the values in the gap cannot occur given an input distribution which encompasses every conceivable pixel, they surely do not occur given any other

distribution. Thus these gaps must appear in natural scenes as well, and if the modal value they delimit is actually attained in the natural image (a likely occurrence), a spurious modes results. Further predictive power of this assumption derives from the fact that, given the knowledge that a mode is spurious, this distribution provides a zeroth order (constant term) approximation to both the location and the relative magnitude of other nearby spurious modes.

It can be shown that under this hypothesis, then, that false modes occur in the saturation distribution where min(R, B, G)/sum(R, B, G) is expressible as a very simple fraction in lowest terms (call this fraction n/d). The expected count of pixels having a saturation value corresponding to this ratio is approximately proportional to (d-3n)/((d-2n)(d-n)). (See Appendix A). Since this value decreases when either d or n or both are incremented (or nearly always: see Appendix A), the count of pixels with saturation values corresponding to fractions with small integer numerators and denominators is proportionately higher than the count at other fractions.

The digital saturation distribution of the hypothetical uniform tristimulus distribution is shown in figure 3-1. The false modes occur most strongly at those saturation values corresponding to simple min/sum fractions, at the values given in appendix A (e.g. saturation = 1, .25, .4, .5, etc.). Relative magnitudes are also in close agreement to the analytically derived values.

The applicability of the assumption of uniformity to natural scenes is demonstrated by figures 3-2 to 3-4, which are the digital saturation distributions of natural scenes. False modes, appearing as spikes, also occur at the values predicted, with approximate predicted relative magnitudes (as seen in comparison with figure 3-1).

3.4. Spurious Gaps

The false modes in the distribution are accompanied by false gaps, one on either side of the mode. Depending on the scale factor allowed the digital saturation, they can extend over several digital units above and below the modal value. This result is independent of the assumptions made of the input distribution (uniformity is not required), and the extent can be quantified.

Gaps above and below a mode at a min/sum ratio of n/d are due to the digital granularity of the input: only certain fractions can be formed. The gap represents the distance from the modal value to the nearest attainable fraction. This latter can be thought of as a small perturbation of the numerator and/or denominator of a large member of the modal equivalence class. It can be shown that the extent of the gap bordering the spurious mode at n/d is approximately of length $(3/2)(S/M)(d-n)/d^2$. (See Appendix A). Again, this value is maximized for simple fractions n/d, as is verified by figures 3-1 through 3-4 (at saturation = 1, 0, .25, .4, etc.).

The above discussion suggests that any image having totally saturated pixels in it will show a mode at total saturation, separated by a gap of length (3/2)(S/M) from

the rest of the distribution. If $S \geq 2M/3$, such a gap would be apparent in the digital distribution. But such a gap is an artifact of the choice of S; no other image with totally saturated pixels could ever possibly have a smaller one.

Thus, S should be set no larger than M/2 (or so). This result is purely a negative one, in that this choice of S does not guarantee good behavior; it simply avoids the bad behavior that is guaranteed if S were larger. That this choice of S often works well in removing spurious gaps from the digital saturation distributions of large natural scenes is shown in figures 3-5 through 3-7, which are the rescaled distributions of the same scenes as figures 3-2 to 3-4.

Two considerations cloud the issue, however. The first is the following. Natural images, given filter imperfections and the reflectance characteristics of natural objects, tend to be unsaturated. Thus, often the worst behaved spurious mode discussed above, the one at total saturation, does not appear. Its special case behavior, then, does not disturb the distribution; the next most difficult mode is at $n/d = 1/3$ (zero saturation). This generates a gap of $(1/3)(S/M)$, which is significantly smaller, and which is no longer guaranteed when $S \leq 2M$.

The second consideration is more significant. It is impossible to reverse the sense of the negative result. That is, there is no value S, however small, which guarantees that no spurious mode will appear. Several observations bear this out. Best case conditions rarely occur. Often in an image, the range of pixels does not span the allotted range of from 0 to M. Even if they do (or, even if M is replaced by M', the true maximum value attained), in a given subarea of an image the pixel values may attain a local maximum that is considerably less. Gaps in this specific region then would be, inversely, considerably greater, and spurious modes more pronounced. Thus, in anything less than the total image, saturation is very likely to behave worse than the theoretical best case. This is particularly true of regions of low intensity; here all three pixel coordinates are small, and spurious modes and gaps predominate. As the subdistribution contributes additively to the full distribution, an image with a sufficiently large area of low intensity regions will still be beset with spurious modes.

(It should be noted that the above analysis only indirectly addresses the problem of those spurious modes which are delimited, though less distinctly, by valleys (of which gaps are an extreme case). This further complication of the distribution, evident in the figures, is of significance to those methods that work directly with the distribution histogram (clustering and region-splitting) and demands further caution in their use of saturation.)

Solutions are difficult when this transformation is used alone. If S is set before analysis begins, some logic must be employed to be aware of these effects when the regions under study become a smaller and less intense subset of the image. If saturation is generated on a regional basis, some logic must be employed to adjust either the value S or the method of handling saturation as a measure, in correspondence with the local maximum. In either case, these problems complicate the analysis, making the intensity-normalized aspect of saturation a mixed advantage in segmentation.

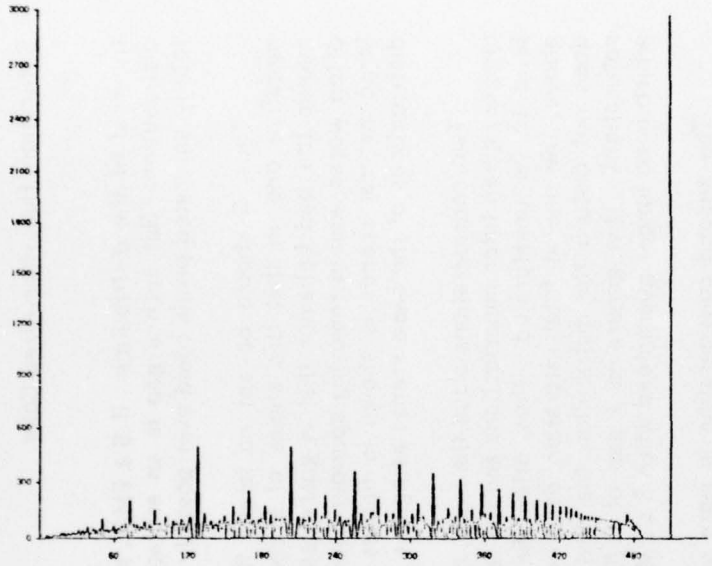


FIGURE 3-1. DIGITAL SATURATION OF UNIFORM INPUT, M = 31, S = 511, BLACK NOT SHOWN

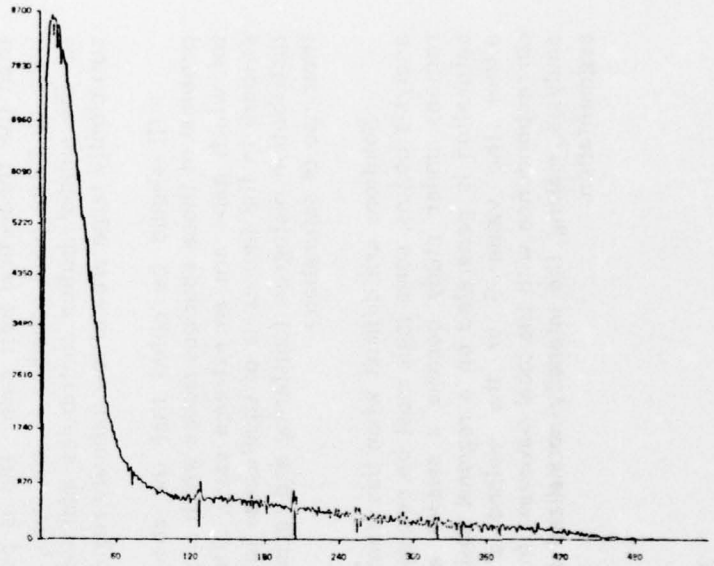


FIGURE 3-2. DIGITAL SATURATION OF OHLANDER BEAR SCENE, M = 255, S = 511, BLACK NOT SHOWN

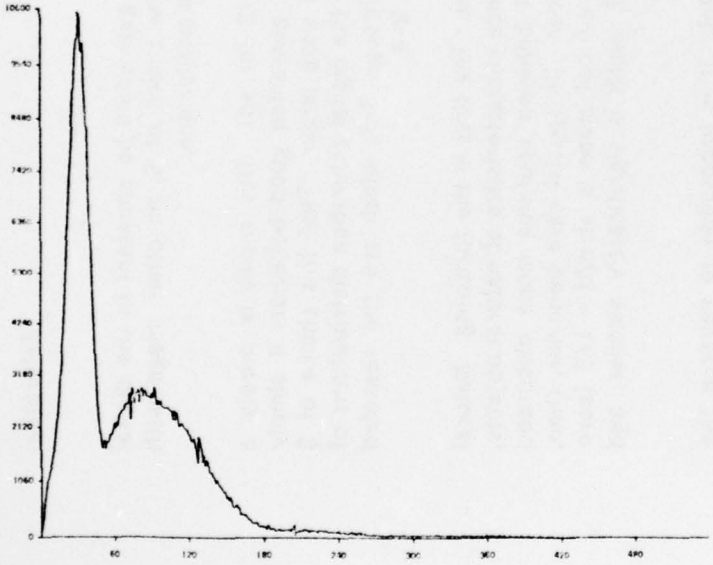


FIGURE 3-3. DIGITAL SATURATION OF OHLANDER AUTO SCENE, M = 255, S = 511, BLACK NOT SHOWN

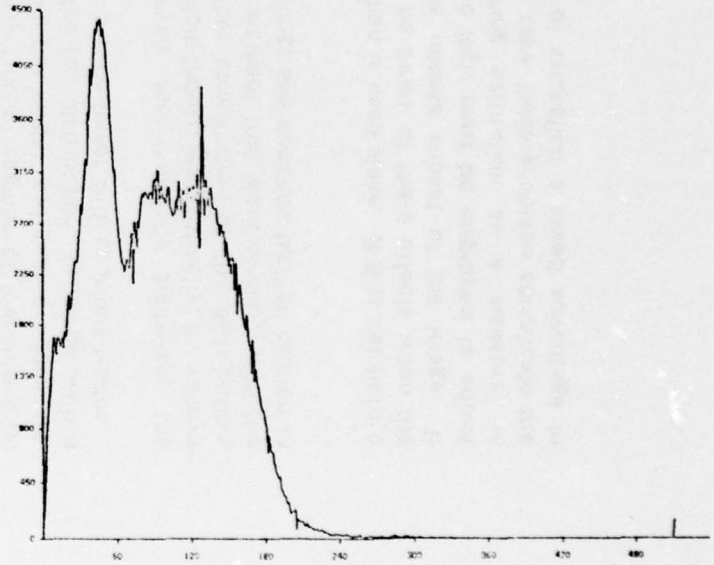


FIGURE 3-4. DIGITAL SATURATION OF OHLANDER CITY SCENE, M = 255, S = 511, BLACK NOT SHOWN

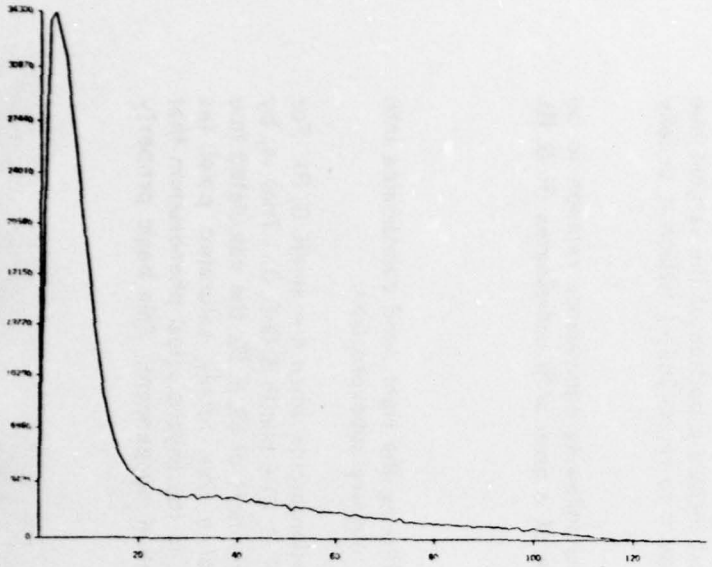


FIGURE 3-5. DIGITAL SATURATION OF OHLANDER BEAR SCENE, M = 255, S = 127, BLACK NOT SHOWN.

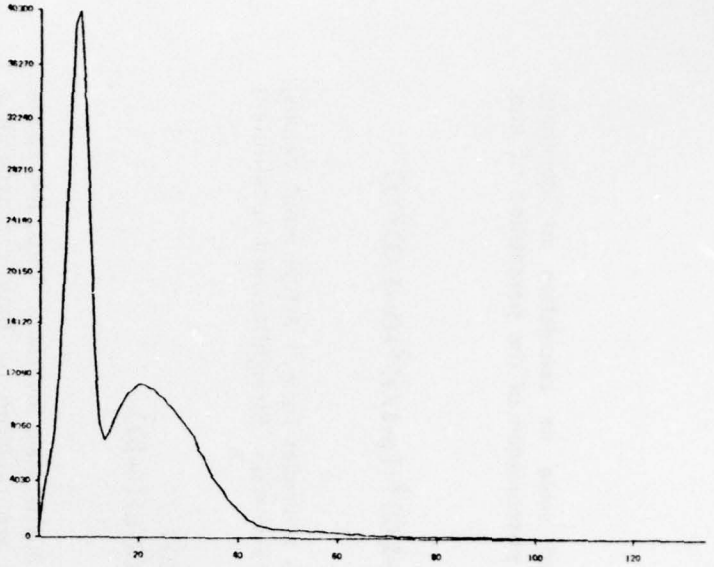


FIGURE 3-6. DIGITAL SATURATION OF OHLANDER AUTO SCENE, M = 255, S = 127, BLACK NOT SHOWN.

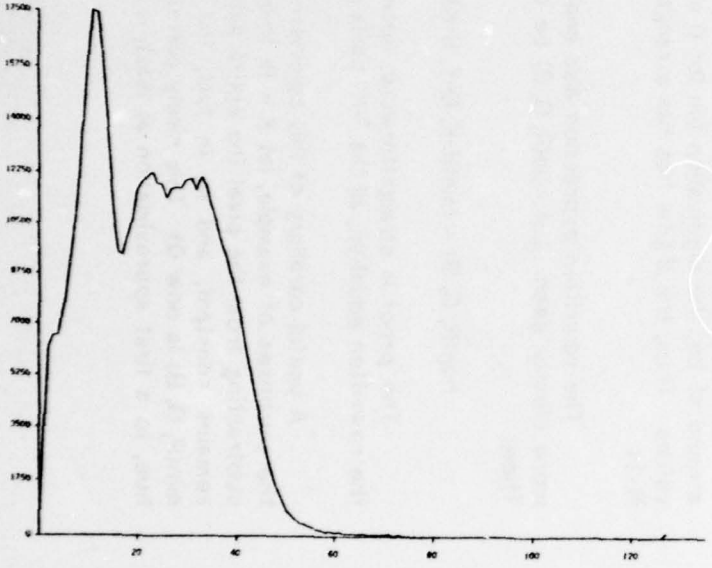


FIGURE 3-7. DIGITAL SATURATION OF OHLANDER CITY SCENE, M = 255, S = 127, BLACK NOT SHOWN.

4. Hue

The "hue" transformation is so called since its calculation of dominant wavelength is analogous to the psychological phenomenon of the perception of hue. Its formula [Tenenbaum] is:

```
hue := arccos((2r-g-b) /
                (sqrt(6) sqrt((r-1/3)^2+(g-1/3)^2+(b-1/3)^2)))
if b > g then hue := 2pi-hue
```

The above formula can be expressed in a simpler form if actual input values, rather then normalized values, are used in its calculation. Straightforward substitution yields:

```
hue := arccos((1/2)((R-G)+(R-B)) /
                sqrt((R-G)(R-G)+(R-B)(G-B))
if B > G then hue := 2pi-hue
```

In this form, it is more easily seen that hue is scale invariant. Thus, the above formula is also valid if all tristimulus values are replaced by their corresponding chromaticity coordinates.

As given, the value of hue ranges from 0 to 2π. If hue is to be digitized, then:

```
digitalhue := round(S * hue) mod round(S * 2pi)
```

Usually S is of the form $N/(2\pi)$, for some N. The inclusion of the mod function is necessary to properly handle those hue values which lie near 2π. As hue wraps around at 2π, the digitization bin for 0 also must include a portion of the largest hue values. Thus, the digital hue has a range of from 0 to round(2πS)-1 (which is usually N-1).

The rewritten expression also enables the following equivalence relation to be more clearly seen. Let hue(R, G, B) be the hue of a pixel with coordinates (R, G, B). Then:

hue(R, G, B) = hue(R-K, G-K, B-K), $\forall$K.

The proof is straightforward: upon substituting the right hand coordinates into the rewritten equation, all the "-K" parts cancel in every subexpression.

A useful corollary of this equivalence relation occurs when K = min(R, G, B). For the purposes of example, let K = B; then hue(R, G, B) = hue(R-B, G-B, 0). That is, by subtracting from the pixel the entire white component of (B, B, B), the calculated hue remains constant, and is, in fact, the hue of a now totally saturated pixel (as min(R, G, B) is now 0). This nicely corresponds to the psychological phenomenon that hue, to a first approximation at least, is saturation independent. This basic property

will allow an easier study of this transformation's basic properties (and, in fact, suggests a more efficient algorithm for its calculation: See Appendix B).

4.1. Essential Singularity

Like saturation, this transformation is ill-conditioned. However, its essential singularity is along an entire line in tristimulus space, at R = B = G. The formula is undefined anywhere along this axis, as is easily seen from the rewritten formula. As before, this singularity is not removable as the following shows.

Consider hue((1-r)x, rx, 0), where $0 \le r \le 1$ (that is, totally saturated pixels along the red-green line of the color triangle). Then:

$$\lim_{x \to 0} \text{hue}((1-r)x, rx, 0) = \arccos((2-3r)/(2\text{sqrt}(1-3r+3r^2)).$$

But, as $0 \le r \le 1$, hue((1-r)x, rx, 0) = 0 if r is 0, $2\pi/3$ if r is 1, and, as the expression is continuous in the unit interval (the discriminant of the quadratic is negative, thus the fraction is always defined), it attains every value in between for appropriate values of r. Thus, hue(0, 0, 0) cannot be uniquely defined to remove the singularity.

The problem is actually more severe. By permuting the above coordinates, it can be shown that, like the singularity of the saturation transformation, the limit can take on every value in the range of hue values (0 to 2π). Further, by the equivalence relation, it is seen that the singularity at hue(x, x, x), for all x, is also not removable.

The psychological analogue to this problem occurs when a tricolor stimulus, two of whose sources are equal, is brought to an achromatic white by adjusting the third source. The apparent hue remains constant until no hue is perceived. Continuity would require the hue of the resulting achromatic stimulus, then, to be equal to the initial choice of hue. But this choice is infinitely variable.

Achromatic points pose the same problems to analyses based on this feature as do the black points of the saturation feature. Being undefined, they cannot contribute to any analysis based upon hue as a metric. (It may be noted here in passing that it is difficult to conceive of a "hue edge," due to the manner in which hue is defined. Hue distances must be calculated on a circle, not a line; the problems of wraparound are not easy to resolve when using an edge detector). As with saturation, ad-hoc definitions of achromatic points also fail, and can give arbitrarily bad results when used in segmentations. The only way of accurately dealing with them is to assign them a unique value. If, for example, hue is digitized by converting radians to degrees before rounding, thus giving digital hue a range of 0 to 359 (9 bits), "achromatic" can be defined as 511.

Although the black points which trouble saturation can be relatively rare, it is much harder to imagine any image, or any preprocessed image, that is totally free of achromatic ones. Whereas only one pixel value, (0, 0, 0), had to be avoided for

saturation, and entire line of pixels in tristimulus space is at fault here. The problem is more acute at low dynamic ranges that at high ones; in these cases proportionately more of the possible pixel types are achromatic ones. However, within any digitization range, whites, greys, and blacks still abound in natural scenes, especially since deeply saturated colors seem to be relatively rare in nature.

4.2. Effects of Input Perturbations

Like saturation, hue also is intrinsically unstable near its singularity. The minimum change of ± 1 in any coordinate can cause hue to jump from any one of six well defined hues (pure red, green, blue, and their complements) to one that is undefined; this occurs at (x, x, x-1), (x, x, x+1), and their permutations. In addition, the minimum possible change of ± 1 in any one of the coordinates can cause the output to change as much as $\pi/3$. For example, pixels of the form (x+1, x, x) have a hue of zero; when the second coordinate is incrementally perturbed, hue becomes $\pi/3$. Decrementing the first coordinate of (x+1, x+1, x) has the opposite effect. Further, pixels of the form (x+1, x, 0), $x \geq 0$, when perturbed by +1 in the second coordinate, exhibit a wide range of hue changes depending on x. The resulting pixel always has hue = $\pi/3$. However, the original pixel has hue = $\arccos((1/2)(x+2)/\mathrm{sqrt}(x^2+x+1))$. That is, its hue is determined by a function that runs continuously from x = 0 (where hue = 0) to x = M (where hue approaches $\pi/3$); thus the hue changes lie in the interval 0 to $\pi/3$. By the equivalence relation, this also holds for the many pixels of the form (n+x+1, n+x, n), for all n, and all their permutations.

Input perturbations are easily caused by sensor noise or minor reflectance variations. The large class of pixels that are so affected make the instabilities of hue less avoidable, and even more significant than the similar problems of saturation.

4.3. Spurious Modes

The formula for hue in Appendix B is helpful in analyzing the non-linear behavior of the transformation. Since the function has, except for the added constant "baseline" terms, a basic sixfold symmetry (as does the normalized color triangle), the basic behavior can be studied in one such sixth. Let $R \geq G \geq B$. This is actually somewhat larger than a sixth, as it includes the boundaries and the singularity R = G = B, but the white point can be handled separately. In this region:

```
hue := pi/3+arctan(sqrt(3) (G-R)/(G-B+R-B))
```

As with saturation, the use of a normalizing division causes a highly non-uniform distribution of hue values, even given a uniform distribution of red, green, and blue values. False modes with neighboring false gaps, in fact, occur much more unpleasantly with hue: they are more pronounced and harder to remove.

Assume again a uniform distribution over the tristimulus space. (The justification is the same as with saturation). Under this hypothesis, false modes also occur in hue

distributions where (G-R)/(G-B+R-B) is expressible as a simple fraction in lowest terms (call this fraction n/d). The expected count of pixels having a hue value corresponding to this ratio is approximately proportional to 1/(n+d) if n and d are of the same parity, and half that value if they are not. (See Appendix A). Since this value decreases when either d or n or both are incremented (or nearly always: see Appendix A), the count of pixels with hue values corresponding to fractions with small integer numerators and denominators is proportionately higher that the count at other fractions. Again, this corresponds to the observation that, given limited input, more simple fractions can be formed than exotic ones.

A digital hue histogram of the hypothetical tristimulus uniform distribution is shown in figure 4-1. False modes are more pronounced than the false digital saturation modes. The false modes occur most strongly at those hue values corresponding to simple fractions for (G-R)/(G-B+R-B) and its two analogous ratios, at the values given in Appendix A (e.g. hue = $n\pi/3$, $(2n+1)\pi/6$, $(2n+1)\pi/6\pm\alpha$, etc.). Relative magnitudes are also in close agreement to the analytically derived values. The applicability of the assumption of uniformity to natural scenes is demonstrated by figures 4-2 to 4-4, which are digital hue histograms of natural scenes. False modes, appearing as spikes, also occur at the values predicted, with approximate predicted relative magnitudes (as seen in comparison with figure 4-1).

4.4. Spurious Gaps

The false modes in the distribution are accompanied by false gaps, one on either side of the mode. Depending on the scale factor allowed the digital hue, they can extend over several digital units above and below the modal value. This extent can be quantified, and again, this result is independent of the assumptions made of the input distribution; uniformity is not required.

Hue gaps are caused by the same digitization effects that create saturation gaps: due to the input granularity, only certain fractions can be formed. It can be shown that the extent of the gap bordering the spurious mode at a (G-R)/(G-B+R-B) ratio of n/d is approximately of length $\mathrm{sqrt}(3)(S/M)(n+d)/(d^2+3n^2)$ if n and d are of the same parity, and half that value if they are not. (See Appendix A). This value is maximized for simple fractions n/d, as is verified by figures 4-1 through 4-4 (at hue = $n\pi/3$, $(2n+1)\pi/6$, $(2n+1)\pi/6\pm\alpha$, etc.).

The above discussion suggests that any image having any pure primary colors or their complements (hue = $n\pi/3$) will show a mode at their corresponding hue value, separated by a gap of length $(\mathrm{sqrt}(3)/2)(S/M)$ from the rest of the distribution. Thus, if $S \geq 1.16M$, such a gap would be apparent in the digital distribution. Again, such a gap is artificial.

Therefore, S should be set no larger than M (or so). Again, this is a purely negative result; it avoids otherwise assured difficulties but does not guarantee good behavior. The use of even 1/4 of this choice of S is shown in figures 4-2 through 4-4. The persistence of spurious modes and gaps in parts of the distributions is noticeable.

Several reasons account for their recalcitrance. As with saturation, the sense of the negative result cannot be reversed; no value of S guarantees an absence of spurious modes. Best case conditions rarely occur. Subimages attain lower maxima than the full image and are therefore locally worse behaved. In addition, natural scenes tend to be unsaturated and this further effectively decreases the local maxima. (Through the equivalence relation, hues are determined by the amounts that the maximum and median coordinate of a pixel are in excess over the minimum coordinate; low saturation implies these amounts will be very small. Put another way, the tristimulus values of the hue-equivalent, totally saturated pixel (R', G', B') are very small, implying a very small local M). The additive contribution of the subdistributions of sizable areas of low saturation, then, generates a full image hue distribution which is still characterized by spurious modes.

(Again, this analysis only indirectly addresses the problem of spurious valleys, a further complication evident in the figures.)

Solutions are more difficult for hue, when used alone, than with saturation; the equivalence relation and the extended singularity make hue more ill-conditioned. Any logic to handle the effects of the locally varying maximum would have to be even more sophisticated. Thus, the normalization with respect to both intensity and saturation that characterizes hue is also a mixed advantage in segmentation.

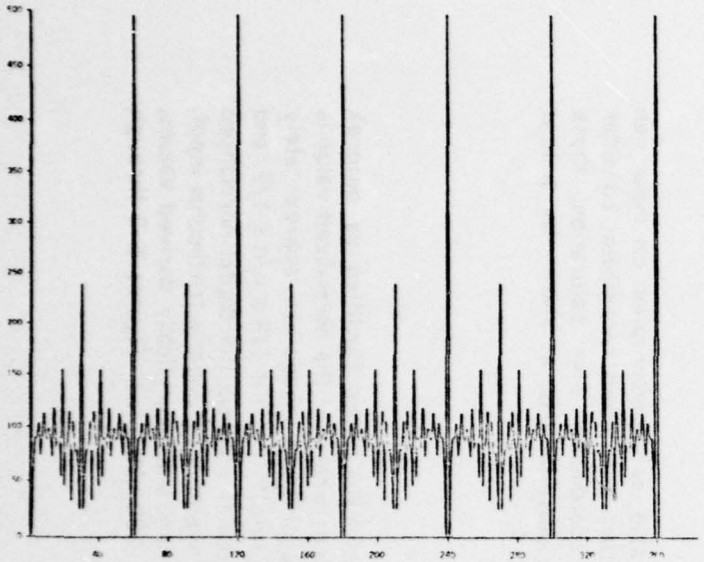


FIGURE 4-1. DIGITAL HUE OF UNIFORM INPUT, M = 31, S = 360/(2PI), ACHROMATIC NOT SHOWN

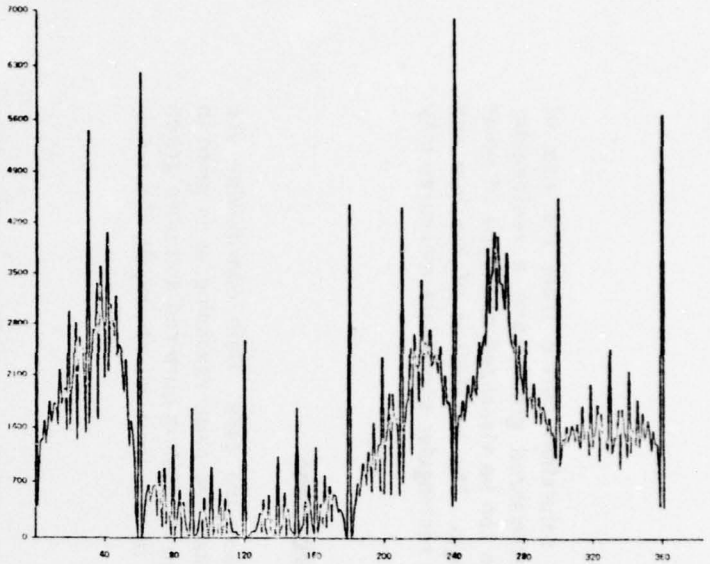


FIGURE 4-2 DIGITAL HUE OF OHLANDER BEAR SCENE, M = 255, S = 360/(2PI), ACHROMATIC NOT SHOWN

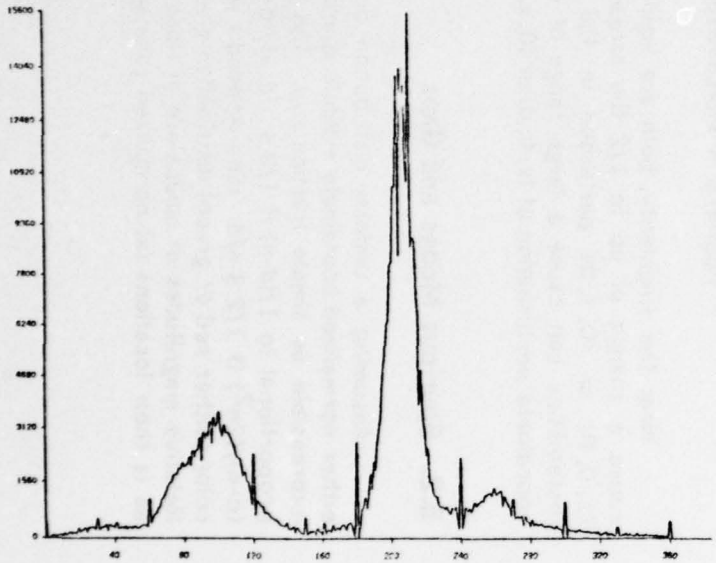


FIGURE 4-3 DIGITAL HUE OF OHLANDER AUTO SCENE, M = 255, S = 360/(2PI), ACHROMATIC NOT SHOWN

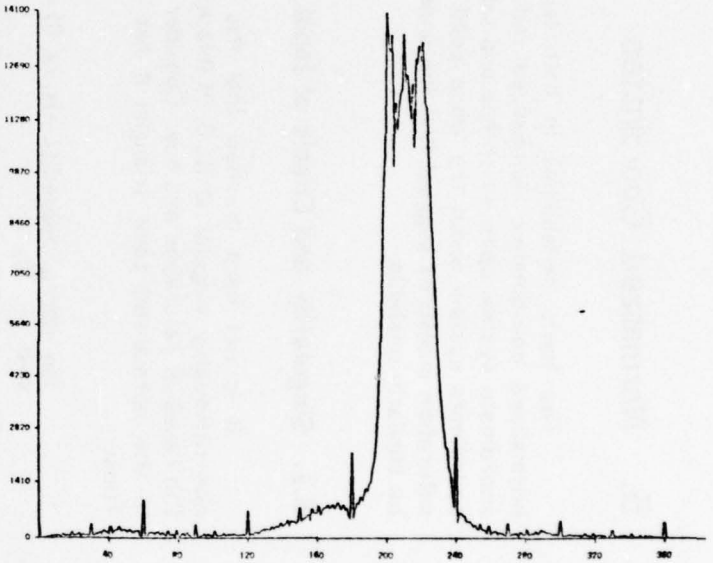


FIGURE 4-4 DIGITAL HUE OF OHLANDER CITY SCENE, M = 255, S = 360/(2PI), ACHROMATIC NOT SHOWN

5. Normalized Coordinates

The basic instabilities in both hue and saturation derive from the use of normalized coordinates. Normalized red and normalized green form a rectangular coordinate system upon which hue and saturation can be visualized as a type of polar coordinate system about the white point as origin. The instabilities of the hue and saturation measures suggest that the underlying rectangular coordinate measures may be similarly unstable.

5.1. Singularity and Effects of Input Perturbations

It is not hard to show that this is indeed the case. Both coordinates are non-removably singular at (0, 0, 0) (black), by much the same reasoning as is seen in the cases of saturation and hue. Consider the line connecting pure red and pure green in the normalized color triangle; it has pixels of the form ((1-r)x, rx, 0), $0 \le r \le 1$. Thus:

$$\lim_{x \to 0} \text{normalizedred}((1-r)x, rx, 0) = 1-r$$

The limit in the case of normalized green is r. Both singularities are therefore nonremovable, and black is thus unrepresentable in this system of coordinates. To properly calculate digital normalized red, the following, where "black" is outside the range 0 to S, is necessary:

```
digitalnormalizedred :=  if max(R, G, B) = 0 then
        black
    else
        round(S * (R/(R+B+G))
```

Near the singularity, both are highly unstable, as a perturbation on input can cause a change of up to 1/2 the range of either normalized coordinate; consider (1, 0, 0) or (0, 1, 0) perturbed in the first coordinate. Like saturation, these instabilities can cause a large range of output perturbations, as shown by a first coordinate perturbation of (x, 0, 0) or (0, x, 0), $x \ge 1$.

5.2. Spurious Modes and Gaps

Assuming a uniform distribution over tristimulus space (justified as before), either normalized coordinate exhibits spurious modes wherever the normalized value is expressible as simple fraction n/d. The heights of these modes are approximately proportional to $1/(d-n)$ if $1/3 \le n/d$, $(1/(d-n))(1-(3n-d)^2/(2n^2))$ if $1/3 \le n/d \le 1/2$, and $(d-n)/(2n^2)$ if $1/2 \le n/d$. (See Appendix A). Figure 5-1 shows the digital normalized color (either red or green) distribution given the hypothetical uniform tristimulus input. Relative magnitudes of modes are in close agreement with analytically derived values, as is their locations (at normalized color = 0, .5, .333, .25, etc.). Figures 5-2 through

5-4 are the digital normalized red distributions of natural scenes; spurious modes and their relative heights are comparable to those predicted in figure 5-1.

Spurious gaps bordering the mode at n/d are approximately of length $(1/2)(S/M)(d-n)/d^2$ if $n/d \leq 1/3$, and $(S/M)(n/d^2)$ otherwise. This value is maximized for simple fractions n/d, as is verified by figures 5-2 through 5-4 (at normalized color = 1, 0, .5, .667, etc.). The value for S above which these gaps are guaranteed to appear in the digital distribution is approximately M; the effect of such a setting is shown in figures 5-5 through 5-7.

Cautions concerning the use of normalized color are very much the same as those put forth for saturation and hue. Best case conditions rarely occur. (Valleys also persist in the distribution). When this transformation is used alone, some additional logic is needed to adjust for varying local maxima. Again, normalization with respect to intensity is an advantage not without its difficulties.

5.3. Distribution of (Normalized Red, Normalized Green) Ordered Pairs

The analysis of the two-dimensional distribution of ordered pairs of the form (r, g) is difficult, even given the hypothesis of tristimulus uniformity. However, spurious modes and gaps appear in the projections of this distribution onto either of the normalized color axes. Therefore, it must be true that the two dimensional distribution is also troubled by false modes and gaps. Figure 5-8 shows that, under the uniformity hypothesis, this is indeed so.

Similarly, the analysis of the distribution of (hue, saturation) pairs is difficult. But as its projections onto either the hue or saturation axis is characterized by spurious modes and gaps, it itself must also be so characterized. Such is the case, under uniform input, as seen in figure 5-9.

Both these results have implications for higher dimensional methods of segmentation; see Section 6.

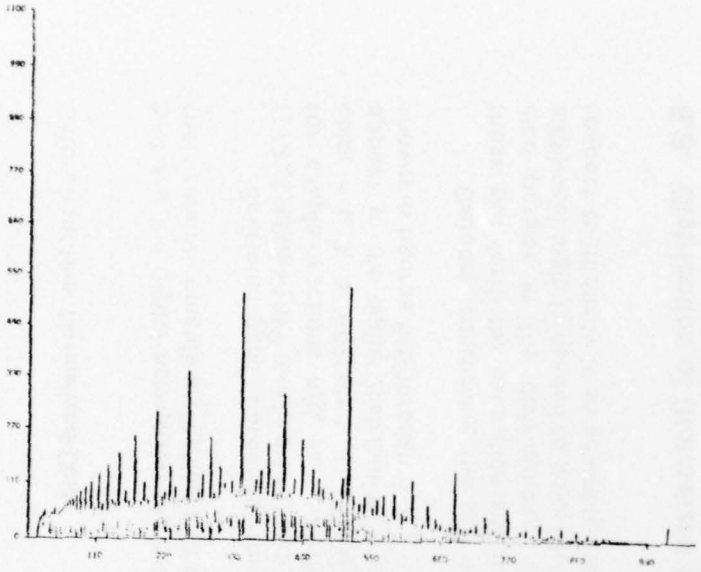


FIGURE 5-1 DIGITAL NORMALIZED RED OF UNIFORM INPUT, M = 31, S = 1023, BLACK NOT SHOWN

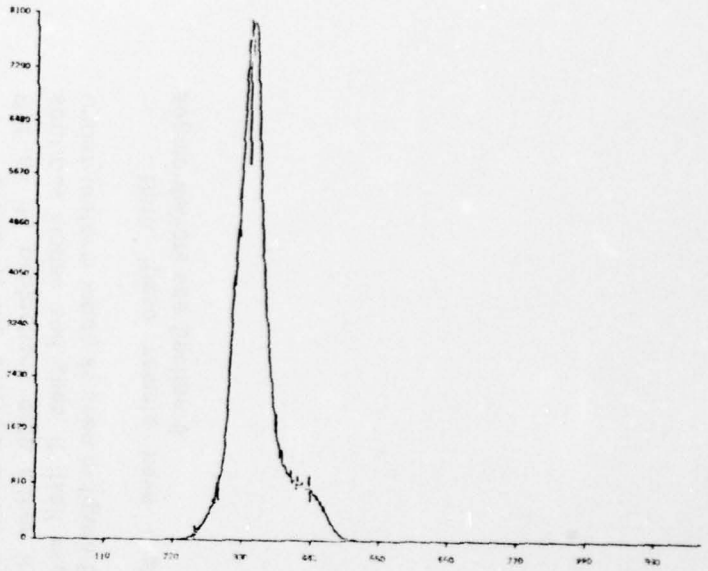


FIGURE 5-2 DIGITAL NORMALIZED RED OF OHLANDER BEAR SCENE, M = 255, S = 1023, BLACK NOT SHOWN

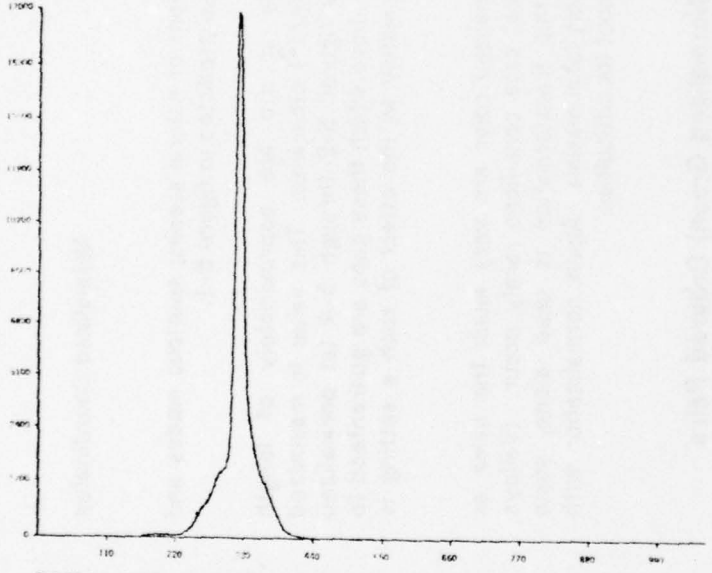


FIGURE 5-3 DIGITAL NORMALIZED RED OF OHLANDER AUTO SCENE, M = 255, S = 1023, BLACK NOT SHOWN

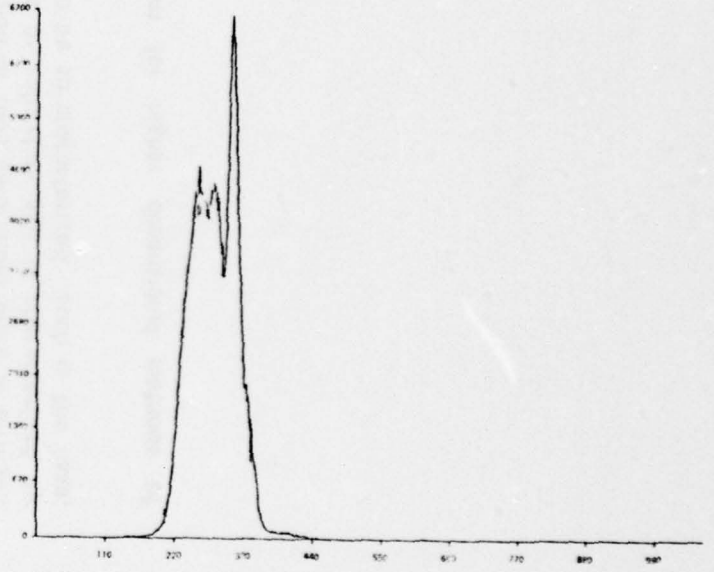


FIGURE 5-4 DIGITAL NORMALIZED RED OF OHLANDER CITY SCENE, M = 255, S = 1023, BLACK NOT SHOWN

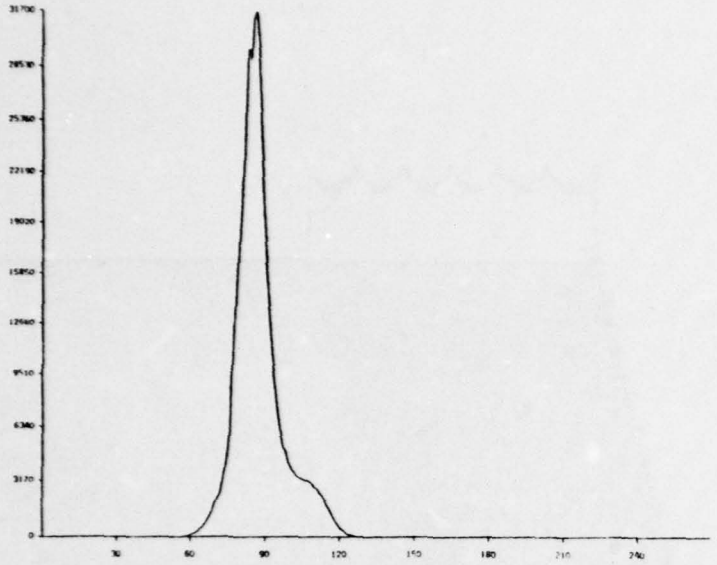


FIGURE 5-5. DIGITAL NORMALIZED RED OF OHLANDER BEAR SCENE, M = 255, S = 255, BLACK NOT SHOWN.

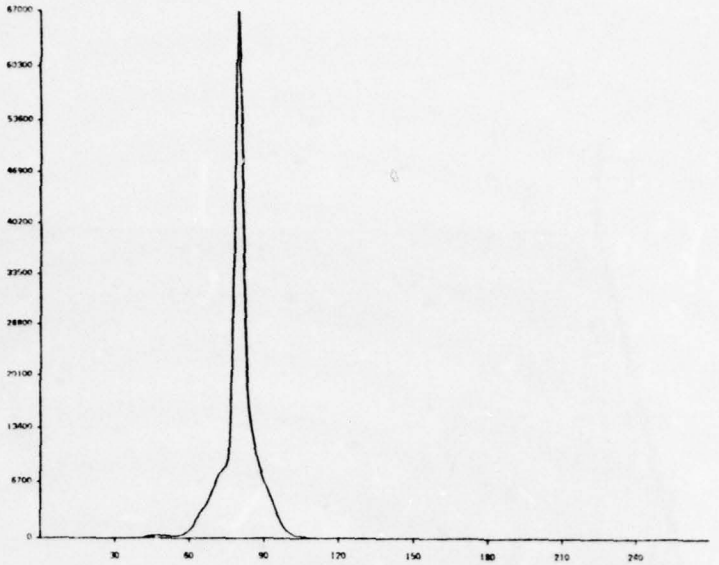


FIGURE 5-6. DIGITAL NORMALIZED RED OF OHLANDER AUTO SCENE. M = 255, S = 255, BLACK NOT SHOWN.

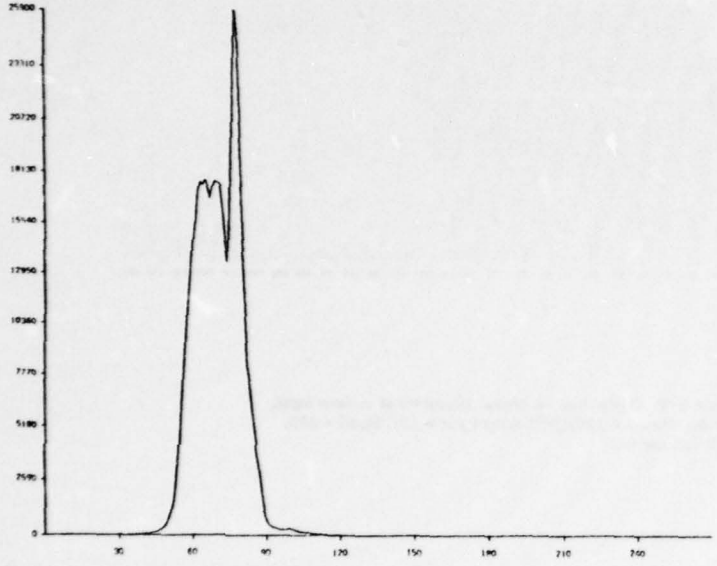


FIGURE 5-7. DIGITAL NORMALIZED RED OF OHLANDER CITY SCENE, M = 255, S = 255, BLACK NOT SHOWN.

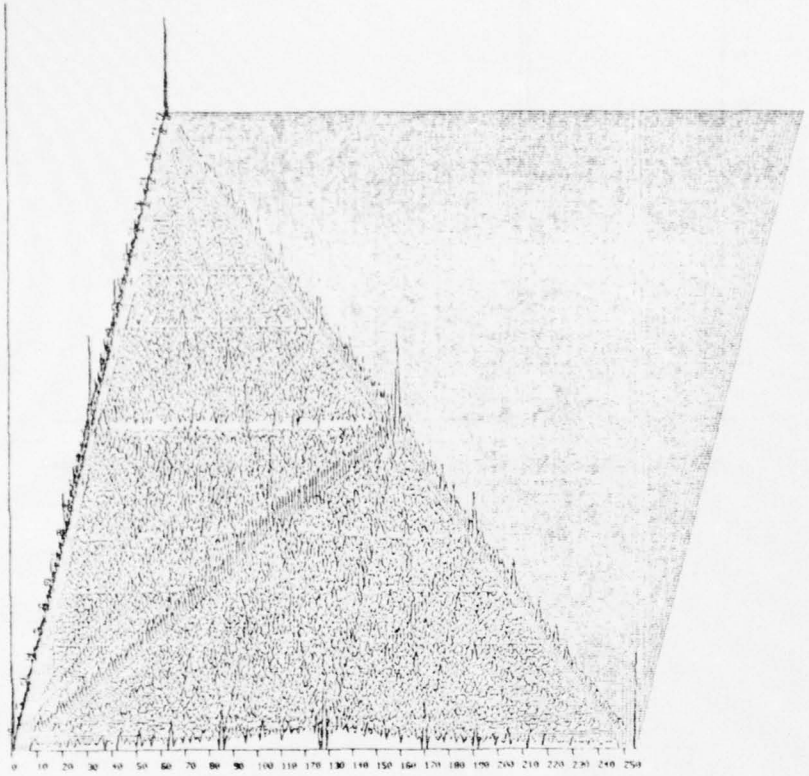


Figure 5-8: Digital normalized red vs. digital normalized green of uniform input, M = 31, S = 255, black not shown.

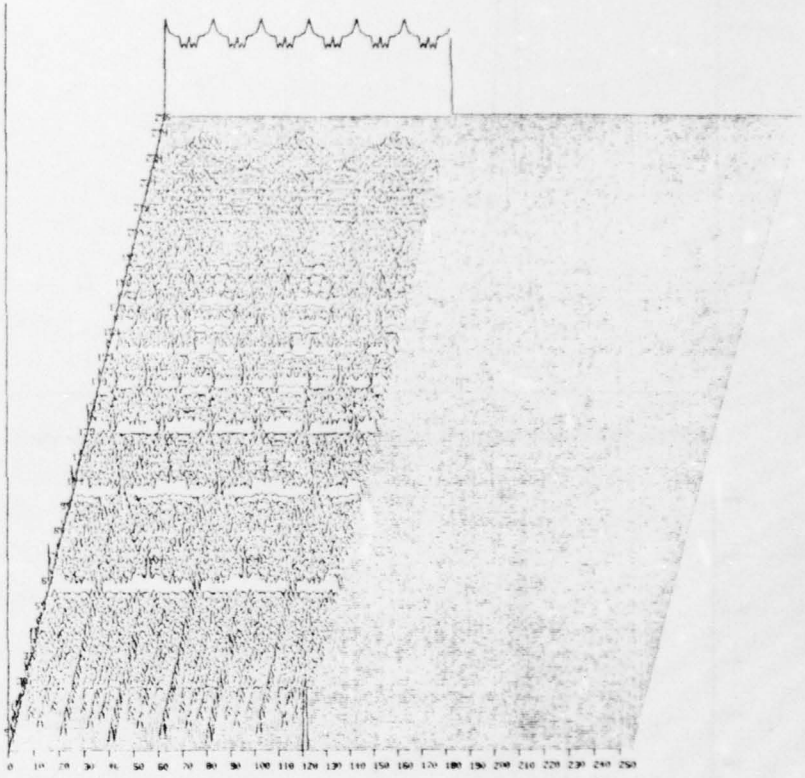


Figure 5-9: Digital hue vs. digital saturation of uniform input, M = 31, S(hue) = 120/(2PI), achromatic = 121, S(sat) = 255, black not shown.

6. Handling the (Intrinsic) Difficulties

The analytic bad behavior of the four transformations is not merely a theoretical issue. Singularities and digitization errors deeply affect the four segmentation techniques of edge detection, clustering, region splitting, and region growing. Arbitrarily bad segmentations can result from errors induced by these transformations' undefined values, unstable spectral neighborhoods, and spurious modes and gaps.

6.1. Use in Segmentation: What Goes Wrong

It is impossible to smoothly redefine the values of the transformations at their singularities. Ad-hoc definitions that attempt to locate these singular points somewhere along the range of the transformation will always cause some segmentations to fail. As an example, if black (the saturation singularity) is considered to have zero saturation, then a dark green area (low in reflectance, but spectrally pure) obtained by thresholding above a certain saturation value (or by region growing using this value as a seed, or by detecting saturation edges, or clustering) will exclude these points, although many may well belong to the region and, in fact, appear black because of sensor noise. Other examples would show any other assignment for this singularity causes analogous problems, and likewise for the singularities of hue and normalized color. Again, the singular points are not "close" to any other points, and must be treated separately (or avoided altogether).

Unstable areas around the singularities impact edge detection by generating spurious strong edges from very small tristimulus variations. Clustering and region splitting are similarly impaired by the scattering throughout the transform space of points which vary little in tristimulus values. Thus, a region of spectral uniformity may never be detected. Although edge detection is not seriously affected by spurious modes and gaps, those methods which work directly with the statistical distributions of the transforms are. In clustering or region splitting, spurious modes and gaps may lead to spurious groupings of pixels, and thus to spuriously delimited regions. (It should also be noted that the use of any of these four transformations as an axis for n-dimensional clustering or region splitting will also create difficulties, regardless of the choice of measure for the other axes. Cognizance must be taken of both the singularity-induced instabilities and the digitization-induced preferences for simple fractions along that axis, and, by inheritance, within the entire distribution.)

The technique of region growing also is sensitive to the irregularities of these transformations. A given object can have a range of pixels that would produce a tight cluster in tristimulus space, and yet (especially if it is dark and/or unsaturated) span the entire range of saturation, hue, or normalized color. Therefore, any criterion for stopping region growth must be too severe to grow these objects. Likewise, dissimilar adjacent objects may permit regions to grow across object boundaries if one object's spectral distribution is near a singularity. This latter condition generates a large range of transformed values, some of which must necessarily overlap the other object's (possibly) well-behaved ones. Lastly, spurious gaps can cause premature cessation of a region's growth, as pixels within a certain range of transform values may not exist at

all, due to digitization effects. These phenomena may explain the failure of hue and saturation to provide meaningful segmentations in a recent attempt of their use in region growing [6].

6.2. Ameliorating Stratagems

It should first be pointed out that all four transformations are intrinsically unstable. Saturation, hue, and normalized color all must have singularities as they are all scale invariant and non-constant [f(R, G, B) = f(kR, kG, kB), ∀k > 0]. Secondly, as all four employ a central normalizing division, all four must respond to digital input with spurious modes and gaps in their resulting output. No rewritting of their formulae, no use of higher precision, no refinement of any programming technique will eradicate either problem: the transformations, as given, are not well-posed. However, specific care in their use, and a somewhat broader view of what they attempt to measure are two stratagems for minimizing both classes of difficulties.

6.2.1 Avoiding the Singularities

It is desirable that small tristimulus changes cause small transform changes. Such behavior does obtain in these transformations away from their singularities. Recalling the definition of brightness as mean coordinate value, the central formulae of the four transformations can be rewritten as:

```
saturation :=  . . . 1 - min(R, G, B) / brightness

hue :=  . . . pi/3 + arctan(sqrt(3)(G-R) /
                          (brightness * saturation)) . . .

normalizedcolor :=  . . . (color/3) / brightness
```

The above indicates that saturation and normalized color are normalized with respect to brightness, and hue is normalized with respect to both brightness and saturation, in close agreement with psychological phenomena. However, unlike the psychological phenomena of the perception of saturation only above certain levels of brightness, and of hue only above certain levels of brightness and saturation, no such singularity-avoiding behavior is evident in the above expressions. It can, however, be artificially imposed by altering the use of the transformations by segmentation techniques.

The basic idea is that of ordering the application of the measures. Brightness is applied first, then saturation (or normalized color), and lastly hue. For various techniques, this takes the following forms. In region growing, first regions are grown using brightness; saturation is used to grow subregions only in those regions of uniform high brightness; hue is a third pass on bright, deeply saturated regions. Region splitting is similar, with thresholding permitted along the saturation feature only if and where there has been a thresholding at high brightness (or evidence of an

absence of low brightness pixels); hue is similar. Clustering would recluster bright pixels along saturation, and recluster for hue those of high brightness and deep saturation. Edge detection would first seek brightness edges, then in areas of high brightness, saturation edges. (Hue edges are undefined. Perhaps normalized color, sought in bright regions, would compensate).

Two comments are in order. First, this increases computation time, especially for the one pass algorithms of clustering, region growing and edge detection (if it is indeed possible to augment this last technique). It appears that only region splitting, intrinsically recursive in nature, could easily integrate such a priority scheme into its control flow without added expense; it already requires a similar hierarchical scheme for the selection of optimal modes.

Secondly, "high" brightness or "deep" saturation can be quantified, after a fashion. Although there are evident no sharply defined or preferred cutoff values (as there are none, apparently, in psychological phenomena), the formulae can be used to estimate the effectiveness of a particular brightness or saturation cutoff in stabilizing the subsequently applied transform. As the above expressions indicate, the larger the brightness (and saturation) values at a pixel, the smaller the effect of a tristimulus perturbation on the calculated output. As a first approximation, consider brightness to be a value high enough so that it (and saturation * brightness) can be considered constant under small input perturbations. Recall that "S" represents each digital transformation's scale factor, and need not be the same value for all three; "S" is really "$S_{transformation}$". Now, it can be seen that a ± 1 change in any coordinate can cause a maximum perturbation of S_s/brightness in digital saturation, sqrt(3)S_h/(brightness * saturation) in digital hue, and (1/3)S_n/brightness in digital normalized color.

It is reasonable to define "high" brightness as that value which guarantees, in response to a unit input change, an output change of no more than one unit of output digitization scale. Then "high" in the case of saturation is brightness $\geq S_s$, and, for normalized color, brightness $\geq S_n/3$. A similar reasonable definition of "deep" saturation requires that saturation $\geq$ sqrt(3)S_h/brightness. Depending on the values of each S, however, any of these reasonable definitions may be impossible to attain.

6.2.2 Undigitizing the Input

The above suggestions do not address the effects of the use of digitized input. As figure 6-1 shows, avoiding the singularities of hue does not minimize irregularities in the distribution, even when the input is assumed to be the hypothetically uniform distribution. One procedure for smoothing out such deviations is to "undigitize" the input. If the number of pixels is large enough, the appropriate randomization of the data to approximate a more continuous image distribution can effectively simulate an image transducer of much higher precision. (This idea is similar to the one sometimes used in the generation of halftone pictures upon a binary output device). The randomization vastly diminishes the quantization interval, so M is effectively increased. It should be expected that modes are leveled and gaps are closed up.

Given no other information about the analog-to-digital converter, it is reasonable

to assume that it is of truncating type, as its range is from 0 to M. Thus, the one digital value of x is the representative for all analog values in the interval [x, x+1). Given no other information about the image at a particular pixel (such as statistics in an immediate neighborhood), it can be assumed that each analogue value in the interval is as likely as any other to have been the actual value "seen" by the transducer. Thus, except for those digital values of M indicating possible device over-saturation, the use of the digital value added to a random number uniformly selected from the unit interval is an approximation to the actual real value, with an average error of zero. Assuming channel independence, this new real value of x+random can be created separately for each of three inputs.

This device appears to eliminate digitization-induced irregularities in the distributions of all four transformations (including spurious valleys, which adjusting S cannot do), as seen in their digital histograms under the hypothetically uniform input (figures 6-2 through 6-4). Further, it appears to preserve those properties of the distribution which are in fact due to genuine features of a natural image, as is attested by the comparison of the digital hue histograms of randomized input (figures 6-5 through 6-7) with their nonrandomized counterparts (figures 4-2 through 4-4). Similar results are found with natural scenes when this method is applied to the other (less sensitive) measures of saturation and normalized color (figures 6-8 and 6-9).

Some comments about the use of this procedure are possible. Although it can be used alone, if it is used only after the singularity avoidance cautions have been employed, the output perturbation due to each small randomizing increment will not exceed one digital scale unit. The procedure takes time: in some segmentation techniques, such as region splitting, its cost is incurred only once, though its benefits are repeatedly experienced; for others, it becomes a question of time versus accuracy tradeoff. It seems, however, to be a relatively cheap way to accurately diminish digitization effects. Although trying to compensate for the spurious modes and gaps by working only with the transformed values themselves (directly, or by their distributions) would be cheaper, it would necessarily be inaccurate in some part of the transform range. This is principally because of the nonlinearities of the formulae, and the loss of information due to the large number of pixel types that produce the same transformed value. Examples of these faster but less accurate methods would include randomizing the output value (by how much?), or smoothing the histogram (by what point spread functions?).

The statistical features of the randomization (for example, the variance of the errors in the output distribution, given n pixels) is difficult to analyze. Even merely deriving a closed form expression for the continuous saturation, hue, and normalized color distributions under uniform input is difficult. Calculating the dispersion is an intricate affair, both because of the transformations' nonlinearity, and because some digitization issues still remain. For example, the uniform discrete distribution under the randomization scheme above ensures that the number of pixels within a cube of length n will be exactly in the interval $(n-1)^3$ to n^3. As even this very simple approximation, then, has a complex analysis, about all that can be said is that under the randomizing procedure given above, the calculated distribution will more closely approach the actual image distribution as the number of pixels transformed increases. It is likely that the variance of the errors is inversely proportional to pixel count.

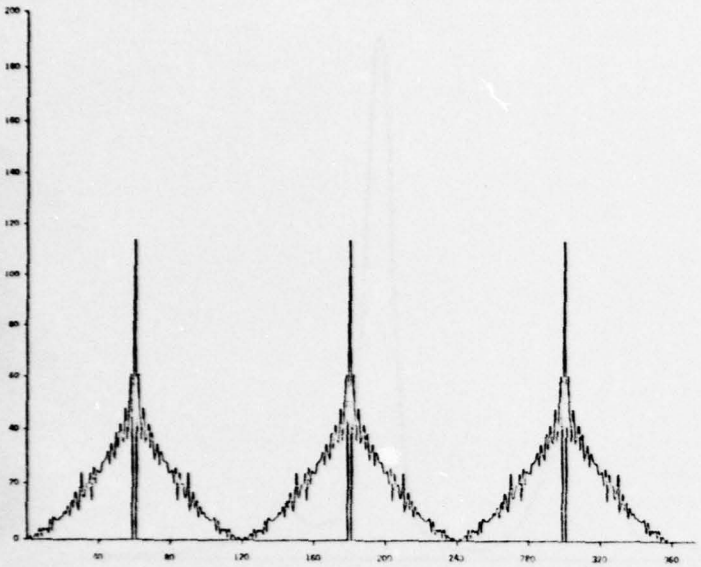


FIGURE 6-1: DIGITAL HUE OF UNIFORM INPUT, WHERE BRIGHTNESS >= M/2 AND SATURATION >= 5, M = 31, S = 360/(2PI)

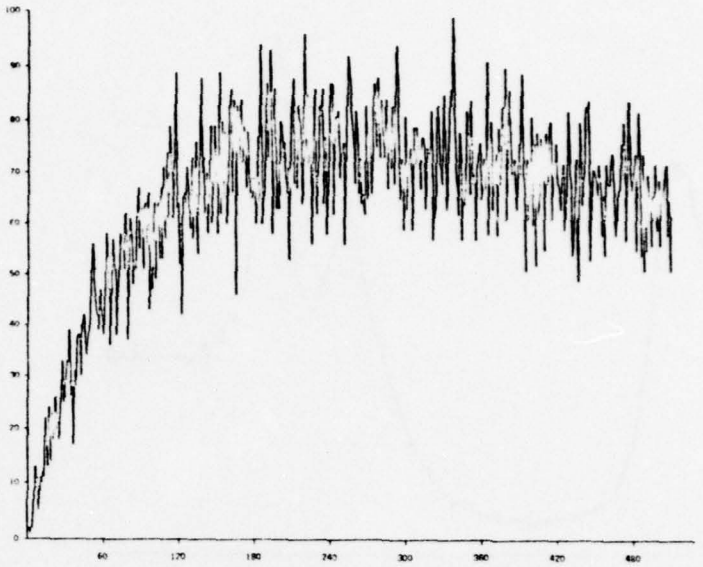


FIGURE 6-2: DIGITAL SATURATION OF UNIFORM INPUT, RANDOMIZED, PARAMETERIZED AS IN FIGURE 3-1.

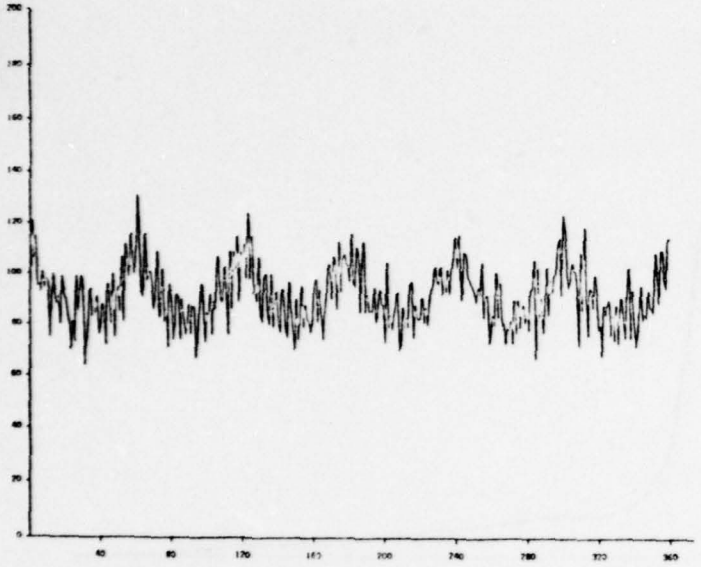


FIGURE 6-3: DIGITAL HUE OF UNIFORM INPUT, RANDOMIZED, PARAMETERIZED AS IN FIGURE 4-1.

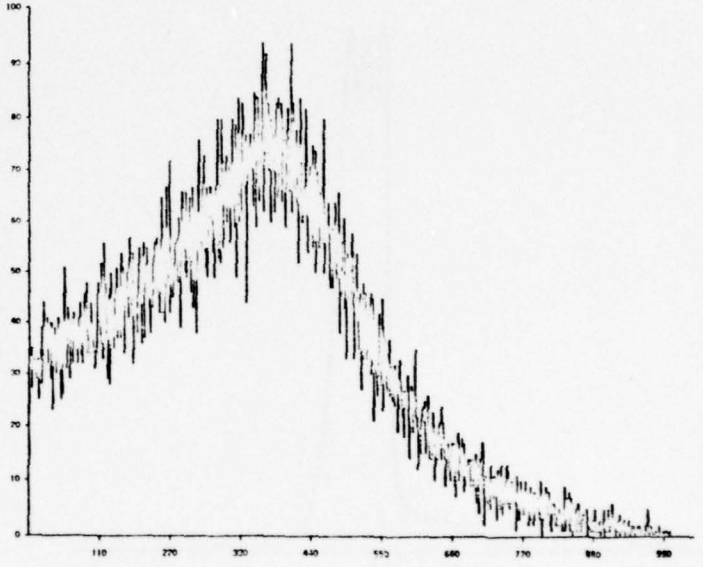


FIGURE 6-4: DIGITAL NORMALIZED RED OF UNIFORM INPUT, RANDOMIZED, PARAMETERIZED AS IN FIGURE 5-1.

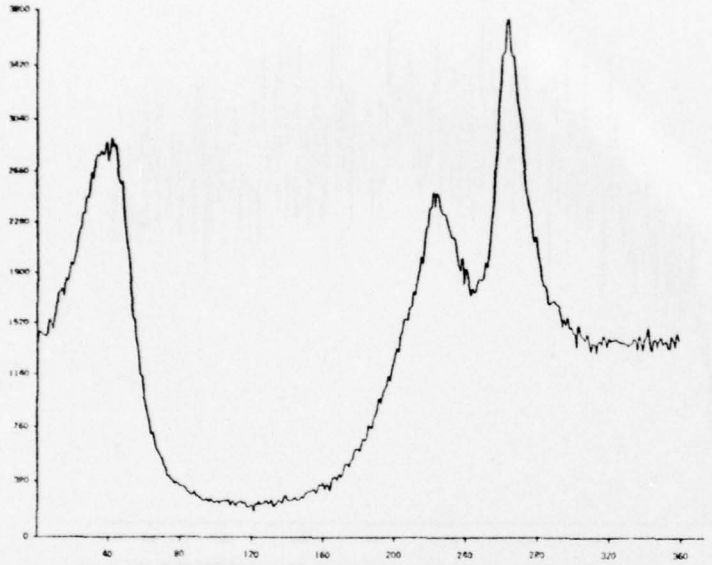


FIGURE 6-5 DIGITAL HUE OF OHLANDER BEAR SCENE, RANDOMIZED, PARAMETERIZED AS IN FIGURE 4-2

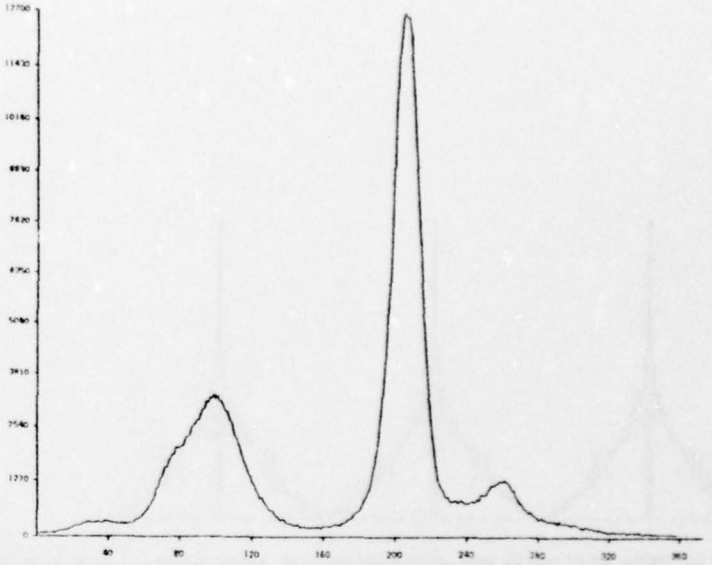


FIGURE 6-6 DIGITAL HUE OF OHLANDER AUTO SCENE, RANDOMIZED, PARAMETERIZED AS IN FIGURE 4-3

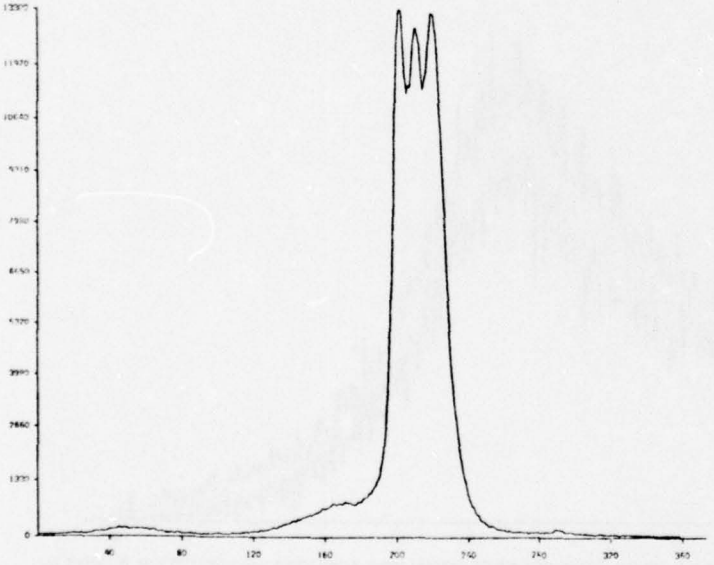


FIGURE 6-7 DIGITAL HUE OF OHLANDER CITY SCENE, RANDOMIZED, PARAMETERIZED AS IN FIGURE 4-4

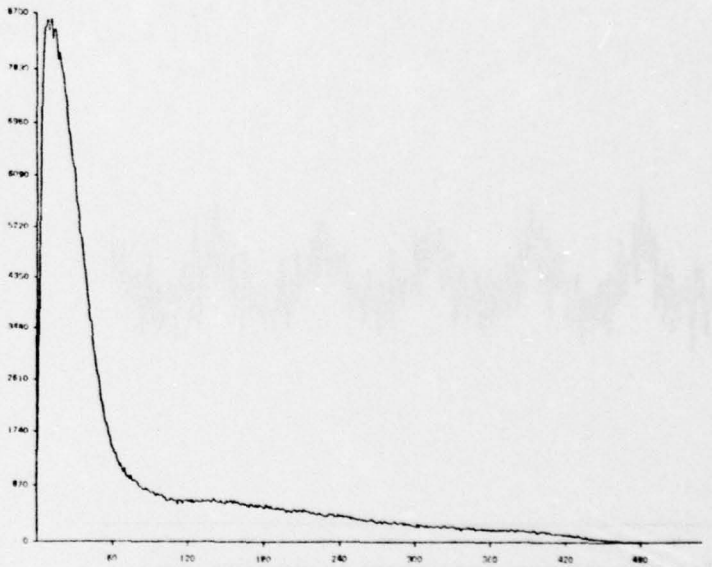


FIGURE 6-8 DIGITAL SATURATION OF OHLANDER BEAR SCENE, RANDOMIZED, PARAMETERIZED AS IN FIGURE 3-2

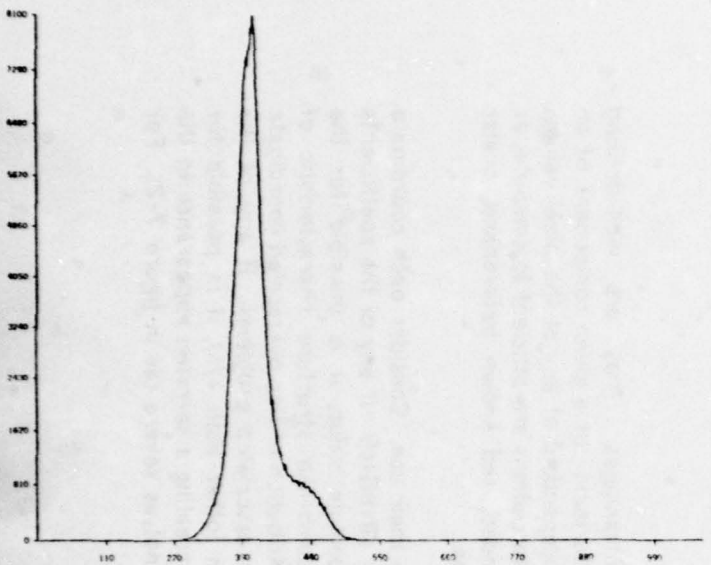


FIGURE 6-9 DIGITAL NORMALIZED RED OF OHLANDER BEAR SCENE, RANDOMIZED, PARAMETERIZED AS IN FIGURE 5-2

7. Other Transformations

Since "color" is basically a psychological phenomenon, many ways have been defined to recombine the tristimulus information into other three-dimensional coordinate systems that attempt to reproduce various aspects of the response of the human visual system. One desired property of these various transformation systems is that of quantifying the concept of color distance: the mathematical formulation of how far apart two colored stimuli are from each other, as perceived by a human observer. These systems, such as Wyszecki 1964 CIE($U^*V^*W^*$)-system; Adams chromatic valence; Glasser cube-root color coordinate system; Hunter L, a_L, b_L system, etc., are, on the average, equivalent formulations that weigh and relate brightness with chromaticity [9]. If it is so desired, they would provide a more accurate (with respect to human visual phenomena) color metric than does the hue, saturation, and brightness system. However, since many of these systems use normalizing divisions in their calculation (often through the requirement that the input be expressed in CIE chromaticity coordinates), analogous digitization effects such as singularities and spurious gaps are certain to appear.

An alternative to the above transformations, which attempt to maximize their similarity to human standards, is the use of linear transformations, which can be used to maximize data separability. The best linear transformation for the tristimulus data would most likely be the principal component (Karhunen-Loeve) one; this, however, is a function of the image and expensive to compute. There is another linear transformation, which is constant over all images, and which has the added advantage of incorporating some properties of subjective perception; it is used as a basis for color television transmission. This YIQ transform is briefly discussed to indicate the freedom of linear transformations from the singularity- and digitization-induced problems, and their utility in manipulating chrominance information.

7.1. Linear Transformations

Linear transformations have several advantages. They are well-defined everywhere, and have no singularities. The perturbation of a given component of an input pixel has an effect on the output that is independent of any of the pixel values, and this effect is readily quantified. Linear transformations are efficient to compute; as the digital domain of the input is small, bounded, and known beforehand, scalar multiplications can be converted into table lookups.

Some cautions are necessary, however, in their use. Consider each coordinate of the transformed vector separately: that is, $T = aR+bG+cB$. If any of the coefficients used to calculate T are greater that 1 in absolute value, it is possible for the distribution of that coordinate to have the comb-like structure characteristic of scaled-up digital data (as in figure 7-1). The contribution of this overscaled coordinate would cause false modes and valleys, and their associated problems. If any of the coefficients are very close to a simple fraction (other than $1/n$), it is possible for "beating" to occur due to digitization error, also creating a serrated appearance in the output distribution, but with modes and valleys not as severe (as in figure 7-2). For

example, the contribution of aR, when a = 2/3 and R is uniform, would have every other output bin twice as full as the intervening ones. That is, the R values 1, 2, 3, 4, 5, 6, . . . would be transformed to the aR values 2/3, 4/3, 6/3, 8/3, 10/3, 12/3, . . . which rounds to 1, 1, 2, 3, 3, 4, ...

Since singularity or ill-condition implies a loss of dimensionality, the transformation matrix should be well-conditioned. Further, the coefficients in each row of the matrix should be scaled so that the largest one of them has an absolute value of 1. This preserves, at least at one of the input coordinates, the maximum amount of dynamic range and granularity in the input, without any arbitrary compression and resulting loss of information and discrimination at the output. This value of 1 also guarantees that a perturbation of ± 1 in any coordinate will create an output perturbation no larger than ± 1; under usual conditions, no false gaps (or valleys) can form.

7.2. The YIQ Transformation

The YIQ linear transformation was devised by the color television industry as a way of minimizing signal bandwidth while retaining subjective color fidelity. It is therefore a type of psychological principal components transform. It is based, in brief, on the phenomena that the eye at very narrow angles of view is achromatic, and that before full color perception is attained at large view angles, it is bichromatic along an orange-cyan axis. The input signal is therefore linearly transformed to provide a brightness signal (Y, after CIE Y), an (approximate) orange filtration (I, for "in phase"), and an (approximate) magenta filtration orthogonal to orange (Q, for "quadrature"), using the matrix below [3]:

$$\begin{matrix} Y \\ I \\ Q \end{matrix} = \begin{matrix} .299 & .587 & .114 \\ .596 & -.274 & -.322 \\ .211 & -.523 & .312 \end{matrix} \begin{matrix} R \\ G \\ B \end{matrix}$$

The above transformation also assumes certain properties of the televising camera and the receiving phosphors, so in any other domain the matrix is strictly only an approximation, albeit a close one, to pure achromatic brightness measure coupled with an (unnormalized) rectangular chromaticity coordinate system. Note also the effect of the bandwidth restraints; for use in segmentation, the matrix needs to be rescaled (particularly the Q component). Following the recommendation above finds:

$$\begin{matrix} Y \\ I \\ Q \end{matrix} = \begin{matrix} .509 & 1.000 & .194 \\ 1.000 & -.460 & -.540 \\ .403 & -1.000 & .597 \end{matrix} \begin{matrix} R \\ G \\ B \end{matrix}$$

The digitization of the real values calculated above is accomplished simply by rounding. Both I and Q can attain negative values; for convenience, their digital values are translated in a positive direction by M, to handle the extreme possible negative output. (This increases the output byte size by 1). Thus, for example:

```
digitalq := round(q)+M
```

The smooth behavior of these three transformations, given the hypothetically uniform input, is evidenced in figures 7-3 through 7-5. The transform applied to a natural scene is shown in figures 7-6 through 7-8.

A recent application of the region-splitting algorithm to several natural scenes has indicated that I and Q have more sharply delimited modes in their distribution (as determined by an automated mode selector) than do hue and saturation, after all four measures have been scaled according to the methods previously discussed [6]. This is probably due to the fact that most natural scenes are of low saturation. This tends to cluster the pixels about the hue singularity, which disperses their transformed values throughout the distribution, decreasing the ability to discriminate them by their modes.

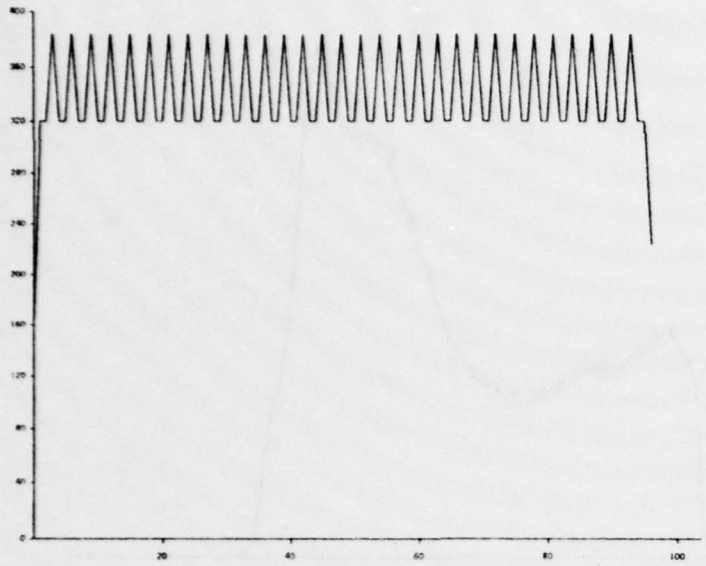


FIGURE 7-1 DIGITAL T = ROUND(3R+.1G+0B) OF UNIFORM INPUT, M = 31.

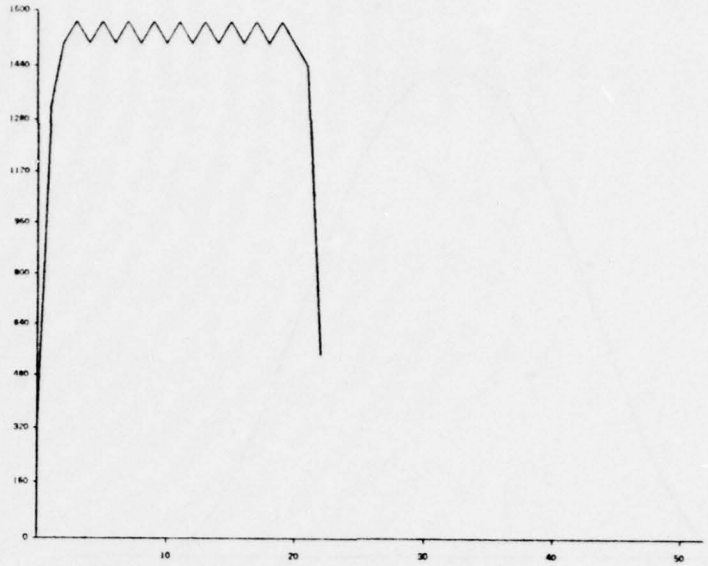


FIGURE 7-2 DIGITAL T = ROUND((2/3)R+.05G+0B) OF UNIFORM INPUT, M = 31.

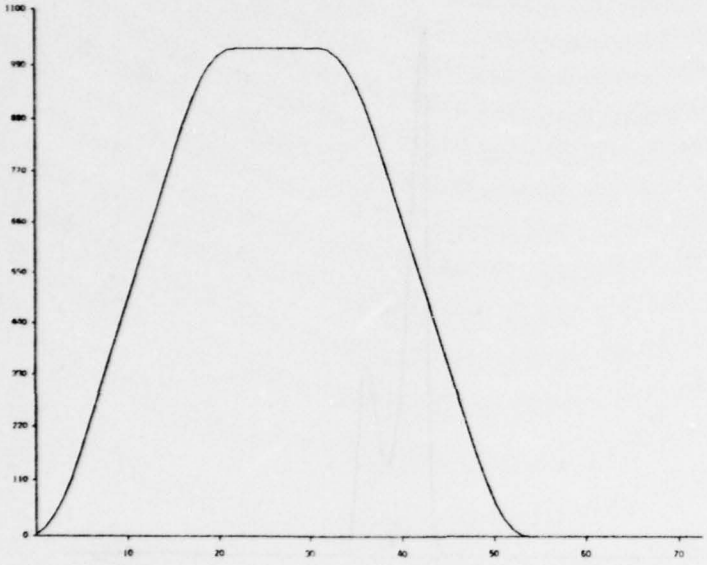


FIGURE 7-3 DIGITAL RESCALED Y OF UNIFORM INPUT, M = 31.

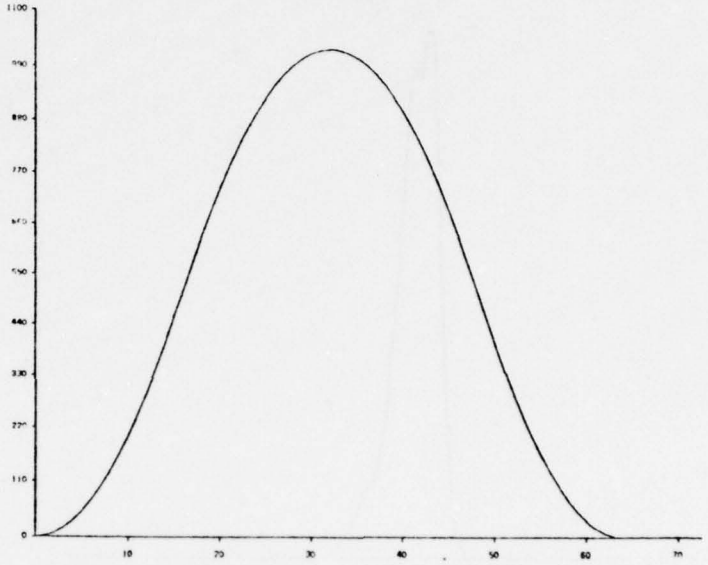


FIGURE 7-4 DIGITAL RESCALED I OF UNIFORM INPUT, M = 31.

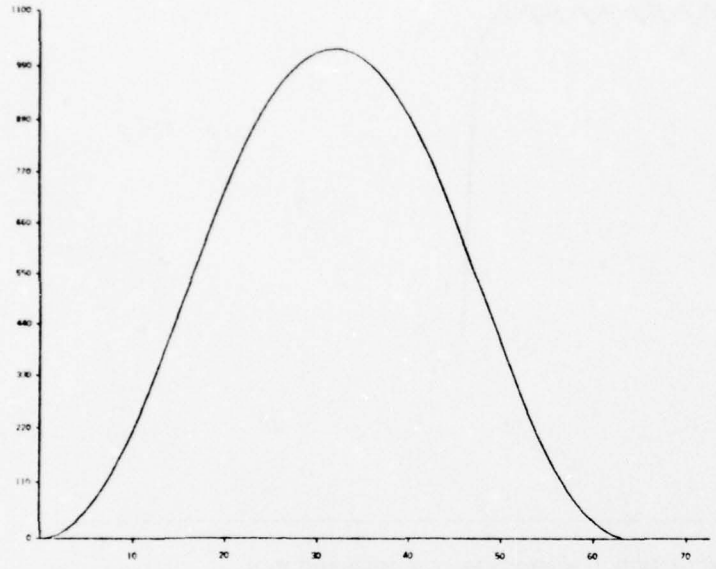


FIGURE 7-5. DIGITAL RESCALED Q OF UNIFORM INPUT, M = 31.

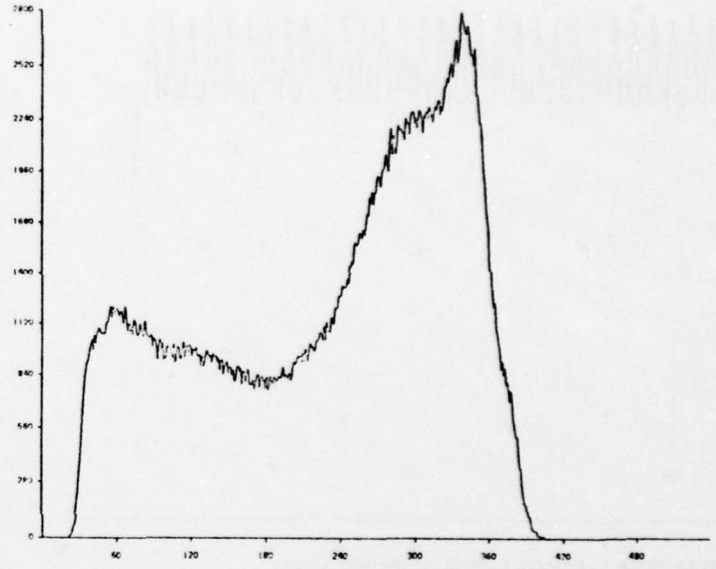


FIGURE 7-6. DIGITAL RESCALED Y OF OHLANDER BEAR SCENE, M = 255.

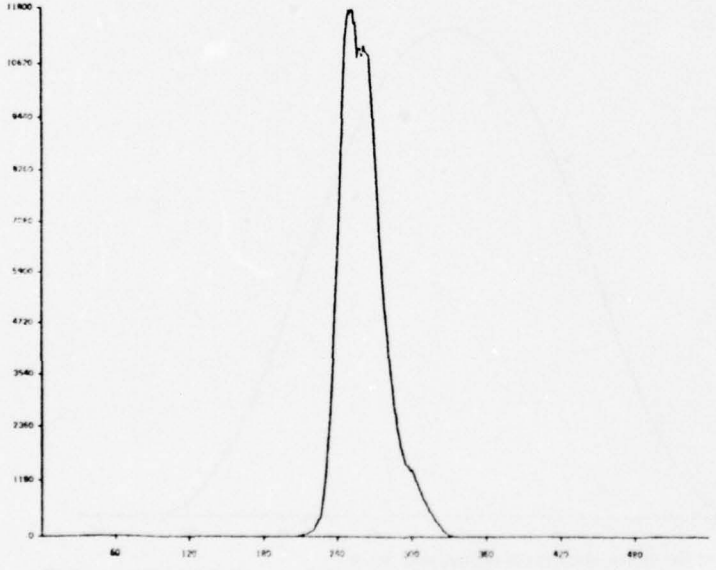


FIGURE 7-7. DIGITAL RESCALED I OF OHLANDER BEAR SCENE, M = 255.

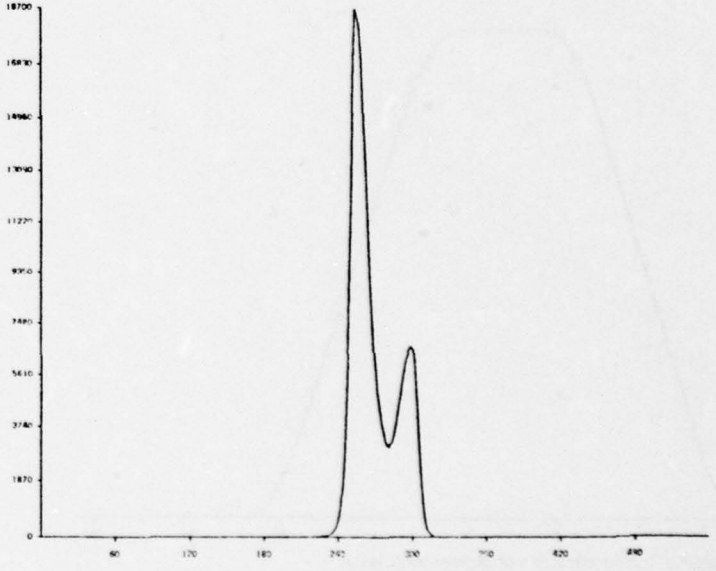


FIGURE 7-8. DIGITAL RESCALED Q OF OHLANDER BEAR SCENE, M = 255.

8. Summary

Three types of transformations used in the analysis of tricolor natural scenes were analyzed: saturation, hue, and normalized color. All three were seen to have nonremovable singularities, near which they are highly unstable. Further, given digital input, the distribution of their transformed values is highly nonuniform, characterized by spurious modes and gaps. These effects were quantified, and illustrated with examples using both theoretical and natural scenes. (Analytic results are summarized in figure 8-1). In addition, the study of hue resulted in a significantly faster algorithm for its calculation.

More empirically, the image segmentation techniques of edge detection, region growing, clustering, and region splitting were seen to be affected arbitrarily badly by such problems. Some stratagems were illustrated that help minimize the bad behavior. Further, linear transformations were presented as a generally favorable alternative to these three nonlinear ones.

Acknowledgements

Keith Price and Steve Rubin read and commented on an earlier draft of this paper. Alan Frisch, Bruce Weide, and David Lamb provided programming, statistical, and documentary assistance, respectively, and helpful discussion. I thank my advisor, Raj Reddy, for his advice and encouragement.

	Saturation	Hue	Normalized Red
Range	[0, 1]	[0, 2π)	[0, 1]
Singularity	R = G = B = 0	R = G = B	R = G = B = 0
Max perturb.	Full range	1/6 range	1/2 range
Modes and gaps, n/d =	min/sum	(G-R)/(G-B+R-B) and analogues	R/sum
Mode height (proportion, uniform input)	(d-3n)/(d-2n)(d-n)	(a/2)/(n+d) [a=2: n,d same parity] [a=1: n,d opposite]	$1/(d-n)$ $\in [0,1/3]$ $(1/(d-n))(1-(3n-d)^2/(2n^2))$ $\in [1/3,1/2]$ $(d-n)/(2n^2)$ $\in [1/2,1]$
Max mode at	1	nπ/3	0
Next three highest at	.25 .4 .5	(2n+1)π/6 (2n+1)π/6±α [α=arctan(sqrt(3))/9]	.5 .333 .25
Gap length (S/M units)	$(3/2)(d-n)/d^2$	$(a/2)(\text{sqrt}(3))(n+d)/(d^2+3n^2)$	$(1/2)(d-n)/d^2$ $\in [0,1/3]$ (n/d^2) $\in [1/3,1]$
Max gap at	1	nπ/3	1
Next three largest at	0 .25 .4	(2n+1)π/6 (2n+1)π/6±α	0 .5 .667
Largest S	(1/2)M	M	M

Figure 8-1: Summary of Analytical Results.

9. References

1. R. O. Duda, and P. E. Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, New York, 1973.

2. A. R. Hanson, and E. M. Riseman, "Preprocessing Cones: A Computational Structure for Scene Analysis," Technical Report 74C-7, Computer and Information Science, University of Massachusetts, Amherst, September, 1974.

3. R. W. G. Hunt, The Reproduction of Colour, John Wiley and Sons, London, 1967.

4. T. Ito, "Color Picture Processing by Computer," Proceedings of 4-IJCAI, pp. 635-642, 1975.

5. R. Ohlander, "Analysis of Natural Scenes," Ph. D. Thesis, Computer Science Department, Carnegie-Mellon University, Pittsburgh, June 1975.

6. K. Price, "Change Detection and Analysis in Multiple Images," Ph. D. Thesis, Computer Science Department, Carnegie-Mellon University, Pittsburgh, forthcoming.

7. J. M. Tenenbaum, T. D. Garvey, S. Weyl, and H. C. Wolf, "An Interactive Facility for Scene Analysis Research," Technical Note 87, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, January, 1974.

8. J. M. Tenenbaum, and S. Weyl, "A Region-Analysis Subsystem for Interactive Scene Analysis," Proceedings of 4-IJCAI, pp. 682-687, 1975.

9. G. Wyszecki, and W. S. Stiles, Color Science, John Wiley and Sons, New York, 1967.

10. M. Yachida, and S. Tsuji, "Application of Color Information to Visual Perception," Pattern Recognition, Vol. 3, pp. 307-323, 1971.

11. Y. Yakimovsky, and J. A. Feldman, "A Semantics-Based Decision Theory Region Analyzer," Proceedings of 3-IJCAI, pp. 580-588, 1973.

Appendix A: Mode and Gap Calculations

A.1. Modes

Spurious modes occur in the distributions at those values which correspond to the creation of a simple fraction by the transformations' basic normalizing division. The height of such a mode depends, in large part, upon how many representatives of the fraction can be formed, given the limited dynamic range of the input; in general, the simpler the fraction, the higher the mode. The actual distribution and prominence of modes is highly dependent on the input distribution and will vary from image to image. However, the assumption of a uniform distribution over the color space will indicate which output values are most likely affected. It will also predict the effect of such a disturbance on the modes nearby one that is known to be spuriously present, at least in terms of expected relative magnitudes.

Advantage is taken of the uniformity of the assumed input distribution, which permits the use of counting arguments in place of probabilistic ones. (Any pixel triple is as likely to occur as any other). The analysis of the modes in all three of the following distributions depends on the counting of the number of representatives of the mode fraction n/d that can be formed, subject to the constraints imposed by each particular distribution on the formation of such fractions.

A.1.1 Saturation

$$\text{Prob}\{\min(R, G, B)/\text{sum}(R, G, B) = n/d \mid n/d \text{ reduced}\} =$$

$$\sum_{i \in I} (\text{ways: } \min = i*n \wedge \text{sum} = i*d \wedge 0 \le R, G, B \le M) \; / \; (M+1)^3$$

The set I is defined implicitly by the above three constraints. Explicitly, the lower bound for i is determined from $0 \le n$, $0 < d$ and $0 \le R, G, B$; thus, $0 \le i$. Further, $1 \le i$ to avoid the singularity. (The singularity itself is a special case. There, min/sum = 0/0 and can be attained in exactly 1 way, giving a probability of about $1/M^3$). The upper bound is seen as follows. The first constraint implies $in \le M$, thus $i \le M/n$. The second constraint implies $id \le 3M$, thus $i \le 3M/d$. The two together imply, as min = in, sum = in+a+b = id for $\min \le a, b \le M$; thus $i = (a+b)/(d-n)$ or $i \le 2M/(d-n)$. The two together also imply $3n \le d$. Using this last relation to order the three constraints on i, it is seen that the most stringent is $i \le 2M/(d-n)$. Since i is an integer, the upper bound is strictly $u = \text{floor}(2M/(d-n))$. The set I is then $1 \le i \le u$.

Counting the ways of each value of i is a bit complicated. Note that both min and sum are invariant with respect to permutation of their variables. By the standard relation for counting the members in the union of three overlapping sets, it is seen that:

ways: min(R, G, B) = in ≡

3(ways: R = in ∧ in ≤ G, B) - 3(ways: R = G = in ∧ in ≤ B)

+ (ways: R = G = B = in)

The summand can be now broken into three parts, as determined by the above. Combining constraints, the first part reads (except for the factor 3):

ways: in+G+B = id ∧ in ≤ G, B ≤ M ≡

ways: in+G'+in+B'+in = id ∧ in ≤ G'+in, B'+in ≤ M ≡

ways: G'+B' = id-3in ∧ 0 ≤ G', B' ≤ M-in ≡

if id-3in ≤ M-in then id-3in+1 else 2(M-in)-(id-3in)+1 ≡

if i ≤ M/(d-2n) then id-3in+1 else 2(M-in)-(id-3in)+1

The second part is developed similarly (except for the factor -3):

ways: in+in+B = id ∧ in ≤ B ≤ M ≡

ways: B' = id-3in ∧ 0 ≤ B' ≤ M-in ≡

if i ≤ M/(d-2n) then 1 else 0

The third part becomes:

ways: in+in+in = id ≡

if n/d = 1/3 then 1 else 0

Thus, combining terms and their tests, the total number of ways is:

$$\sum_{i=1}^{u} [(\text{if } i \le M/(d-2n) \text{ then } 3((id-3in+1)-1) \text{ else } 3(2(M-in)-(id-3in)+1) + (\text{if } n/d = 1/3 \text{ then } 1 \text{ else } 0)]$$

First consider the case n/d = 1/3 (i.e. n = 1 and d = 3). Direct substitution finds u = M, which is also the value of the summation. This result can be seen directly from the fact that n/d = 1/3 implies zero saturation, which occurs the M times all three coordinates are equal and non-zero.

Consider now the cases where n/d ≠ 1/3. Let c (for "crossover") = floor(M/(d-2n)). It is seen that 0 ≤ c; further, 0 ≤ c ≤ u as a consequence of n/d < 1/3.

Thus, using the value of c to split the sum into two parts, eliminating the if-then-else, and rearranging terms, the number of ways becomes:

$$3[\sum_{i=1}^{c}((d-3n)i) + \sum_{i=c+1}^{u}((d-n)i) + \sum_{i=c+1}^{u}(2M+1)] =$$

$$3[(d-3n)c(c+1)/2 - (d-n)(u-c)(u+c+1)/2 + (u-c)(2M+1)]$$

Given that $M = 2^m-1$ is usually large (≥ 63, say), then $M^2 >> M$ and the above expression is dominated by the M^2 terms. Since u and c are in terms of M, this includes terms of the form c^2, u^2, uM, and cM.

Thus, the dominant term is:

$$3[(d-3n)c^2/2 - (d-n)(u^2-c^2)/2 + 2M(u-c)] =$$

$$3[(d-2n)c^2 - (d-n)u^2/2 + 2Mu - 2Mc]$$

Approximating the integer values of c and u by the real values they would have prior to truncation by the floor function, this simplifies to:

$$3[Mc - 2Mu/2 + 2Mu - 2Mc] =$$

$$3M[u - c] =$$

$$3M^2(d-3n)/((d-n)(d-2n)$$

Therefore, exclusive of the singularity, the probability that min/sum = n/d is about $(3/M)(d-3n)/((d-n)(d-2n))$. (If n/d = 1/3, it is about $(1/M^2)$). Considered as a continuous function of two variables, its behavior can be studied using derivatives. For constant d, it is decreasing for all increasing n. For constant n, it is decreasing for increasing $d \geq 5n$. (Note, however, $d \geq 3n$ by definition, and the use of derivatives is an overly strict test: n/d must remain in lowest terms).

In any case, the function is maximized for small values of n and d, with a global maximum at n/d = 0/1 (saturation = 1), where the probability is 3/M. The next three highest local maxima are at n/d = 1/4, 1/5, 1/6 (saturation = .25, .4, .5), where the probabilities are .5/M, .5/M, .45/M, respectively.

A.1.2 Hue

Counting arguments will again be used for hues inside the (somewhat greater than) one-sixth portion of the color triangle, $R \geq G \geq B$. Thus:

$$\text{Prob}\{(G-R)/((G-B)+(R-B)) = -n/d \mid n/d \text{ reduced}\} =$$

$$\sum_{B=0}^{M} \sum_{i \in I} (\text{ways: } R-G = i*n \wedge (G-B)+(R-B) = i*d \wedge 0 \leq B \leq G \leq R \leq M) \; / \; (M+1)^3$$

The set I is explicitly determined as follows. As before, $1 \leq i$ to avoid the singularity. (The singularity itself is a special case. There, (G-R)/(G-B+R-B) = 0/0, attainable the M+1 times that R = G = B, giving a probability of about $1/M^2$). The upper bound is seen as follows. The first constraint implies $in \leq M$, thus $i < M/n$. The

second constraint implies $id \le 2M$, thus $i \le 2M/d$. The two together imply $i(n+d) = 2R-2B$, thus $i \le 2M/(n+d)$. The two together also imply $n \le d$. Using the last relation to order the constraints on i, it is seen that the most stringent is $i \le 2M/(n+d)$. Since i is an integer, the upper bound is strictly $u = \text{floor}(2M/(n+d))$, and the set I is then $1 \le i \le u$.

Counting the number of ways finds:

ways: $R-G = in \wedge (G-B)+(R-B) = id \wedge 0 \le B \le G \le R \le M =$

ways: $R'-G' = in \wedge G'+R' = id \wedge 0 \le G' \le R' \le M-B$

Solving the above two equations, $R' = i(n+d)/2$, and $G' = i(d-n)/2$. The number of ways is simply 1 or 0, depending on the parity of the right hand sides, since both R' and G' are constrained to be integers. Thus, two cases emerge.

Case 1: n, d same parity (and therefore, both must be odd).

Here, R' and G' are always integers, independent of the parity of i. Thus, the number of ways is:

$$\sum_{B=0}^{M} \sum_{i=1}^{u} 1 = \sum_{B=0}^{M} \text{floor}(2(M-B)/(n+d))$$

Approximating the summand by deleting the floor function gives:

$$(2/(n+d))[M(M-1)- \sum_{B=0}^{M} B] =$$

$$(1/(n+d))(M^2+M-2)$$

As the M^2 term dominates, this is approximately:

$$M^2/(n+d)$$

Case 2: n, d opposite parity.

Here, R' and G' are attainable only when i is even. This happens for about half the values of i. Thus, the number of ways is approximately half the above, or, $(1/2)M^2/(n+d)$.

Therefore, exclusive of the singularity, the probability that $(G-R)/(G-B+R-B) = -n/d$ is about $(a/2)(1/M)(1/(n+d))$, where $a = 2$ if n and d are the same parity, and $a = 1$ if they are not. This function is decreasing for any increase in n or d, and maximized for small values of n and d (with n/d in lowest terms).

Two equal global maxima occur at $n/d = 0/1$ (hue $= \pi/3$) and $n/d = 1/1$ (hue $= 0$), where the probability is .5/M. Thus, over the entire color triangle, there are equal global maxima at hue $= n\pi/3$, for $n = 0, 1, \ldots, 5$, also with probability .5/M. The next

three highest local maximum classes occur at n/d = 1/3 [hue = (2n+1)π/6, for n = 0, 1, . . . , 5] with probability .25/M, and at n/d = 1/2 or 1/5 [hue = (2n+1)π/6±∝, for n = 0, 1, . . . , 5, where ∝ = arctan(sqrt(3)/9) (about 11 degrees)], with probability .167/M.

A.1.3 Normalized Color

As all normalized primaries are calculated in the same way, only normalized red is analyzed. Counting arguments are used again. Thus:

$$\text{Prob}\{R/(R+G+B) = n/d \mid n/d \text{ reduced}\} =$$

$$\sum_{i \in I} (\text{ways: } R = i*n \wedge R+G+B = i*d \wedge 0 \le R, G, B \le M) / (M+1)^3$$

The set I is explicitly determined as follows. As before, $1 \le i$ to avoid the singularity. (The singularity is a special case. There, R/(R+G+B) = 0/0, which is attained exactly once, giving a probability of about $1/M^3$). The upper bound is seen as follows. The first constraint implies $in \le M$, thus $i \le M/n$. The second constraint implies $id \le 3M$, thus $i \le 3M/d$. The two together imply $i(d-n) = B+G$, thus $i \le 2M/(d-n)$. Ordering the constraints reveals that if $n/d \le 1/3$, the most stringent is $i \le 2M/(d-n)$; otherwise, the most stringent is $i \le M/n$. The upper bound is then u = floor(2M/(d-n)) or u = floor(M/n), depending on the value of n/d, which splits the analysis into two cases. In either, $1 \le i \le u$.

Counting the number of ways finds:

ways: R = in ∧ R+G+B = id ∧ 0 ≤ R, G, B ≤ M =

ways: G+B = i(d-n) ∧ 0 ≤ G, B ≤ M =

if i ≤ M/(d-n) then i(d-n)+1 else 2M-i(d-n)+1

Case 1: $n/d \le 1/3$

Let c = floor(M/(d-n)) be the crossover point in the if-then-else. It is seen that $0 \le c \le u$ = floor(2M/(d-n)); in fact, c = u/2, approximately. Then the number of ways becomes:

$$\sum_{i=1}^{c} (i(d-n)+1) + \sum_{i=c+1}^{u} (2M-i(d-n)+1) =$$

$$u+2M(u-c)+(d-n)[\sum_{i=1}^{c} i - \sum_{i=c+1}^{u} i]$$

But as, approximately, 2c = u, this is approximately:

$$u+2Mc+(d-n)(-\sum_{i=1}^{c} c) =$$

$$u+2Mc-(d-n)c^2$$

The dominant terms are the second and third. Approximating c and u by real values finds:

$M^2/(d-n)$

Case 2: $1/3 \le n/d$

Here there is a complication; now $u = \text{floor}(M/n)$. For $1/3 \le n/d \le 1/2$, $c \le u$, and the sum splits. However, if $1/2 \le n/d$, $c \ge u$ and the sum consists only of the "then" part. The number of ways then becomes:

Case 2a: $1/3 \le n/d \le 1/2$

$$\sum_{i=1}^{c} (i(d-n)+1) + \sum_{i=c+1}^{u} (2M-i(d-n)+1) =$$

$$u+2M(u-c)+(d-n)(c^2+c)-(d-n)(1/2)(u^2+u)$$

The dominant terms are those in M^2. Selecting these, and approximating u and c by reals, finds:

$$2M^2(1/n-1/(d-n))+M(M/(d-n))-(d-n)(1/2)(M/n)(M/n) =$$

$$M^2(-7n^2+6nd-d^2)/(2n^2(d-n)) =$$

$$(M^2/(d-n))\,(1-(3n-d)^2/(2n^2)) \text{ or } (M^2(d-n)/(2n^2))\,(2n(2d-3n)/(d-n)^2-1)$$

Case 2b: $1/2 \le n/d$

$$\sum_{i=1}^{u} (i(d-n)+1) =$$

$$u+(d-n)(1/2)(u^2+u)$$

Except for the case $n/d = 1/1$ (where the sum $= u = M/n = M$), the dominant term is the one in u^2. Approximating u finds:

$$M^2(d-n)/(2n^2)$$

Therefore, exclusive of the singularity, the probability that R/sum $= n/d$ generally breaks into three cases. (If $n/d = 1/1$, it is about $(1/M^2)$). If $n/d \le 1/3$, it is about $(1/M)/(d-n)$. If $1/3 \le n/d \le 1/2$, it is about $(1/M)(1/(d-n))(1-(3n-d)^2/(2n^2))$. If $1/2 \le n/d$, it is about $(1/M)(d-n)/(2n^2)$. Consider each case separately.

The first case considered as a continuous function of two variables can be studied by derivatives. It is increasing with increasing n, for constant d. However, by the constraint, $n \le d/3$ and the function is maximized for constant d at $n = d/3$. Here, its value is $(1/M)(3/(2d))$, which is decreasing for increasing d (as is the original expression). Thus, the function is maximized for small values of d, which implies that n will also be small, with n and d in lowest terms.

The third case, again by derivatives, shows the function increasing for increasing d, for constant n. However, by the constraint, $d \le 2n$ and the function is maximized for constant n at $d = 2n$. Here, its value is $(1/M)(1/(2n))$, which is decreasing for

increasing n (as is the original expression). Thus, the function is maximized for small values of n, which implies that d will also be small, with n and d in lowest terms).

The second case, a type of transition between the first and the third, is difficult to analyze direcly. However, as shown above, the number of ways can be expressed in one of two ways. Firstly, it is equal to the product of the number of ways in the first case, and a function that smoothly decreases from a value of 1 at n/d = 1/3 to a value of 1/2 at n/d = 1/2. Secondly, it is equal to the product of the number of ways in the second case, and a function that smoothly increases from a value of 1/2 at n/d = 1/3 to a value of 1 at n/d = 1/2. As both the first and second cases are maximized at small fractions, it is intuitively clear that this case is also maximized there (a conclusion further buttressed by an examination of the digital histogram).

The global maximum occurs at n/d = 0/1 (normalized color = 0), where the probability is 1/M. The next three highest local maxima occur at n/d = 1/2, 1/3, 1/4 (normalized color = .5, .333, .25), where the probabilities are .5/M, .5/M, .333/M, respectively.

A.2. Gaps

Spurious gaps on either side of a mode corresponding to a simple fraction n/d are caused by the digitization of the input and its limited dynamic range: only certain fractions are possible. The extent of the gap depends upon how near the mode it is possible to form a differing fraction; in general, the simpler the fraction, the wider the surrounding gaps. (Thus, in general, the larger gaps surround the larger modes, and the reverse). Note that this effect is independent of the input distribution. Although certain distributions will induce wider gaps, a lower limit of gap size determined solely by input dynamic range always exists, as shown below.

The analysis of the gaps in all three of the following distributions depends on finding that faction x/y, in lowest terms and different from n/d, that minimizes abs(x/y - n/d), subject to the constraints imposed by each particular distribution on the formation of such fractions.

A.2.4 Saturation

The fraction x/y that minimizes abs(x/y - n/d) = abs((dx-ny)/(dy)) will be a nonsimple fraction near the largest permitted representative of n/d (thus y will also be large) that, if possible, makes dx-ny = 1. The largest representative of n/d is (un)/(ud), where u = floor(2M/(d-n)), as seen in the section on modes. By the constraints for saturation, it is possible to perturb fractions of the form min/sum so that x/y can be any of the fractions formed from $(u-1)n \le x < un$, and $(u-1)d \le y < ud$. By number theory, as n and d are relatively prime, there is at least one such x and y pair in the specified range that causes dx-ny = 1. (If n = 0, choose x = 1). Thus, abs((dx-ny)/(dy)) = (1/d)(1/y) = approximately (1/d)(1/ud), by the upper limit on y. By approximating u, this is about $(d-n)/(2Md^2)$. Thus, the digital saturation gap is:

round(S(1-3(n/d))) - round(S(1-3(x/y))) = approximately

$3S(d-n)/(2Md^2)$ =

$(3/2)(S/M)(d-n)/d^2$

The global maximum occurs at n/d = 0/1 (saturation = 1), where the gap is of length 1.5(S/M). The next three highest local maxima occur at n/d = 1/3, 1/4, 1/5 (saturation = 0, .25, .4), with gap lengths of .333(S/M), .281(S/M), .24(S/M), respectively.

A.2.5 Hue

The fraction x/y that minimizes abs(x/y - n/d) = abs((dx-ny)/(dy)) will be a nonsimple fraction near the largest permitted representative of n/d (thus y will also be large) that, if possible, makes dx-ny = 1. The largest representative of n/d is (un)/(ud), where u = floor(2M/(n+d)), as seen in the section on modes. By the constraints for hue, however, fractions of the form (G-R)/(G-B+R-B) are limited in their perturbations; any change in any coordinate yields a new fraction whose numerator versus. denominator parity relation is always conserved. That is, if the original numerator and denominator had the same parity, any perturbation also will; similarly for the case of differing parity. Thus, only about half the fractions x/y which can be formed from $(u-1)n \le x < un$, and $(u-1)d \le y < ud$ are valid perturbations of n/d. Still, if n and d have differing parity, it is possible to find such a valid x and y in the specified range (x and y of differing parity, also) such that dx-ny = 1. (If n = 0, choose x = 1). However, no such x and y pair exists if n and d are of the same parity, as dx-ny is then always even (since x and y are also constrained to be of the same parity). But under these latter conditions, it is possible to find a valid x and y in the specified range such that dx-ny = 2. Therefore, abs((dx-ny)/(dy)) = (a/d)(1/y), where a = 2 if n and d are of the same parity, and a = 1 otherwise. Then (a/d)(1/y) = approximately (a/d)(1/ud), by the upper limit on y. By approximating u, this is about $a(n+d)/(2Md^2)$. Thus, the digital hue gap is:

round(S(arctan(sqrt(3)(n/d)))) - round(S(arctan(sqrt(3)(x/y)))) = approximately

S(arctan(sqrt(3)(n/d))) - S(arctan(sqrt(3)(x/y))) =

S(arctan((sqrt(3)(n/d - x/y)) / (1 + (n/d)(x/y))))

As x/y is very close to n/d, this is approximately:

$S(\arctan(\sqrt{3}a(n+d)/(2Md^2) / (1+(n/d)^2)))$

As this is a small quantity, the arctan can be dropped, yielding:

$(a/2)(\sqrt{3})(S/M)(n+d)/(d^2+3n^2)$

There are two equal global maxima, at n/d = 0/1 (hue = $\pi/3$), and at n/d = 1/1

(hue = 0), where the gaps are of length .5sqrt(3)(S/M). Thus, over the entire color triangle, there are equal global maxima at hue = nπ/3, for n = 0, 1, . . . , 5, also with gap length of .5sqrt(3)(S/M). The next three highest local maxima classes occur at n/d = 1/3 [hue = (2n+1)π/6, for n = 0, 1, . . . , 5], with gap length of .333sqrt(3)(S/M), and at n/d = 1/2 or 1/5 [hue = (2n+1)π/6±∝, for n = 0, 1, . . . , 5, where ∝ = arctan(sqrt(3)/9) (about 11 degrees)], with gap length of .214sqrt(3)(S/M).

A.2.6 Normalized Color

The fraction x/y that minimizes abs(x/y - n/d) = abs((dx-ny)/(dy)) will be a nonsimple fraction near the largest permitted representative of n/d (thus y will also be large) that, if possible, makes dx-ny = 1. The largest representative of n/d is (un)/(ud), where u = floor(2M/(d-n)) for n/d ≤ 1/3, and u = floor(M/n) otherwise, as seen in the section on modes. By the constraints for normalized color, it is possible to perturb fractions of the form R/sum so that x/y can be any of the fractions formed from (u-1)n ≤ x < un, and (u-1)d ≤ y < ud. By number theory, as n and d are relatively prime, there is at least one such x and y pair in the specified range that causes dx-ny = 1. (If n = 0, choose x = 1). Thus, abs((dx-ny)/(dy)) = (1/d)(1/y) = approximately (1/d)(1/ud), by the upper limit on y. By approximating u, this is about (d-n)/(2Md^2) or n/(Md^2), depending on n/d. Thus, the digital normalized color gap is:

round(S(n/d)) - round(S(x/y)) = approximately

(1/2)(S/M)(d-n)/d^2 if n/d ≤ 1/3, and (S/M)(n/d^2) otherwise

The global maximum occurs at n/d = 1/1 (normalized color = 1), where the gap is of length 1(S/M). The next three highest local maxima occur at n/d = 0/1, 1/2, 1/3 (normalized color = 0, .5, .667), with gap lengths of .5(S/M), .25(S/M), .222(S/M), respectively.

Appendix B: Hue Algorithm

B.1. Derivation

The following derivation of a faster algorithm for the computation of hue is based on the equivalence relation proved in the main text. That is, hue(R, G, B) = hue(R-K, G-K, B-K), ∀K. The simplification of the algorithm also depends on the geometry of the color triangle, and the fact that the hue algorithm, as originally stated, accurately reflects the psychological phenomenon of the saturation independence of perceived hue.

First, though, for purposes of comparison it is necessary to rewrite the original algorithm so that it properly handles the singularity. Thus:

```
hue :=  if B > G then
            arccos((2r-g-b) /
                (sqrt(6) sqrt((r-1/3)^2+(g-1/3)^2+(b-1/3)^2)))
        else
            if G > B then
                2pi-arccos((2r-g-b) /
                    (sqrt(6) sqrt((r-1/3)^2+(g-1/3)^2+(b-1/3)^2)))
            else comment: G = B;
                if R > B then
                    0
                else
                    if R < B then
                        pi
                    else comment: R = G = B;
                        achromatic;
```

Here, "achromatic" is any value unattainable by the rest of the computation. The above algorithm avoids the singularity by avoiding the line through pure red and the white point.

A new algorithm can be derived from the corollary to the equivalence relation: let K be the minimum pixel coordinate. Now a corresponding totally saturated, but hue-equivalent, pixel can be formed by subtracting out the white component. Let this new pixel be designated as (R', G', B'). At least one of its coordinates is guaranteed to be zero. Thus, the calculation of hue, which was originally a calculation in three variables, naturally reduces to the selection and execution of one of three possible functions of two variables, where each of the three is simply the general algorithm simplified under the assumptions that the given coordinate is zero. For example, hue(R, G, B), when B is minimum, is calculated as hue(R', G', B') = hue(R-B, G-B, 0).

Further simplification is based on the following observation. Fully saturated

colors must lie on the perimeter of the normalized color triangle (as in figure B-1). Which of the three lines of the perimeter this pixel lies upon is determined by the coordinate whose value is zero. For example, if B (or B') = 0, then the pixel lies on the red-green line, and its hue value must be between 0 and $2\pi/3$. The calculation of hue angle now becomes a function of two variables, the values of which determine exactly where along the given side of the triangle the hue-equivalent pixel lies. Continuing the example, the angle formed in the red-green third of the color triangle (with respect to the baseline axis joining the white point and pure yellow) is arctan(distance of pixel from yellow / length of white-to-yellow line) (as in figure B-1). In this third, a pixel's normalized green value measures its distance from the red point; this value is G'/(R'+G'+B') = G'/(R'+G'), as B' = 0. The angle then is arctan((G'/(R'+G')-1/2) / (sqrt(3)/6)) = arctan(sqrt(3)(G-R)/(G-B+R-B)). As the color triangle exhibits a sixfold symmetry, similar formulae hold for the green-blue and blue-red thirds also.

The angles derived must be adjusted by constant factors, however, to reflect the angular offsets from red of the three baselines from which they were measured. Hence, the entire algorithm can be rewritten, again taking care to avoid the singularity, as follows:

```
hue :=  if R > B and G > B then
            pi/3+arctan(sqrt(3)(G-R)/(G-B+R-B))
        else
            if G > R then
                pi+arctan(sqrt(3)(B-G)/(B-R+G-R))
            else
                if B > G then
                    5pi/3+arctan(sqrt(3)(R-B)/(R-G+B-G))
                else
                    if R > B then
                        0
                    else
                        achromatic;
```

(The above discussion is also valid if all tristimulus values are replaced by chromaticity coordinates, as the normalizing division by R+G+B always cancels out (i.e., hue is scale invariant). Thus the central arctangent formula can read, in the first third of the color triangle, arctan(sqrt(3)(g-r)/(g-b+r-b)), with analogous expressions for the other two thirds. However, since r+b+g = 1, and, in the first third, saturation = 1-3b, this simplifies the expressions to arctan(sqrt(3)(g-r)/saturation) and its analogues: a formula not totally devoid of elegance.)

The time savings are apparent; three multiplications (powers) and a square root are exchanged for the use of three comparisons, and some additions and subtractions are dispensed with altogether.

B.2. Modifications

The calculation of hue is expensive. The following are two suggestions for reducing computation time.

The first is that, as the hue transformation is itself only an approximation to psychological phenomena, perhaps an approximate transformation can be devised that is cheaper to compute. One such algorithm would be accomplished by dropping the arctangent in the calculation. As the arctangent is monotonic, the ratio (G-R)/(G-B+R-B) itself (for hues in the first third), multiplied by the appropriate scale factor $\pi/3$ should be sufficient to discriminate various hues. The same decisions are used as before to determine the proper one of the three formulae. The new algorithm has the effect that now all pixels are projected, instead, onto the periphery of a color wheel circumscribed about the color triangle (as in figure B-2). This new algorithm is still intensity and saturation invariant; in terms of chromaticity coordinates it appears simply as $(\pi/3)(g-r)/\text{saturation}$ and its analogues. The spurious modes and gaps of this transformation are analyzed in almost the same way as hue was in Appendix A. Mode heights are equivalent, though they appear in different locations. However, spurious gaps are exaggerated about some modes, especially at hue = $2n\pi/3$, n = 0, 1, 2; here gap length becomes $(2\pi/3)(S/M)$, or about 2.3 times larger (as shown in figure B-3). Thus, if this algorithm is used, a smaller digitizing scale factor is required, and resolution is cut about in half.

A better and only slightly more expensive modification is possible. Since digitized input has a well defined range and a small number of possible values, table lookup provides an efficient way to exchange a small amount of space for a large savings in time. This well known device allows the calculation of a function of one variable to be done beforehand, and replaces its run-time execution with an indexing operation into a one-dimensional array; this is usually a much cheaper sequence of steps. This idea can be extended to functions of more than one variable, although unless the range of each variable is highly restricted and the function highly complex, the overhead in both filling an n-dimensional array plus the accessing overhead for n-dimensional variables is usually excessive.

Hue is a function of three variables. One of the functions that compose its calculation employs the arctangent, an expensive computation. However, a three dimensional array would be excessively large. Further, there is no apparent easy decomposition of the formula (as opposed to that of matrix multiplication used in linear transformations: if T = aR+bG+cB, then it is more efficient if T = a[R]+b[G]+c[B]). Some savings can occur by collapsing the constant and division into one multiplication, which is then made into a table look-up: arctan((G-R) * sqrt3over[G-B+R-B]). (This device can also be used in the calculation of brightness and saturation). But the problem remains that the argument to the arctangent is not an integer.

Note here, though, that the relatively expensive arctangent calculates too much; most of what it computes is thrown away. If the input picture is to be transformed

and also digitized, the value of the arctangent (plus the baseline constant) is scaled up and rounded. Much of the precision is needed only to select which of two consecutive integers has the least error as an approximation to the scaled up real. What is needed in the computation, then, is a type of reverse table lookup. In this method, a very quick computation predicts, somewhat inaccurately, what the scaled output integer should be. The accuracy of this fast guess is checked by using the integer, through table lookup, to determine what range of arguments to the arctangent would have indeed produced such an integer output. If the original argument lies within the bounds, the guess is correct and accepted as the proper value. If not, nearby integers are attempted. If the initial prediction is a good (and cheap) one, and if one can guarantee what type of error it makes, thereby limiting the resulting search, the computation time will be reduced.

In practice, since the arctangent is monotonically increasing and has odd symmetry, it is possible to arrange matters so that, using only the positive half of the function, the prediction is either exact or always undershoots. The search can then always preceed in one direction (up). Secondly, at the expense of space (in the form of another look-up table), the predictions can be made arbitrarily accurate, so that the necessity for searching is greatly reduced, and, in fact, can become simply a check of the single next highest value.

The hue algorithm presented in this appendix employs the arctangent in three cases, each with its own baseline constant. Rather than using an algorithm which uses the above device to check the estimated integer for acceptability against the proper one of three tables, some savings can be effected by the following. It usually occurs that S is of the form $N/(2\pi)$. If $N = 6K$, $K \geq 1$ (again, usually the case), the three baseline constants become scaled up to K, 3K, and 5K, respectively. As these values are integers, their addition to a scaled-up calculated arctangent does not affect the decision as to which of the two bounding integers the rounded result is nearest to. Thus under these assumptions, it suffices to use only one table to decide which integer best approximates the scaled and rounded arctangent alone; the addition of the scaled baseline is done subsequently.

The following code segment calculates round(S * arctan(ratio)) for -sqrt(3) ≤ ratio ≤ sqrt(3), giving results in the range -K to K.

```
index :=  floor(domaingrain * abs(ratio));
predicted :=  arctanpredictor[index];
corrected :=  if tantable[predicted] >= abs(ratio) then
        predicted
    else
        predicted+1;
return(if ratio >= 0 then corrected else -corrected);
```

Here arctanpredictor[0:floor(domaingrain * sqrt (3))] is an integer array, and tantable[0:K] is a real array. They have been loaded as follows:

```
for i :=  0 thru floor(domaingrain * sqrt (3)) do
    atanpredictor[i] :=  round(S * arctan(i/domaingrain));
```

```
for i := 0 through K do
    tantable[i] := tan((i+0,5)/S);
```

The correctness of this program depends on the accuracy of the estimate, which here is determined by table lookup on an index that is the truncated scaled-up input ratio. This estimater must never overshoot, nor undershoot by more than 1. Overshooting is prohibited by the truncation. The undershooting accuracy is controlled by ensuring that the predictor array (the value of domaingrain) is sufficiently large so that no prediction is ever off by 2 or more; this is possible since the arctangent does not have a steep slope. (It can be noted that such an algorithm is impossible to implement with the original formulation of hue using arccosine, as that function has an unbounded slope at either end of its domain). In the above program there is a truncation in the domain, a mapping through a function whose slope is always ≤ 1, and a truncation in the range. Thus, "sufficiently large" means that the length of the interval of ratios encompassed by a single index should be no larger than the length of the interval of angles encompassed by a single integer output. As this latter is equal to 1/S, each index should span an interval of length 1/S or less, implying a value for domaingrain of at least S. In practice, however, the larger the array, the more efficient the calculation, since the predictions increase in accuracy and less searching is required. As S is usually a relatively small number, even a value of 10S is not excessively large as an array size.

The time savings of this algorithm should be apparent. The ten or so floating point operations, half of them divisions, normally required to calculate the arctangent are replaced with one multiplication, two table lookups, and some comparisons. The overhead for loading the arrays is usually insignificantly small compared to the amount of pixels transformed, as one array-filling calculation is about equivalent to the calculations of one pixel's hue by the direct computation of the arctangent.

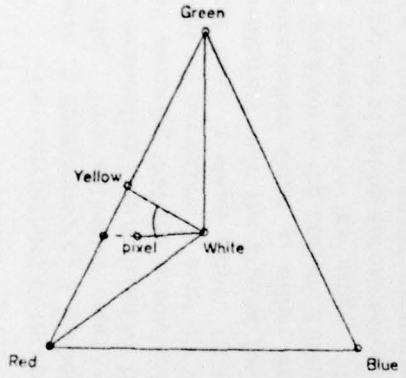


Figure B-1: Normalized color triangle with translated pixel.

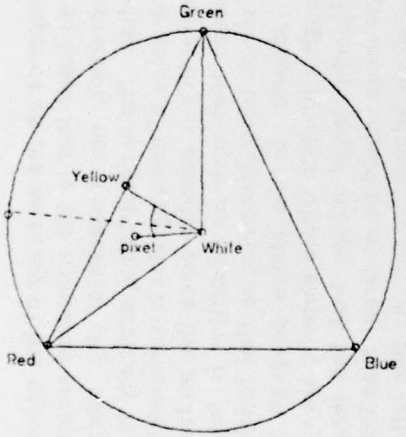


Figure B-2: Circumscribed color wheel with translated pixel.

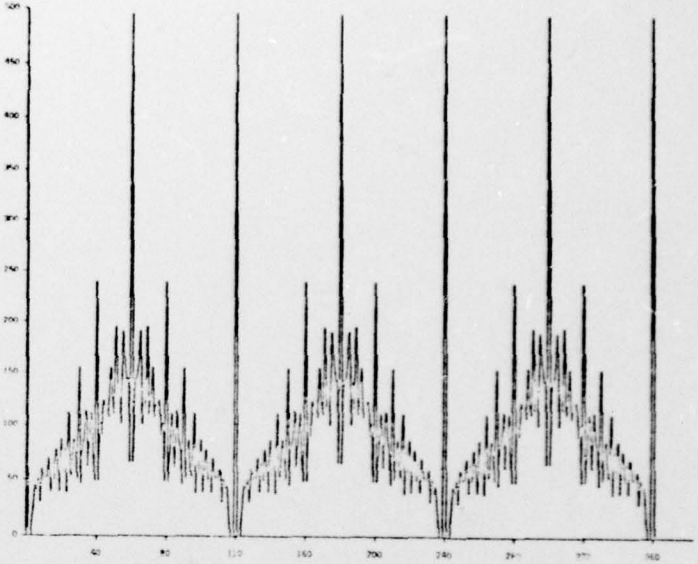


FIGURE B-3 DIGITAL "ARCTANGENTLESS" HUE OF UNIFORM INPUT, PARAMETERIZED AS IN FIGURE 4-1.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

(18) 1. REPORT NUMBER: AFOSR - TR - 77 - 0328

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

(6) 4. TITLE (and Subtitle)

SATURATION, HUE, AND NORMALIZED COLOR: CALCULATION, DIGITIZATION EFFECTS, AND USE

(9) 5. TYPE OF REPORT & PERIOD COVERED

Interim rept.

6. PERFORMING ORG. REPORT NUMBER

(10) 7. AUTHOR(s)

John R. Kender

(15) 8. CONTRACT OR GRANT NUMBER(s)

F44620-73-C-0074

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Carnegie-Mellon University
Computer Science Dept.
Pittsburgh, PA 15213

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

61102F
2304/A2

11. CONTROLLING OFFICE NAME AND ADDRESS

Defense Advanced Research Projects Agency
1400 Wilson Blvd
Arlington, VA 22209

(11) 12. REPORT DATE

November 1976

13. NUMBER OF PAGES

43

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

Air Force Office of Scientific Research (NM)
Bolling AFB, DC 20332

15. SECURITY CLASS. (of this report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

UNCLASSIFIED

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(12) 55p. (16) 2304 (17) A2

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT

Three transformations used in the analysis of tricolor natural scenes are analyzed. All have nonremovable singularities, near which they are highly unstable. Given digital input, the distribution of their transformed values is highly nonuniform, characterized by spurious modes and gaps. These effects are quantified and illustrated. In addition, a significantly faster algorithm for hue is derived. Image segmentation techniques of edge detection, region growing, clustering, and region splitting are affected arbitrarily badly by such problems. Some stratagems are illustrated that help minimize the bad behavior. Linear transformations are presented as a generally favorable alternative to these three nonlinear ones.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

403081

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)


FILM
4