

Statistical approaches to image and scene manipulation

Erik Reinhard*

February 24, 2001

The field of visual perception aims to study the human visual system. It is recognized that in order to understand a system, it is important to understand its input. This has led to the sub-field of natural image statistics, which tries to understand images in statistical terms. Images' statistical properties are usually studied by collecting a number of such images in an ensemble and computing first, second or higher order statistics on them.

One of the most well studied natural image statistics is that of the power spectrum, a second order statistic. By computing the Fourier transform of an image and multiplying each element of the transform by its complex conjugate, the power spectrum is obtained. Averaging over all directions gives power as function of frequency. For most natural images, plots of this quantity result in straight lines with a slope of around $1/f^2$, where f is the spatial frequency (usually measured in cycles per image) [2–4]. This is a special property of natural images, which is generally not obtained for random images (such as for example random noise images, which would produce a flat power spectrum). The $1/f^2$ spectral slope of natural images means that equal power is encoded in each frequency band. It also implies that natural images are statistically scale-invariant.

While the reason for this to occur is still being debated (e.g. [1]), these results are important for graphics applications. It has been shown that certain image interpretation tasks are negatively affected when the spectral slope deviates too much from -2 [6, 7]. It is therefore argued that the human visual system expects to see images that conform to this statistic. As graphics applications produce input to the human visual system, it makes sense to optimize this input to be as easy to interpret as possible. Conforming to second order statistics is one step in that direction.

Although individual images may have spectral slopes that are somewhat different from $1/f^2$, the variation between images is relatively small. Hence, the power spectrum provides a simple and elegant means to assess the quality of an image: if the power spectrum does not yield a straight line and if the slope of this line deviates much from -2 , then the image is less realistic in a statistical sense.

Recent research has shown that the power spectrum is sensitive to modeling, but is quite insensitive to variations in rendering. Different levels of image compression

*University of Utah, reinhard@cs.utah.edu

schemes (gif, jpeg), anti-aliasing, gamma-correction, and choice of rendering algorithm all have at most a small impact on this second order statistic [9]. On the other hand, changes in geometry tend to have a much larger influence on the power spectrum. This establishes the power spectrum as a useful tool to assess quality of modeling. It may therefore be used to guide modelers to construct scenes with content that is perceived to be natural. Applications include parameter optimization for fractal terrains, displacement mapping and solid texture creation.

The quality of modeling is only one factor that influences the perceived naturalness of rendered images. A second issue with many synthetic images involves the selection of color which may affect image quality. Here, statistical methods may be used to impose the look and feel of a photograph upon a rendered image, thus improving perceived realism. Computing simple statistics on the colors of a photograph and applying those to the colors of a rendered image, may convert an obviously synthetic image into a more realistic looking image [8].

The method to transfer the look and feel of one image to another, involves conversion of both source and target images to a perceptually based color space where the three color axes are decorrelated. Decorrelation can be achieved by applying a Principal Components Analysis to each image. However, Ruderman et. al. [10] have shown that for natural image ensembles, the resulting axes have simple forms and interpretations, forming a new color space. The first principal component is an achromatic channel, while the second and third are yellow-blue and red-green color opponent pairs. Rather than computing PCA for each image separately, images given in LMS color space can be converted to a color space which on average decorrelates its axes using the following conversion matrix [5, 10]:

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \log L \\ \log M \\ \log S \end{bmatrix} \quad (1)$$

Note that the conversion also takes the pixels to log space, which conditions the data by no longer requiring the data values to be positive. The data representation is more compact and in log space uniform changes in stimulus tend to be equally detectable. Thus, the $l\alpha\beta$ color space promises to be useful for color manipulation.

The fact that the axes in $l\alpha\beta$ color space are decorrelated is important, because it allows data along each axis to be modified separately, without affecting the other two axes. To transfer statistical properties between images, we compute mean and standard deviation of the pixels along each axis separately for both rendering and photograph. Then, all data points of the rendering are scaled and shifted to assume mean and standard deviation of the photograph. By properly matching a rendering with a photograph, convincing results may be obtained, as shown in Figure 1.

The work summarized in this abstract indicates that research into the statistical properties of natural scenes provides important clues as to how graphics algorithms may be optimized to generate imagery that is well-suited to be interpreted by the human visual system. Statistical approaches may provide a good alternative to assess image

quality, generate scenes and modify existing images. Many more applications for this work can be anticipated.

References

- [1] Rosario M Balboa, Christopher W Tyler, and Norberto M Grzywacz. Occlusions contribute to scaling in natural images. *Vision Research*, 2001. In press.
- [2] G J Burton and I R Moorhead. Color and spatial structure in natural scenes. *Applied Optics*, 26(1):157–170, January 1987.
- [3] Dawei W Dong and Joseph J Atick. Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6(3):345–358, 1995.
- [4] David J Field. Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems. In M Farge, J C R Hunt, and J C Vassilicos, editors, *Wavelets, fractals and Fourier transforms*, pages 151–193. Clarendon Press, Oxford, 1993.
- [5] P Flanagan, P Cavanagh, and O E Favreau. Independent orientation-selective mechanism for the cardinal directions of colour space. *Vision Research*, 30:769–778, 1990.
- [6] C A Párraga, T Troscianko, and D J Tolhurst. The human visual system is optimised for processing the spatial information in natural visual images. *Current Biology*, 10(1):35–38, Januari 2000.
- [7] Stéphane J M Rainville and Frederick A A Kingdom. Spatial-scale contribution to the detection of mirror symmetry in fractal noise. *J. Opt. Soc. Am. A*, 16(9):2112–2123, September 1999.
- [8] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley. Statistical color processing. Submitted.
- [9] Erik Reinhard, Peter Shirley, and Tom Troscianko. Natural image statistics for computer graphics. Submitted.
- [10] D L Ruderman, T W Cronin, and C-C Chiao. Statistics of cone responses to natural images: Implications for visual coding. *J. Opt. Soc. Am. A*, 15(8):2036–2045, 1998.

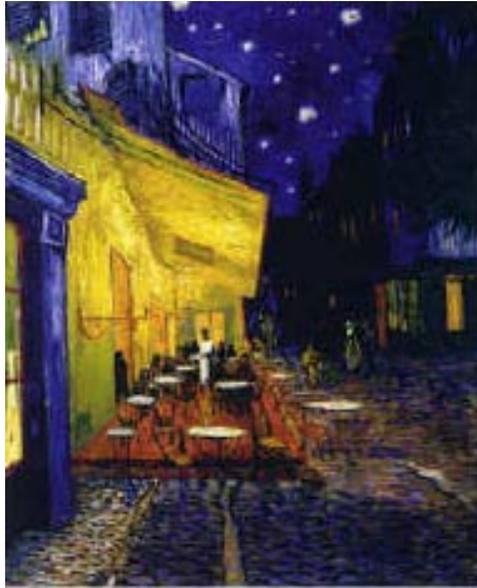


Figure 1: *The atmosphere of Vincent van Gogh's "Cafe Terrace on the Place du Forum, Arles, at Night" (Arles, September 1888, Oil on canvas) applied to a photograph of Lednice castle near Brno. In this example, the colors of the sky in both images, the yellows of the cafe and the castle and the browns of the tables at the cafe and the people at the castle were matched separately using small swatches [8].*