

On the Natural Statistics of Chromatic Images

Zeina Sinno and Alan C. Bovik
Laboratory for Image and Video Engineering
Department of Electrical and Computer Engineering
The University of Texas at Austin

Abstract—The visual brain is optimally designed to process images from the natural environment that we perceive. Describing the natural environment statistically helps in understanding how the brain encodes those images efficiently. The Natural Scene Statistics (NSS) of the luminance component of images is the basis of several univariate and bivariate statistical models. The NSS of other colors or chromatic components have been less well-analyzed. In this paper, we study the univariate and bivariate NSS of luminance and other chromatic components and how they relate.

Keywords—Chromatic Natural Scene Statistics

I. INTRODUCTION

While most mammals are dichromatic, primates are unique in their ability to process trichromatic color information. Scientists believe that this is the result of evolution; early primates gained trichromatic vision because recognizing fresh fruit and immature leaves led to a more nutritious diet. Our visual system has evolved to optimally process images as they are projected from the natural environment. As with luminance properties, the color statistics of natural scenes must have had a significant impact on both the development and the evolution of color vision mechanisms. As observed by Fairchild [1], there is a clear link between the mechanisms of chromatic adaptation and knowledge of objects that are perceived and their illuminated environment. Color adaptation and color constancy are examples of such physiological mechanisms. There has been a significant amount of work towards understanding the color statistics of natural scenes, motivated by the observation that the average color and the spectral distribution trigger physiological mechanisms [2].

Ruderman developed a model to characterize the statistics of cone responses, and their role in the visual encoding of images. Su [3] studied the statistics of color and range for 3D images and modeled the relationship between them. Towards better understanding the efficient visual coding of color information, we study the statistics of multiple chromatic components. Specifically, we model cortical processing of independent chromatic components, and observe how the results correlate with each other. It has been shown that there exists a close correlation between naturalness, colorfulness and the perceived quality of images [4]. This paper is the first step towards building a system that can restore the colors of images suffering from color distortions, by

imposing constraints on the univariate and bivariate NSS of the different color components, and on the correlation between them. Chroma sub-sampling for example, can result in artifacts such as color bleeding in images and videos. We plan using NSS to reconstruct an improved image that would look as natural as possible. Our findings can also be used to improve the performance of existing other luminance based NSS image quality assessment models [5]. For example, it may be possible to weight different chromatic attributes to determine a more holistic measure of perceptual quality [6].

II. RELEVANT WORK

There has been a number of studies of the statistics of the luminance component of natural images when subjected to cortical processing. It has been well established that the distribution of subband images is heavy-tailed. If divisive normalization is applied on subband images, e.g., as a model of the nonlinear adaptive gain control of V1 neuronal responses in visual cortex [7], the distribution becomes more Gaussian as shown in [8]. When distortions are applied to images, this distribution becomes non-Gaussian. As noted by Field and Tolhurst, [9], [10] the amplitude spectra of luminance images follows a $1/f$ model, which underlies several processing models [11], [12]. There are also more recent closed form statistical models [13], [14], [15] that describe the correlation structure of natural images. Here, we study existing univariate [8] and bivariate closed-form NSS models [15] in regards to how they may be applied to the chromatic components of images.

III. NATURAL SCENE STATISTICS MODEL

First, we review the computational steps that form the basis of our NSS models of the chromatic components of images. We used the 29 pristine images of the LIVE IQA database [16] as a resource of high quality images and considered five different color planes; luminance, the two chroma components of the CIELAB space and the two chromatic components of the CIELUV space.

A. The Normalized Bandpass Univariate Model

First, extract a color component of interest from an image. Next, consider a bandpass derivative of gaussian filter DoG filter. At a fixed scale σ , the filter is defined as:

$$DoG(\mathbf{x}, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} - \frac{1}{\pi\sigma^2} e^{-\frac{(x^2+y^2)}{\sigma^2}}, \quad (1)$$

Each image is processed via a bank of DoG filters of scales $\sigma \in [1, 2, \dots, 6]$. The filter window size is increased linearly as a function of σ .

Next, apply divisive normalization on all of the DoG responses to model the nonlinear adaptive gain control of V1 neuronal responses in the visual cortex. The divisive normalization model is defined as:

$$u_j(\mathbf{x}) = \frac{w_j(\mathbf{x})}{\sqrt{s + \sum_{\mathbf{y}} g(j(\mathbf{y})w_j(\mathbf{y}))^2}}, \quad (2)$$

where w_j are the DoG responses from filter indexed j , u are the coefficients obtained after divisive normalization, and $s = 10^{-4}$ is a stabilizing saturation constant. The weighted sum in the denominator is computed over a spatial neighborhood of pixels from the same sub-band, where $g(x_i, y_i)$ is a circularly symmetric Gaussian function having unit volume. To match the increase in scale applied at the steerable filtering step (translated by increasing σ), the variance of $g(x_i, y_i)$ is also increased linearly as a function of σ .

We then modeled the distributions of the obtained coefficients, using a zero-mean generalized gaussian density (GGD):

$$f(x; \phi; \gamma^2) = \frac{\phi}{2\eta\Gamma(1/\phi)} \exp\left[-\left(\frac{|x|}{\eta}\right)^\phi\right] \quad (3)$$

where

$$\eta = \gamma \sqrt{\frac{\Gamma(\frac{1}{\phi})}{\Gamma(\frac{3}{\phi})}} \quad (4)$$

and $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt \quad a > 0. \quad (5)$$

The shape parameter ϕ controls the shape of the distribution, while η controls its variance. We used the moment matching approach [17] to estimate these two parameters from the histograms of each considered image's coefficients.

B. Normalized Directional Bandpass Image Correlation Model

The bivariate NSS model [15] has been developed and validated on the luminance component of images. Here, we extend this model to other chromatic components.

First, deploy a bank of steerable filters [18] to decompose a luminance image. Steerable filters are often used to model bandpass simple cells in primary visual cortex [19]. These bandpass filters, are similar to the DoG filters used in the previous section, but are also directional. A steerable filter at a given frequency tuning orientation θ_1 is defined by:

$$F_{\theta_1}(\mathbf{x}) = \cos(\theta_1)F_x(\mathbf{x}) + \sin(\theta_1)F_y(\mathbf{x}), \quad (6)$$

where $\mathbf{x} = (x, y)$, and F_x and F_y are the gradient components of a two-dimensional unit-energy bivariate isotropic gaussian function having a scale parameter σ :

$$G(\mathbf{x}) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \quad (7)$$

Each analyzed image is passed through a bank of steerable filters of scales $\sigma \in [2, 3, \dots, 6]$ and over 15 frequency tuning orientations $\theta_1 \in [0, \pi/15, 2\pi/15, \dots, \pi]$, yielding 90 bandpass responses per image. We exclude $\sigma = 1$, since steerable filters become less well defined at that scale. Then apply divisive normalization, similarly to (2).

Next, define a window at a fixed position within the cropped image and another sliding window of the same dimensions. Denote the distance between the center of the two windows of bandpass, normalized image samples of interest by d , and the angle between them by θ_2 . Also define the relative angle $\theta_2 - \theta_1$, where θ_1 is the sub-band tuning orientation relative to the horizontal of the bandpass filter. Then, compute the correlation between the two windows. The two windows are separated by horizontal and vertical separations δ_x and δ_y , which are varied over the integer range 1 to 25, i.e, distances of $\sqrt{\delta_x^2 + \delta_y^2}$ at spatial orientations $\theta_2 = \arctan(\frac{\delta_y}{\delta_x})$ (relative to horizontal). We limited the range $\theta_2 \in [0, \pi[$ since the quantities being measured are symmetrically defined. We have observed in [15] that 4 angles (the horizontal, the vertical and the two diagonals) contained the most stable structure information about the images so we limited the θ_2 to these values only.

The correlation function model expresses a periodic behavior in the relative angle $\theta_2 - \theta_1$, and can be modeled as:

$$\rho(d, \sigma, \theta_2) = A(d, \sigma, \theta_2) \cos(2(\theta_2 - \theta_1)) + c(d, \sigma, \theta_2) \quad (8)$$

where $A(d, \sigma, \theta_2)$ is the amplitude, $c(d, \sigma, \theta_2)$ is an offset, d is the spatial separation between the target pixels, σ is the steerable filter spread parameter, and θ_2 is as before.

Then, define the peak correlation function:

$$P = \max(\rho) = A + c. \quad (9)$$

wherein we may rewrite (8) as:

$$\rho(d, \sigma, \theta_2) = A(d, \sigma, \theta_2) \cos(2(\theta_2 - \theta_1)) + [P(d, \sigma, \theta_2) - A(d, \sigma, \theta_2)]. \quad (10)$$

Lee, Mumford and Huang [20] systematically observed that the sample covariances of bandpass image pixels follow an approximate reciprocal power law, of the form $\frac{1}{|d|^b}$. Here, we model the peak correlation function as having a more stable form $\frac{1}{|d|^{\beta+1}}$. Given a fixed spatial orientation θ_2 and a scale σ , define

$$\hat{P}(d, \sigma, \theta_2) = \frac{1}{\left(\frac{d}{\alpha_0(\theta_2)*\sigma}\right)^{\beta_0} + 1} \quad (11)$$

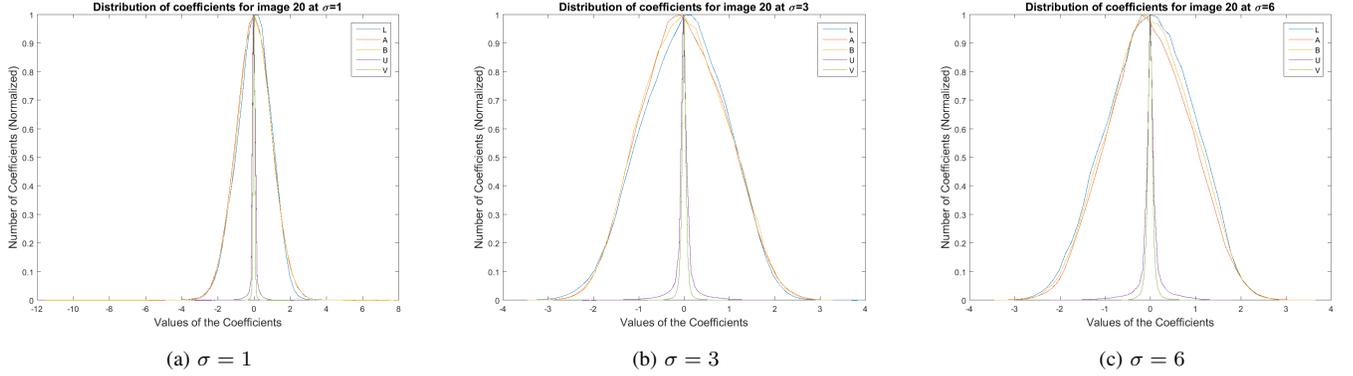


Figure 1. Histograms normalized bandpass coefficients for an exemplar image at three scales.

where $\{\alpha_0, \beta_0\}$ are parameters that control the shape and fall-off of the peak correlation function, which depend on the spatial orientation θ_2 .

We model A as the difference of two functions of the form (11):

$$\hat{A}(d, \sigma, \theta_2) = \frac{1}{\left(\frac{d}{\alpha_1(\theta_2)*\sigma}\right)^{\beta_1(\theta_2)} + 1} - \frac{1}{\left(\frac{d}{\alpha_2(\theta_2)*\sigma}\right)^{\beta_2(\theta_2)} + 1} \quad (12)$$

where $\{\alpha_1, \beta_1, \alpha_2, \beta_2\}$ are parameters that control the shape of A and are functions of θ_2 .

Our next goal is to find, for a fixed spatial orientation θ_2 , the values of the parameters $\{\alpha_0, \beta_0\}$ that produce the best fit to (11) and the parameters $\{\alpha_1, \beta_1, \alpha_2, \beta_2\}$, yielding the best fit to (12), resulting in the least mean squared error. We form two optimization systems for P and A that account for scale to find those optimal values, that minimize the error. Denote by D the set of distances for a given spatial orientation θ_2 . Our optimization systems are expressed as:

$$\begin{aligned} \min_{\alpha_0, \beta_0} & \sum_{d \in D} \sum_{\sigma=2}^6 (P(d, \sigma, \theta_2) - \hat{P}(d, \sigma, \theta_2))^2 \\ \min_{\alpha_1, \beta_1, \alpha_2, \beta_2, b_2} & \sum_{d \in D} \sum_{\sigma=2}^6 (A(d, \sigma, \theta_2) - \hat{A}(d, \sigma, \theta_2))^2 \end{aligned} \quad (13)$$

We validated the bandpass correlation model in [15] and we were able to verify that we can reconstruct the correlation with very low mean squared error.

IV. COMPARISON OF THE STATISTICS OF DIFFERENT COLOR SPACE COMPONENTS

Here, we study and compare the univariate and bivariate statistics of the chromatic components of natural images. We begin with the normalized univariate bandpass statistics. We

observed for each image and scale how the the distribution of the coefficients varies as a function of the color planes. We noticed a high degree of similarity in the distributions of the luminance component and the chromatic a and b components, and a high similarity between the chromatic u and v components. Fig. 1 shows an example of this behavior based on one content from the LIVE IQA database, ‘‘Sailing2’’. We considered three scales: $\sigma = 1, 3$, and 6. Similar observations were also observed on the other 28 contents and for the other considered σ values.

Next, we wanted to determine whether meaningful correlations exist between the measured GGD parameters (shape and variance) in (3) of the chromatic planes. We did not find meaningful correlations between the parameters, except for η_A and η_B which seem to be linearly correlated.

Next, we considered the correlation ρ obtained from the normalized directional bandpass image correlation model. Fig. 2 plots ρ at $d = 1$, and $\theta_2 = 0$ for $\sigma = 2, 4$, and 6 for the ‘‘Sailing2’’ content. Notice that there is a considerable overlap between the curves resulting in a resemblance in the behavior of P between the different components as well as A . Luminance appears to be slightly less correlated at smaller scales as compared to the other color planes. This may be expected since the luminance component contains the most detailed structural information. This diversity leads to slightly less overall correlation. Also as d is increased, the curves overlap more. As a next step, we plan to extend this analysis to images suffering from color distortions, and explore whether it would be possible to correct for those artifacts using chromatic NSS.

V. CONCLUSION

We compared univariate and bivariate NSS models on different color planes. We found the univariate NSS of the different color planes (CIELAB and CIELUV) to be distinguishable, but found similar behavior among the luminance, chroma (a, b) components and among the chroma (u, v) components. We also found the bivariate correlation NSS

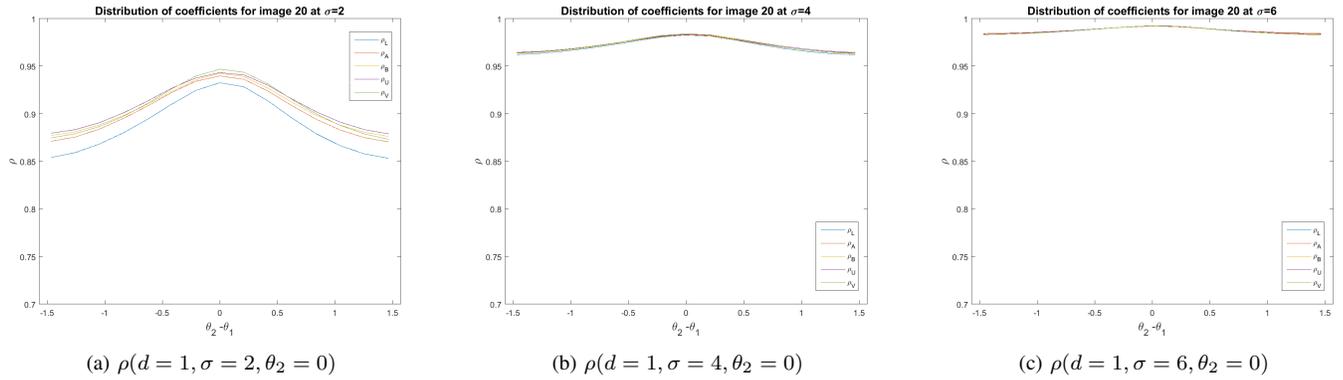


Figure 2. $\rho(d=1, \sigma=4, \theta_2=0)$ for $\sigma=2, 4,$ and 6 for the “Sailing 2”

of the different colors spaces to be very similar. As a next step, we plan to study the NSS of chromatic images suffering from color distortions and augment other luminance based NSS image quality assesment models [5] to improve their performance.

REFERENCES

- [1] M. D. Fairchild and P. Lennie, “Chromatic adaptation to natural and incandescent illuminants,” *Vision R.*, vol. 32, no. 11, pp. 2077–2085, 1992.
- [2] J. Von Kries, “Chromatic adaptation,” *Sources of Color Vision*, pp. 145–148, 1970.
- [3] C. C. Su, A. C. Bovik, and L. K. Cormack, “Natural scene statistics of color and range,” in *18th IEEE Inter. Conf. Imag. Process. (ICIP)*. IEEE, 2011, pp. 257–260.
- [4] E. A. Fedorovskaya, H. de Ridder, and F. J. J. Blommaert, “Chroma variations and perceived quality of color images of natural scenes,” *Col. Res. App.*, vol. 22, no. 2, pp. 96–110, 1997.
- [5] Z. Sinno, C. Caramanis, and A. C. Bovik, “Second order natural scene statistics model of blind image quality assessment,” *IEEE Conf. Acous. Spee. Sig. Proc. (ICASSP)*, Apr. 2018.
- [6] P. G. Engeldrum, “Extending image quality models,” *IS AND TS PICS CONFERENCE*, pp. 65–69, 2002.
- [7] M. Carandini, D. J. Heeger, and A. J. Movshon, “Linearity and normalization in simple cells of the macaque primary visual cortex,” *J. Neurosci.*, vol. 17, no. 21, pp. 8621–8644, 1997.
- [8] A. K. Moorthy and A. C. Bovik, “Blind image quality assessment: From natural scene statistics to perceptual quality,” *Image Processing, IEEE Transactions on*, vol. 20, no. 12, pp. 3350–3364, 2011.
- [9] D. Field, “Relations between the statistics of natural images and the response properties of cortical cells,” *J. Opt. Soc. Am.*, vol. 4, no. 12, pp. 2379–2394, 1987.
- [10] D. J. Tolhurst, Y. Tadmor, and T. Chao, “Amplitude spectra of natural images,” *Opt. Phys. Opt.*, vol. 12, no. 2, pp. 229–232, 1992.
- [11] H. B. Barlow, “The coding of sensory messages,” in *Curr. Prob. in Anim. Behav.*, 1961.
- [12] B. A. Olshausen and D. Field, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [13] Z. Sinno and A. C. Bovik, “Generalizing a closed-form correlation model of oriented bandpass natural images,” *IEEE Glob. Conf. Sig. Info. Proc.*, Dec 2015.
- [14] Z. Sinno and A. C. Bovik, “Relating spatial and spectral models of oriented bandpass natural images,” *Southw. Symp. Image Anal. Interpret.*, Mar 2016.
- [15] Z. Sinno, C. Caramanis, and A. C. Bovik, “Towards a closed form second-order natural scene statistics model,” *IEEE Trans. Imag. Proc.*, in revision.
- [16] H. Sheikh, M. Sabir, and A. Bovik, “A statistical evaluation of recent full reference image quality assessment algorithms,” *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov 2006.
- [17] K. Sharifi and A. Leon-Garcia, “Estimation of shape parameter for generalized gaussian distributions in subband decompositions of video,” *IEEE Trans. on Circ. Sys. Vid. Tech.*, vol. 5, no. 1, pp. 52–56, 1995.
- [18] W. T. Freeman and E. H. Adelson, “The design and use of steerable filters,” *IEEE Trans. Pattern Anal. Machine Intell.*, no. 9, pp. 891–906, 1991.
- [19] M. Clark and A. C. Bovik, “Experiments in segmenting texton patterns using localized spatial filters,” *Patt. Recog.*, vol. 22, no. 6, pp. 707–717, 1989.
- [20] A. B. Lee, D. Mumford, and J. Huang, “Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model,” *Int. J. Comput. Vision*, vol. 41, no. 1-2, pp. 35–59, 2001.