# Surround Statistics and the Perception of Intensity and Color

James J. Clark and Diana Hernandez
Centre for Intelligent Machines
McGill University
Montréal, Québec, Canada

September 19, 2003

**Abstract**

We investigate conditional statistics of image patches. In particular we look at the conditional distribution of surround pixel values given the value of the center pixel. We propose that a process of statistical learning of these conditional distributions underlies the Stevens effect. We present simple ecological models for the conditional distributions, employing Jaynes Maximum Entropy principle. These models treat the processes of occlusion, shadowing, and inter-reflection, and qualitatively account for empirically observed relations between surround entropy and mean intensity. We also present a PCA appearance-based method for estimating patch hue from conditional models of surround spatial structure.

# 1 Introduction

In this paper we explore the statistical nature of the surround, or context, of a small image patch. Our goal is to see whether the surround statistics can provide information about point quantities associated with the center pixel in the patch, such as its lightness or color.

The reasons for wanting to obtain information about point quantities from surrounds are many. Measurement of image quantities over a small region (possibly a single pixel) may be noisy, whereas measurements taken over a larger surrounding region can be aggregrated to reduce the effect of noise. Furthermore, it is our view that the statistical learning of surround distributions provides a possible explanation of the Stevens effect, a well-known perceptual scaling of contrast with intensity. Finally, there are statistical measures, such as entropy, that are invariant to some transformations of the surround values, such as inversion. Thus, a system that perceives intensity through these measures of the surround will also be invariant to these transformations. This has implications for philosophical theories of perception.

The layout of the paper is as follows. The second section describes the Stevens effect and related effects. The third section motivates the use of surround entropies in modeling perception. Some experimental results taken on the van Hateran image database (van Hateran and van der Schaaf (1998)) are provided, illustrating the relationship between surround entropy and intensity. Section 4 looks at modeling the conditional statistics of the surround given the central (or mean) intensity of an image patch. We just provide a qualitative overview and emphasize the ecological nature of the physical processes which affect the conditional distributions. We focus our modeling mainly on the following processes: occlusion, shadowing, inter-reflection, and do not consider shading or albedo variation (texture). The final section of the paper describes preliminary work on a PCA-based method for modeling surround textures conditioned on mean hue.

# 2 Stimulus Entropy and the Stevens Effect

In a study that produced the effect bearing his name, Stevens (1961, 1963) found that subjects viewing a gray patch in a white surround perceived the contrast between the patches to increase as the intensity of the illumination. The background brightness was perceived to increase via a power law, with exponent 0.33, with respect to its luminance. The brightness of the gray patches, on the other hand, had a variable exponent, which became negative for darker patches. Overall, the effect is that the perceived contrast increased with the illumination intensity. Hunt (1995) observed an analogous effect in the perception of colored patches - he found that as overall intensity increased so did the perceived colorfulness.

It has long been informally conjectured that a form of "inverse-Stevens effect" exists. That is, *perceived intensity increases with image contrast.* As Fairchild (1999) points out, photographers often underexpose a high contrast scene (e.g. a dim indoor scene) and overexpose a low contrast scene (e.g. a bright outdoor scene). Fairchild did a psychophysical study to investigate this conjecture (Fairchild 1999). His results were inconclusive, however, showing a wide inter-subject variability. Some subjects had the supposed contrast-intensity relation while others had no relation, and still others had a relation in the direction opposite to that supposed.

Weber's law implies a visual system whose contrast sensitivity is constant with respect to intensity. Pattanaik *et al* (1998) propose that the variation of perceived contrast with intensity arises from an automatic gain mechanism, which results in a *sub-Weber's law* sensitivity behavior.

In the Pattanaik *et al* model contrast is measured using a pyramid of local differential operators. We take a different view, which perhaps has little practical difference, but a significant philosophical difference. We propose to use surround entropy as a measure of contrast. In particular, we want to suggest that Stevens effect is statistical in genesis, and is actually relating subjective intensity to the entropy of the visual stimulus. The idea is that, through visual experience, an observer learns an association between surface brightness and the entropy of the surface patch intensity values. Our thinking is motivated by the ideas of Norwich (Norwich 1984) who suggests that perception arises through reduction of uncertainty. In his view, a more intense stimulus has more uncertainty, and hence higher entropy. Furthemore, he proposes that the *subjective* impression of the intensity of a stimulus is related to the entropy of the stimulus.

If we take Norwich's ideas at face value are led to the hypothesis that the subjective impression of increased contrast with brighter images that comprises the Stevens Effect is a result of a learned association between contrast and entropy. That is, high contrast patches in natural images statistically tend to have higher entropies than low contrast image patches. This association is then linked to subjective intensity in Norwich's theory.

The entropy could be estimated through operations on a statistical model of variations obtained through experience. For example, intensity of a light could be related to the variability of scenes experienced when the illumination is of a given strength.

Irrespective of whether Norwich's ideas accurately model what is going on in the brain, the possibility of a statistical correlation between image patch entropy or contrast and average image patch intensity remains.

## 2.1 Surround Entropy in Natural Images

To see whether a case could be made for entropy variations with mean intensity as the cause of the Stevens effect, we carried out an empirical study of the conditional statistics of surrounds in a database of natural images. For our study we used a set of 136 images from

Figure 1: Samples of the images used in the empirical study. These images were taken from the van Hateran image database (van Hateran and van der Schaaf, 1998).

the van Hateran database (van Hateran and van der Schaaf, 1998). Each of these images had a size of 1024x1536 pixels. Figure 3 shows four of the images that were used in our study.

The raw image values were scaled by calibration factors provided with the database. These factors account for variations in sensitivity caused by aperture settings and shutter speeds, and permit us, to some extent, compare intensities across images in the database. We smoothed the scaled images with a 5x5 averaging kernel before computing entropies. This is to remove the effects of a non-uniform quantization scheme which would otherwise create a relative decrease in entropy with intensity, due to the greater spread between quantization levels at high intensities than at low. We also eliminated images which exhibited noticeable saturation at the high intensity end, as indicated by examination of the images' intensity

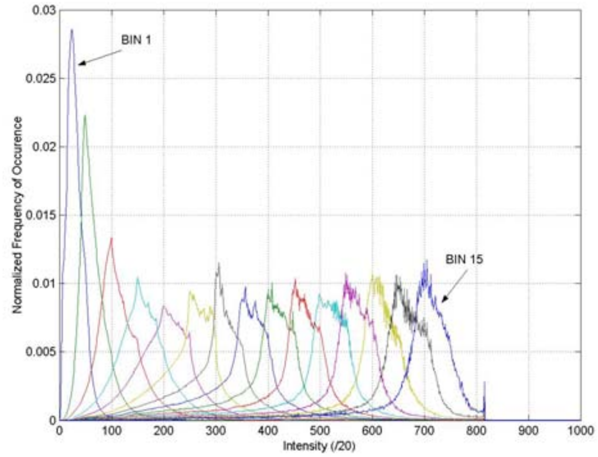


Figure 2: The conditional histograms of the surround pixel values given the central pixel value (for 15 different ranges of central pixel values).

histograms. Saturation results in an excessive number of pixels having the same value, which would reduce the compute entropy values, especially at high intensity levels. Although the images that were used in the study did not appear to exhibit any saturation, examination of the conditional histograms (in figure 4) show that there is still at least a low level of saturation, as indicated by the blip on the high end of the highest intensity conditional histogram.

To construct the conditional histograms for a given central value $I$, we searched the database images for pixels with values in the range $[I, I + \Delta I]$. Then we computed the histogram of the pixels in an 11x11 neighborhood centered on these pixels. These individual surround histograms were then summed to give the overall conditional histogram for the value $I$.

The conditional entropy is shown in figure 5. It is seen to rise, almost linearly, for low intensities, and then flatten out, and finally to drop once again.

It is important to realize that the observed variation in entropy with intensity is *not* due to a lack of dynamic range in the sensor. As we mentioned earlier, we deliberately omitted images from our test set which exhibited noticeable sensor saturation. The conditional variation of entropy must be due to other effects, and in the next section we propose that these effects are ecological in nature.

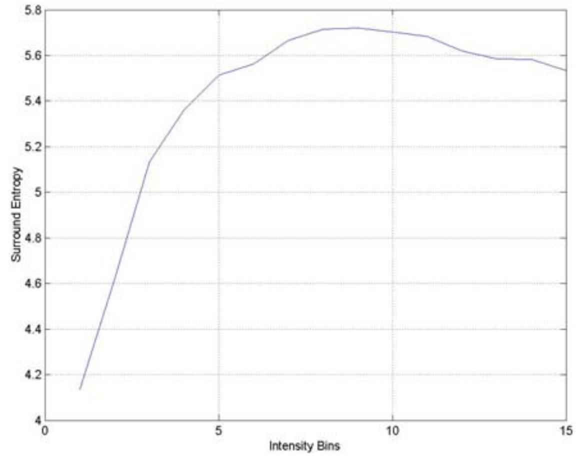


Figure 3: The entropy of the conditional distributions of surround values given the central value, for 15 different central value ranges.

# 3 Ecological Models of Surround Distribution

One could point out the initial rise in entropy in figure 5 as support for our theory of the genesis of the Stevens Effect, but one could also point out the late flattening and decrease as evidence against it. Our view is that the situation is complicated, as there are many factors which act to determine the surround distribution and hence its entropy. Not all of these are dependent on the surround brightness. It may be that the human visual system is able to factor out these various contributions and isolate those that are related to intensity. In this section we will look at some of these factors, and see what effect they have on surround entropy.

We will begin by attending to the problem of modeling the conditional distribution of the surround values given the central pixel value, $p(c|I)$. That is, given a patch intensity $I$, what is the distribution of surround values?

Let us first make the simple assumption that the mean of the surround for a given image patch intensity is the same for all surrounds, and that this mean is related to the patch intensity by some function $f(I)$. We could determine $f(I)$ empirically, by computing means over all surrounds of a given image patch intensity on a set of natural images, for example. This is shown in figure 6, where the data was obtained from a subset of the van Hateran natural image database.

It is evident that, at least for the empirical data that we gathered, $f(I) \approx I$. So, in the following, we will interchangably refer to the intensity of the center pixel, $I$ and the mean $f(I)$ of the patch.

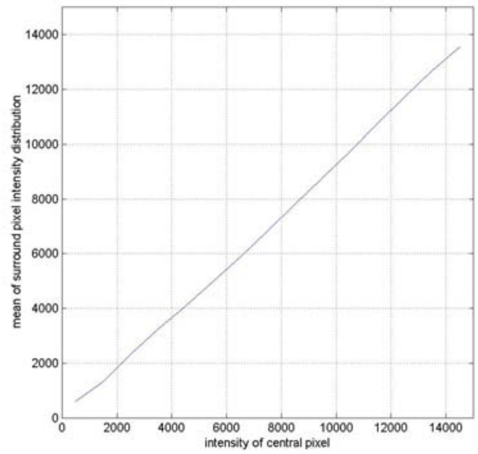


Figure 4: The mean of surround intensities over the image database conditioned on the value of the central pixel.

Following Jaynes (1968), we will construct our models using the Maximum Entropy approach. Jaynes argues that the form of a distribution is given by the distribution that has maximum entropy over all distributions that are consistent with all prior knowledge about the distribution.

Following this approach, we could start by assuming that the only knowledge we have is the empirical results depicted in figure 6, i.e. that the mean of the surround values, $< c >$, is equal to the value of the central pixel, $I$. In (Jaynes 1968) it is shown that the maximum entropy distribution subject to the constraint $< c >= I$ has an exponential form, $p(c|I) = \frac{1}{Z}\exp(\lambda c)$. The constant $\lambda$ can be computed so as to enforce the constraint. The term $Z$ is a normalization to ensure that $p(c|I)$ sums to one over all $c$ (assumed to take on non-negative integer values). The result is that, under this simple model,

$$p(c|I) = \frac{1}{1+I}\left(\frac{I}{1+I}\right)^c$$

The log of $p(c|I)$ is linear, with a negative slope, $\log(p(c|I)) = -\log(1+I) - c\log((1+I)/I)$. If we plot the log of an empirically obtained histogram of surround values, the result should also be linear with a negative slope, if the simple Jaynesian model is accurate.

In figure 7, we show the histogram of the (log) surround values over all of the images, and over all values of the central pixel value. It is clearly linear with a negative slope. Thus, the overall surround value statistics are well modeled as being constrained only by the mean being equal to the central value. This is effectively a smoothness or coherence constraint.

Figure 10 shows the entropy of the exponential distribution $p(c|I)$ as a function of the center pixel intensity (the $\alpha = 0$ curve). It can be seen that the simple exponential distribution models quite well the empirically obtained entropy curves. There are still differences,

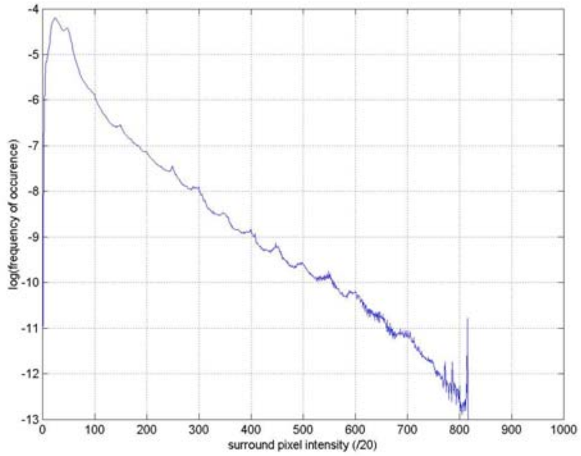


Figure 5: The log histogram of the surround intensities over the image database.

however, particularly in the middle and high intensity ranges.

Figure 8 shows the (log) histogram of the difference between surround pixel values and the center pixel value. This was obtained by shifting the conditional histograms of figure 4 to have their peaks line up with their respective center pixel intensities and then summing. It is very similar to the distribution found by Ruderman (1993) and Zhu and Mumford (1997) for the gradient of intensity, or that found by Lee *et al* (2001) of two point differences.

The theoretical entropy and difference histograms are misleading in their apparent good fit to the measured data, however. They hide differences in the underlying conditional distributions. In figure 9 we show the (log) conditional histograms for various values of the central pixel. The high surround intensity falloff is approximately linear, but there is a falloff at low surround intensities, which is stronger at high center pixel intensities. Therefore the conditional statistics are not well modeled by the simple Jaynes prior with mean constraint. Additional constraints must be provided to account for the falloff at low intensities.

To come up with a better model for the conditional statistics we have to examine the physical mechanisms by which the surround pixel values are generated, and see how, if at all, the center pixel affects the surround.

The primary ecological interactions affecting the spatial distribution of image intensities are:

- albedo variation (surface markings)
- occlusions
- shadowing

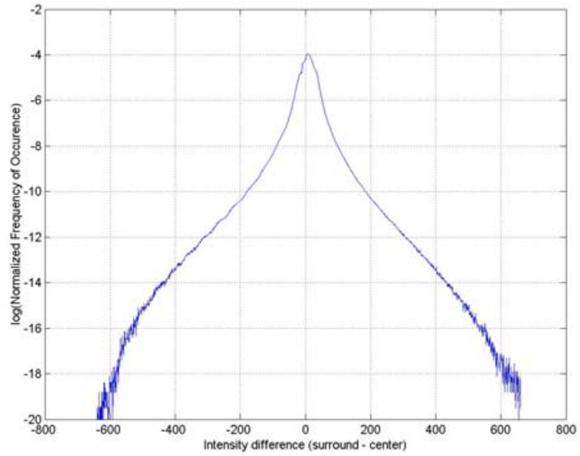


Figure 6: The log histogram of the differences between the surround and center pixel intensities over the image database. This histogram was obtained by shifting the conditional histograms of figure 4 to have their peaks line up with their respective center pixel intensities (and then taking the log).

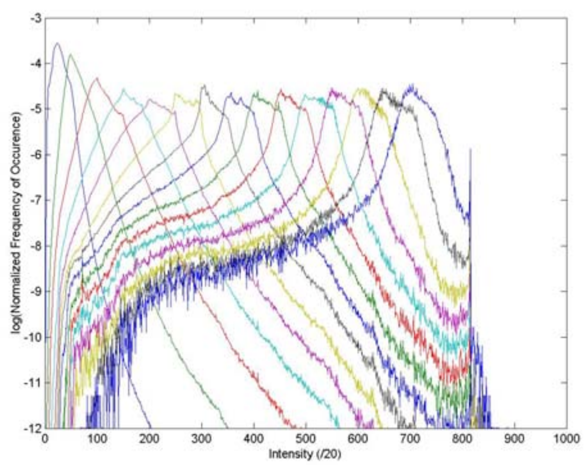


Figure 7: The log conditional histograms of the surround pixel values given the central pixel value (for 15 different ranges of central pixel values). These are the same as figure 3, save that the log of the histograms are shown.

- inter-reflection
- shading (surface normal variation)

In this paper we will concentrate on statistically modeling the effects of occlusion and shadowing, and inter-reflection. Occlusion and shadowing can be considered to be similar processes, in that occlusion is an interrupting of rays from the surface patch to the viewer, and shadowing is an interrupting of rays from the surface patch to the light source. Thus we expect that the statistical modeling of these two processes to be similar.

We will take a qualitative approach here, and focus on understanding the effects that each process has on the surround entropy - center pixel intensity relationship.

## 3.1 Occlusion

Occlusion is the process in which physical points in the scene are blocked from view by object that lie along the ray of light from the scene point to the viewer. A detailed statistical model of occlusion would have to model the size and distribution (and opacity) of objects in the world. Such a model was developed by Lee *et al* (2001). This 'Dead-Leaves' model closely simulates natural image statistics.

In its published form, the Dead-Leaves model is not quite suited for our problem, as it is not conditioned on the center pixel values (or on the surround mean). Thus, an extension to the model is needed, which provides the conditional distribution of the surround values given the surround mean or the central value, and from which the published distributions could be obtained through marginalisation.

We will not attempt to provide an extension of the Dead-Leaves model here, but will instead present a much simpler qualitative model, which is nonetheless based on similar principles. We will assume that a surround patch includes at most one occlusion edge, and that this single edge is straight. Thus, if an occlusion edge is present, the surround is divided into two sections, the larger of which includes the center pixel. To use the simplest possible model, let us assume that the section containing the center pixel (call it section 1) is unconstrained except that $< c_1 >= I$, where $I$ is the intensity of the center pixel. Likewise, let the other section (call it section 2) be constrained only by $< c_2 >= \tilde{I}$, where $\tilde{I}$ is the mean intensity over all pixels in the database. Let $\alpha < 0.5$ be the relative proportion of surround pixels in section 2 to the total number of surround pixels. The distribution of surround pixels is seen to be a mixture of two exponential distributions:

$$p(c|I;\alpha) = (1-\alpha)\frac{1}{1+I}\left(\frac{I}{1+I}\right)^c + \alpha\frac{1}{1+\tilde{I}}\left(\frac{\tilde{I}}{1+\tilde{I}}\right)^c$$

We could proceed further with our modeling, and specify a model for the distribution of $\alpha$ and then compute the marginal to give $p(c|I)$. We will not do that here, as all we are

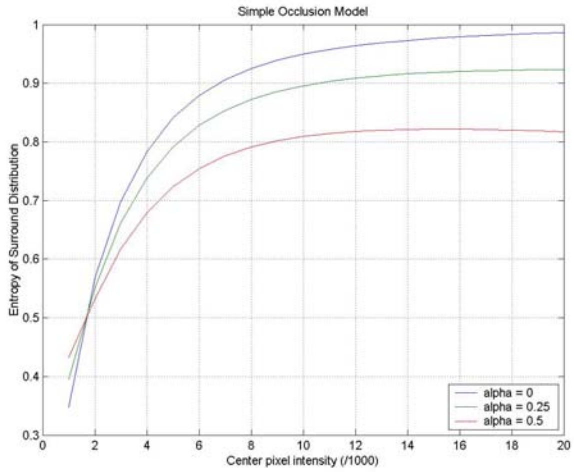


Figure 8: The conditional entropies computed from the simple two-region occlusion model. Shown are curves for three different occluding region proportions.

interested at the moment is to observe the qualitative effect of an occlusion on the entropy of $p(c|I)$. Clearly, if $\alpha = 0$ there will be no effect, and the effect will increase as $\alpha$ increases. In figure 10 we plot $p(c|I;\alpha)$ for a range of $\alpha$ values, with $\tilde{I}$ set to the empirically determined value of 1562.

In the figure we can see that the primary effect of occlusion in our simple model is to increase the amount of flattening of the entropy at high intensities. For $\alpha = 0.5$ there is even a slight downturn in entropy at the highest intensity.

## 3.2 Shadowing

The 'Dead-Leaves' model only models occlusion. It does not directly model shadowing, since the model creates images as a 2-D collage of overlapping disks. All shadows are occluded from view by the disks themselves. In order to introduce shadowing into the model, one needs to construct a 3-D collage model (a 'mobile' model?). While this is a worthwhile undertaking, we will not attempt it here. Instead, we will merely look at the general effect that shadowing will have on the entropy of the surround pixel distribution.

As in our simple model of occlusion, we will assume that a surround contains at most one shadow edge, and that this edge is straight. Thus, the edge divides the surround into two sections. The larger section, which we will call section 1, will contain the center pixel. Unlike the case of occlusion, where we took the surround pixels in section 2 to be constrained only by the overall mean, in the case of shadowing, the mean of the two sections will be related by a scale factor, which we call $\beta$. That is because, in the absence of occlusion, a shadow edge is assumed to fall on a single physical surface. Thus, the distribution of the

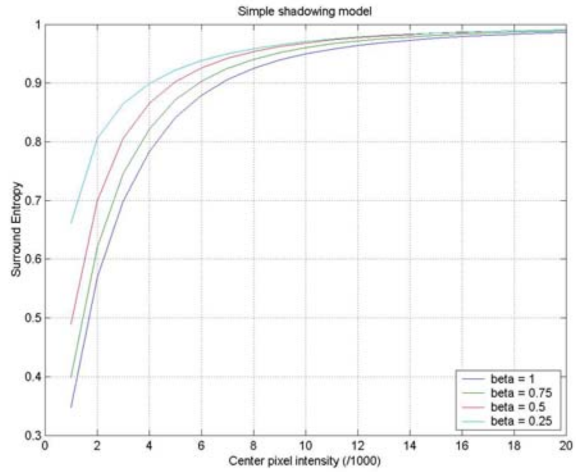


Figure 9: The conditional entropies computed from the simple shadowing model. Shown are curves for three different lit-shadow ratios.

pixels in section 1 should be similar to those in section 2, save for a scaling. The value of $\beta$ can range from $[0, \infty]$, as the central pixel can fall in the shadowed region ($\beta > 1$) or in the unshadowed region ($\beta < 1$).

$$p(c|I; \alpha, \beta) = (1-\alpha)\frac{1}{1+I}\left(\frac{I}{1+I}\right)^c + \alpha\frac{1}{1+\beta I}\left(\frac{\beta I}{1+\beta I}\right)^c$$

In figure 11 we plot $p(c|I; \alpha, \beta)$ for a range of $\beta$ values with $\alpha$ set to 0.5. It can be seen from the figure that the primary effect of shadowing in this simple model is to raise the entropy of low intensity regions. This is due to the identification of the pixels in section 2 (the region not containing the center pixel) as being out of the shadow, and therefore brighter than the center pixel. This spreads out the distribution, reducing the entropy. For bright center pixels, the effect is much smaller as there is already a significant frequency of low intensity pixels in the surround.

## 3.3 Inter-reflection

Inter-reflection invloves illumination of a surface patch by an illuminant through many paths. There is the direct path from the illuminant, as well as indirect paths bouncing off of other surfaces in the environment. Let us make a simple model of this process, where we consider only the direct path and first order reflections. Let us assume a Lambertian reflectance, where the reflectance is the same for all view angles. This means that the amount of light reflected from a surface patch to another is the same as that reflected from the surface patch to the viewer. It is not strictly neccessary to make this assumption, but it simplifies our development.

Let us model the conditional distribution $p(c|I)$of the intensities, $c$ of a single surround pixel. We will, as before, impose the constraint that $< c >= I$. Now, we will use a first order inter-reflection model for $c$, where we will assume that the surround pixel patch has an inter-reflection component of its illumination coming from the central pixel patch. Let the irradiance of the light source be $L$ and the reflectance of the pixel patches be $r$. Then, the surround pixel intensity can be expressed as:

$$c = L(r + r^2)$$

where the linear term corresponds to the direct reflection and the square term corresponds to the first order inter-reflection.

Let $< c >=< Lr + Lr^2 >= I$ as given earlier, and define $I' =< Lr >$. You can think of $I'$ as the mean surround intensity in the case that there is no inter-reflection. Thus we can write:

$$I =< Lr > + < Lr^2 >= I' + < Lr^2 >$$

and therefore

$$< Lr^2 >= I - I'$$

We can assume that $I$ will be proportional to $I'$ and that $I \geq I'$ since the effect of inter-reflection is to increase intensities, and the increase is proportional to the original intensity (bright surfaces provide more illumination than dark ones). Thus we can reasonably assume that

$$I - I' = \gamma I; \qquad \gamma < 1$$

The square of the surround intensity can be approximated to second order in $r$ by

$$c^2 = [L(r + r^2)]^2 = L^2(r^2 + 2r^3 + r^4) \approx L^2 r^2$$

Thus, the expected value of $c$ is

$$< c^2 >= L < Lr^2 >= L(I - I') = L\gamma I$$

Now, there is a linear relation between the illuminant irradiance $L$ and the mean intensity $I$. The more intense the illuminant, the more intense the reflected light. Therefore we can write $I = kL$, or $L = I/k$, with $k < 1$, and hence:

$$< c^2 >= k\gamma I^2$$

Thus the effect of inter-reflection is to impose an additional constraint on the distribution of $c$. Once again we can employ Jaynes' Maximum Entropy principal and find the distribution $p(c|I)$ which maximizes the entropy subject to the constraints $< c >= I$, $< c^2 >= k\gamma I^2$. It is straightforward to show that this distribution is a Gaussian, with mean $I$ and variance:

$$\sigma^2 =< c^2 > - < c >^2 = I^2(k\gamma - 1)$$

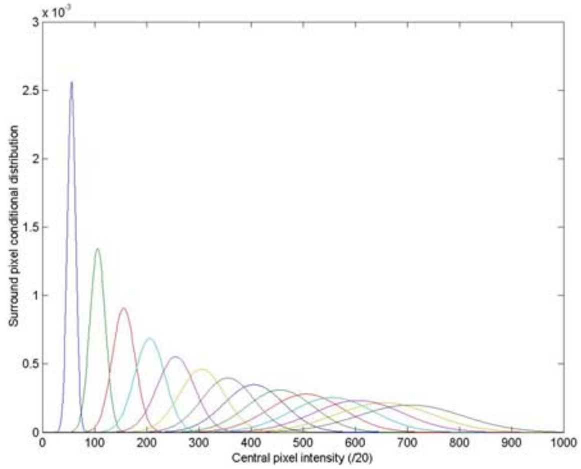


Figure 10: The conditional distributions computed from the simple inter-reflection model (for $k\gamma = 1.02$.

Note that the Gaussian form for the conditional distributions imply that there is a falloff away from the mean value. In particular, there is a falloff for small values of $c$. This replicates the empirical finding, and we can hypothesize that the falloff observed in the natural image data is due to inter-reflection. This makes intuitive sense, as the effect of inter-reflection is to brighten shadowed parts of the scene, thereby reducing the proportion of low brightness regions.

Further detailed modeling is needed to determine the value of $k\gamma$. We can see, however, that $k\gamma$ should be greater than 1 for there to be a falloff in the conditional distribution for small values of $c$. In figure 12 we show the conditional distributions predicted by this approach (for $k\gamma = 1.02$).

The empirically obtained conditional histograms shown in figure 4 are clearly not Gaussian, as they exhibit some asymmetry, especially for low $I$. But one can fit them quite well with Gaussians nonetheless. Thus we could reasonably hypothesize that inter-reflection does impose a Gaussian shape on the conditional distributions, with the empirically observed deviations from Gaussianity coming from shadowing and occlusion effects and unmodeled aspects of inter-reflection (higher order reflections).

The entropy of a Gaussian distribution is easily seen to be

$$H_{Gaussian} = \frac{1}{2}\log(2\pi) + \frac{1}{2} + \log(\sigma)$$

Thus, the entropies of the conditional distributions according to our simple inter-reflection model are given by:

$$H(I) = \frac{1}{2}\log(2\pi) + \frac{1}{2} + \log(k\gamma - 1) + 2\log(I)$$

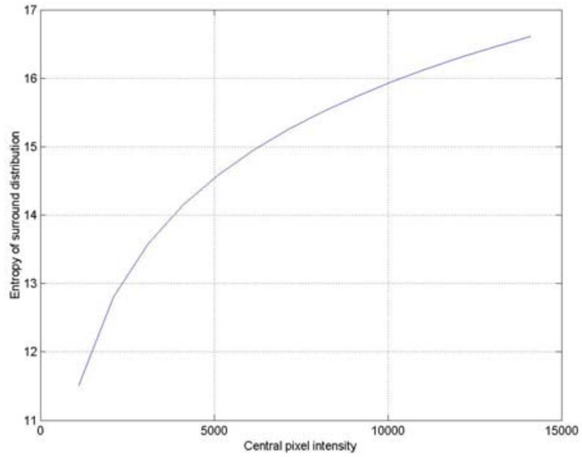


Figure 11: The conditional entropies computed from the simple inter-reflection model (for $k\gamma = 1.02$.

We plot the variation of these entropies with the central pixel intensity $I$ in figure 13. It matches the logarithmic shape of the empirical results (figure 5) for low intensity values. At higher intensities, other factors, such as occlusion, presumably are playing a role.

### 3.4 Shading

Shading induced variations in intensity arise from a dependence of the reflected light on the relative orientations of the surface normal vector of a surface patch, the view vector, and the illuminant vector. The surface normal vector presumably can be modeled as having a uniform distribution. Modeling of the view vector distribution is complicated by occlusions, and modeling of the illuminant vector distribution is complicated by shadowing. The modeling task must also include a model of the reflectance map, which provides the relationship between these vectors and the measured intensity. We will not provide any model for shading in this paper, but it remains a possible target for future research.

### 3.5 Albedo Variation (Texture)

In none of the preceding models did we consider the spatial structure of the pixels in the surround. They were all considered to be i.i.d. In natural scenes, however, there is significant spatial correlation between pixels in a surround. The physical processes that give rise to these correlations are many and varied. It is impractical to come up with a simple general model that can describe all of these processes.

Some progress has been made, however. Zhu *et al* (1998) describe a technique for modeling the local structure of image data using a Markov-Random-Field approach. In their approach they also use Jaynes' Maximum Entropy principal to define the surround distribution. Instead of applying mean or variance constraints, they apply constraints on the output of various filters acting on the surround. They treat the distribution so obtained as a sample of a Markov Random Field, which is then a model for the ensemble of surrounds obtainable in natural images. Some work will need to be done to adapt their technique to our problem, as their method provides joint distributions, while we need conditional distributions.

# 4 Surrounds Conditioned on Color

Up to this point in the paper we have been concentrating on the conditioning with respect to intensity. In color images we can consider conditioning with respect to color-related quantities such as hue and saturation. We propose conditioning on the hue, to produce a statistical learning algorithm that can estimate the (mean) hue of an image pixel from measures of its surround. Our current approach is to select a hue (e.g. red), and find all pixels in a database of images that have that hue. We then take the set of surrounds of each of these pixels and use these to compute principal components. This gives us a linear model for pixel surrounds conditioned on the hue value. That is, for each hue, $h$, we have a linear model for a surround $c(x, y)$

$$\hat{c}(x, y; h) = \sum_{i=1,N} w_i \phi_h^{(i)}(x, y)$$

where the $\phi_h^{(i)}, i = 1, ..., N$ are the $N$ principal components of the surrounds of pixels with hue $h$. Given an arbitrary surround, $c(x, y)$, we can project it onto the principal components to obtain the coefficients $w_i$, which can be thought of as coordinates in an eigenspace spanned by the $\phi_h^{(i)}$:

$$w_i = \int_S \left(\phi_h^{(i)}(x, y) c(x, y)\right) dx dy$$

where $S$ is the domain of the surround image.

With the linear models for a set of different hue values, we can apply the appearance-based recognition scheme of Nayar *et al* (1995) to a given surround to estimate the hue of the patch. The approach taken with this technique is to project the surround onto each of the sets of principal components, or eigenspaces, one for each hue. This will give a set of coefficients for each eigenspace. We then reconstruct the surround using the coefficients and the principal components. We can then determine which reconstruction best matches the original surround. We can then identify the hue associated with the eigenspace that produces the best reconstruction as the (estimated) hue of the surround.

$$\tilde{h}(c) = \arg\min_h ||c(x, y) - \sum_{i=1,N} w_i \phi_h^{(i)}(x, y)||$$

where $||\cdot||$ is a suitable norm on the space of surround sub-images.

A similar idea was used by Turk and Pentland (1991) to do face localization. They projected image patches onto a face eigenspace and measured the reconstruction error. Patches that were similar in form to faces were well represented in the eigenspace and therefore resulted in a low reconstruction error. Our method uses the same principal; we are localizing areas of a given hue instead of faces. From a statistical learning viewpoint the idea is that, through visual experience one can form eigenspace models of surrounds conditioned on quantities such as intensity and hue. Then, when presented with an arbitrary image patch, expectations will arise as to the intensity or hue of the patch.

To test our idea, we searched for 25x25 pixel surrounds around pixels with three different hues: Red, Green, Blue, in a database of 300 natural color images. There were about 2000 Red image patches, 600 Green image patches, and 500 Blue images patches selected from these images. From the selected surrounds we generated three sets of principal components, one for each of the three hues we considered. The first three eigen-surrounds for each of the three hues is shown in figure 14. Note that we have three components to each eigenimage. These correspond to the (R,G,B) channels in the image. try not to confuse the R,G, and B image channels with the Red, Green, and Blue hues! We also investigated eigen-images based on (H,S,V) image representations. The results are not shown here.

Once we constructed our hue eigenspaces, we extracted a number of image patches from an image that was not in the training set. Such an image is shown in figure 15, with the boundaries of the test patches superimposed. Figure 16 shows the reconstruction error for each of the patches in the test image for each eigenspace.

The hue associated with the eigenspace producing the minimum reconstruction error for each test patch is shown in Table 1.

| Patch Number | Hue Estimate | Comment |
|---|---|---|
| P1 | Blue | |
| P2 | Blue | (but Red and Green also give low error) |
| P3 | Blue | |
| P4 | Red | |
| P5 | Blue | |
| P6 | Red | (strong difference) |
| P7 | Green | (all errors quite high) |
| P8 | Red | (all errors quite high) |
| P9 | Red | |
| P10 | Blue | (strong difference) |
| P11 | Blue | (strong difference) |
| P12 | Green | |

Table 1. Hues associated with the hue eigenspaces that produce the minimum reconstruction

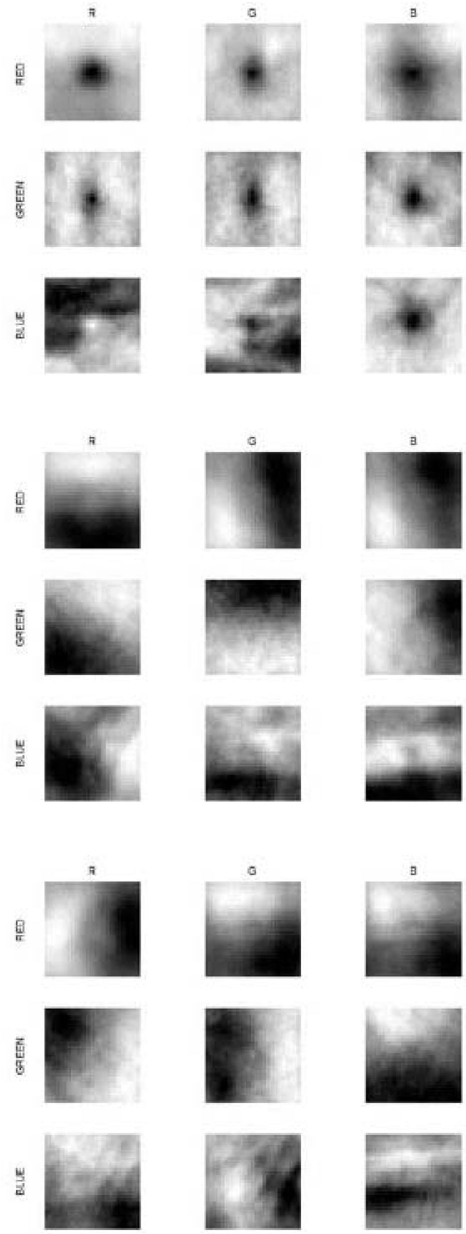


Figure 12: The first three eigenimage sets (one for each of the R,G,B channels) for the hue=RED, hue=GREEN, hue=BLUE eigenspaces. Top: first principal component, Middle: second principal component, Bottom: third principal component.

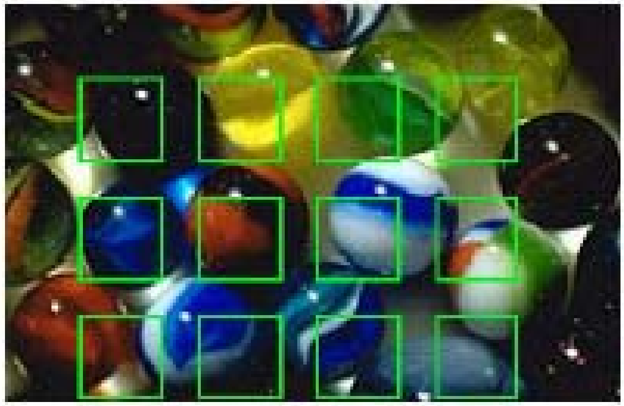

Figure 13: Image used to test the hue-from-surround estimation algorithm.

errors for the image patches outlined in figure 15. The numbers (P1, P2,...) corresponds to a left-to-right, top-to-bottom numbering of the patch outlines in figure 15.

The patch numbers given here are associated with the right-to-left, top-to-bottom ordering of the boxes in figure 16. We can see that the hue estimate scheme produces reasonable answers. Where the estimate might appear to be wrong (e.g. P2 - Blue instead of a Red/Green mix, or P7 - Green instead of Blue) is seen to to be in situations where the hue is other than one of the categories used (e.g. Yellow, or White), or in places where there is a mixture of hues (as in P8).

# 5 Summary

We proposed that the Stevens effect (and its analogous effect in colored scenes - the Hunt effect) has as its basis a statistical learning of the association between the mean intensity and contrast of a patch of pixels. Following the ideas of Norwich (1984) we propose that humans learn to associate the entropy of a stimulus with the intensity of that stimulus. We examined a small set of natural images and observed that there was, indeed, a relationship between mean surround intensity and the entropy of surround pixels which would explain the Stevens effect.

We derived a set of simple models that could explain the form of the conditional distributions observed in the natural images. These models described the effect of various ecological processes, namely occlusion, shadowing, and inter-reflection. Our models indicate that occlusion tends to reduce surround entropies for high mean intensities, that shadowing raises surround entropies for low mean intensities, and that inter-reflection imposes a low-intensity falloff in the conditional distributions. Taken together, these models can qualitatively account for the form of the empirically observed conditional distributions, as well as account

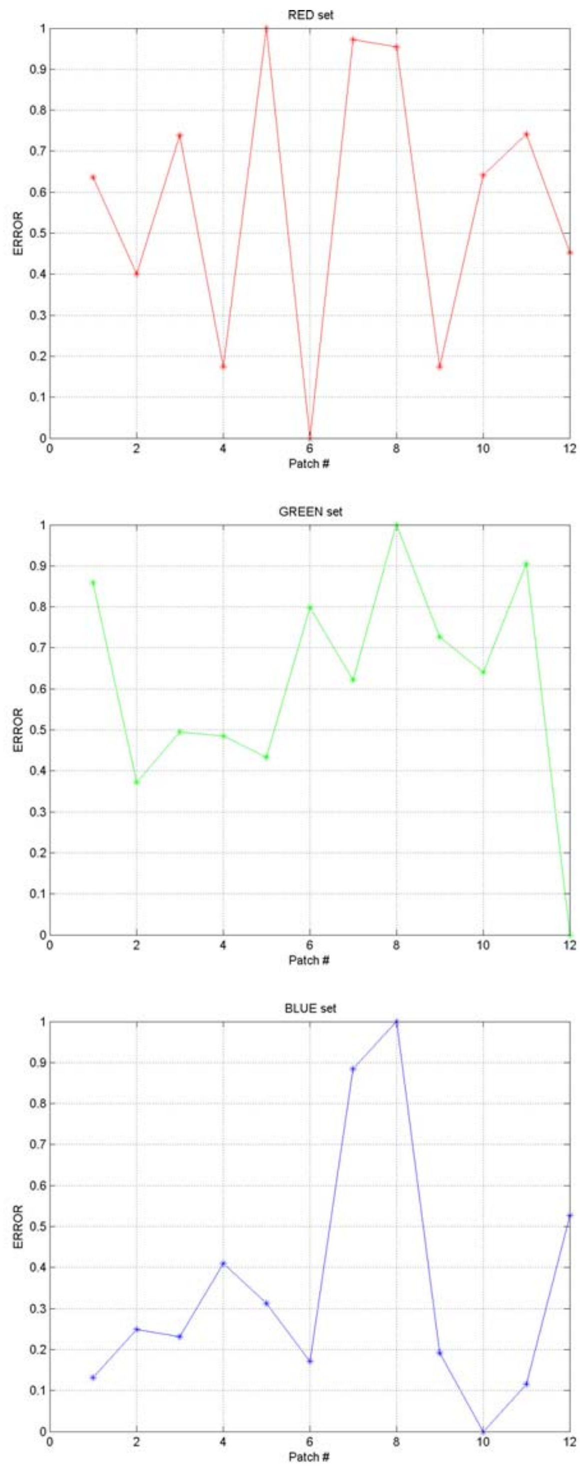


Figure 14: Reconstruction errors for each of the three hue eigenspaces for the image patches shown in figure 15.

for the variation of surround entropy with mean intensity observed in the natural images.

We presented an appearance-based method for obtaining statistical models of image patches conditioned on hue. This technique constructs linear models of the contexts by performing principal component analysis on hue-contingent surrounds extracted from a database of color images. An arbitrary patch can then be projected on to the hue eigenspaces and reconstructed. The eigenspace which produces the best reconstruction gives an implicit estimate of the patches' hue.

There is much that can be done to improve upon the work presented in this paper. In particular, the 'Dead Leaves' model of Lee *et al* could be extended to provide conditional distributions for occlusion, and to model the effects of shadowing. Higher order models of inter-reflection could be developed. The development of conditional texture models is a challenging and interesting possibility. The MRF texture models of Zhu *et al* (1998) seem suited for this. We plan to continue our investigation of conditional PCA models especially in its application to color.

# 6 Acknowledgements

The authors would like to acknowledge the valuable discussions on the topic of this paper held with Kevin O'Regan and Tal Arbel.

Funding for this project was provided by the Natural Sciences and Engineering Research Council (NSERC) of Canada, and by the Institute for Robotics and Intelligent Systems (IRIS).

# 7 References

Fairchild, M.D., "A Victory for Equivalent Background – On Average" IS&T/SID 7th Color Imaging Conference, Scottsdale, 87-92, 1999.

Hunt, R.W.G., **The Reproduction of Color**. 5th edition, Kingston-upon- Thames, England: Fountain Press, 1995.

Jaynes, E.T., "Prior Probabilities,' IEEE Trans. on Systems Science and Cybernetics, SSC-4, 227, 1968.

Langer, M. S., "A model of how interreflections can affect color appearance". Technical Report No. 70, Max-Planck Institute for Biological Cybernetics, Tuebingen, Germany, 1999.

Lee, A.B., Mumford, D., Huang, J., "Occlusion Models for Natural Images: A Statistical Study, of a Scale-Invariant Dead Leaves Model", International Journal of Computer Vision, 41(1/2), pp 35-59, 2001.

Nayar, S.K., Nene, S.A., and Murase, H., "Subspace Methods for Robot Vision," CUCS-06-95, Technical Report, Department of Computer Science, Columbia University, New York, February 1995.

Norwich, K., "The psychophysics of taste from the entropy of the stimulus", Perception and Psychophysics, 35 (3), 269-278, 1984.

Pattanaik, S.N., Ferwerda, J.A., Fairchild, M.D., and Greenberg, D.P., "A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display", Proceedings of SIGGRAPH'98, pp. 287-298, Orlando, July 1998.

Ruderman, D.L., "The statistics of natural images". Network, 5(4):517–548, 1993.

Stevens, S.S.. "To Honor Fechner and Repeal His Law". Science, 133, 13 Jan., pp 80-86, 1961.

Stevens, S.S. and Stevens, J.C., "Brightness function: Effects of Adaptation", Journal of the Optical Society of America, 53(3), p375, 1963.

Turk, M. and Pentland, A., "Face recognition using Eigenfaces", Proc. IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, 1991.

van Hateren, J.H. and van der Schaaf, A., "Independent component filters of natural images compared with simple cells in primary visual cortex". Proc.R.Soc.Lond. B 265:359-366, 1998.

Zhu, S.C., and Mumford, D.B., "Prior learning and Gibbs reaction-diffusion", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.19, no.11, pp1236-1250, Nov. 1997

Zhu, S.C., Wu, Y.N. and Mumford, D.B., "FRAME: Filters, Random field And Maximum Entropy: — Towards a Unified Theory for Texture Modeling", International Journal of Computer Vision, 27(2) pp.1-20, March/April, 1998.