Eurographics/ IEEE-VGTC Symposium on Visualization (2006)
Thomas Ertl, Ken Joy, and Beatriz Santos (Editors)

Natural Visualizations

Steve Haroz* and Kwan-Liu Ma*

Department of Computer Science, University of California at Davis

Abstract

This paper demonstrates the prevalence of a shared characteristic between visualizations and images of nature. We have analyzed visualization competitions and user studies of visualizations and found that the more preferred, better performing visualizations exhibit more natural characteristics. Due to our brain being wired to perceive natural images [SO01], testing a visualization for properties similar to those of natural images can help show how well our brain is capable of absorbing the data. In turn, a metric that finds a visualization's similarity to a natural image may help determine the effectiveness of that visualization. We have found that the results of comparing the sizes and distribution of the objects in a visualization with those of natural standards strongly correlate to one's preference of that visualization.

Categories and Subject Descriptors (according to ACM CCS): H.1.2 [Models and Principles]: User/Machine Systems, I.4.8 [Image Processing and Computer Vision]: Scene Analysis, H.5.2 [Information Interfaces and Presentation]: User Interfaces

1. Introduction

Many have tried to define or explain what properties make for an effective visualization [Sea95][Shn96][Tuf90]. Elements such as aspect ratio, density, color usage, and typeface variation have been examined separately and in combination with varying degrees of success, but no one has given a concrete reason as to why these aspects of a visualization are important [Sea95][Bra97]. Even when measuring proximity and clustering, little evidence is given to explain why these features are helpful [Bra97]. Layout is also a commonly scrutinized aspect of visualizations, but the comparison of layout techniques frequently requires exhaustive user studies involving eye-tracking [Sea95], which is expensive and time-consuming. Though these user studies can give concrete proof that one implementation better conveys information than does another, they are incapable of explaining the underlying reasons behind why some visualizations provide more insight then do others. We therefore aim to explain a more fundamental property that correlates strongly with preferred visualizations.

Visualization is ultimately a field of translation. The goal of every visualization is to convert raw binary data from a machine-readable encoding to a neural encoding understandable to the mind. In order to be successfully converted, this data must pass through the visual system. Therefore, tuning visualizations to the visual system should help in their effective translation or perception. In this paper, we demonstrate that theories known to neuroscientists who study vision can be effectively utilized to find patterns in the effectiveness and preference of visualizations. As a first step, we replicate computational neuroscience experiments and extend them to show that their image analysis techniques can be applied to images with properties similar to those of visualizations. After inspecting measurements of these images, we apply those measurements to actual visualizations to find a correlative pattern. We then examine how this pattern's categorization of visualizations is similar to that of the InfoVis contest results [FGP04][GCD*05] and a timed user study . These techniques could potentially contribute to the advancement and expansion of our understanding of visualization perception and may help influence future visualization developments and applications rooted in neuroscience foundations.

* {sharoz | klma} @ucdavis.edu

2. Background

For over two decades now, neuroscientists have studied images of nature to better understand how our visual system perceives them [Fie87][Fie93]. They have found that these natural images contain distinctive statistical regularities that random images lack and that our brains are wired to perceive this natural stimulus [OF96][SO01]. One of the fundamental distinctions of natural images is their sizing and spacing which can be analyzed by observing the images' spatial frequencies [Fie87].

2.1. Spatial frequency

Spatial frequencies are similar to sound frequencies. Sound frequencies are a measurement of compression varied over time, whereas spatial frequencies are a measurement of intensity varied over distance [Fie87]. Since spatial frequencies can only measure a single intensity value, brightness is commonly used.

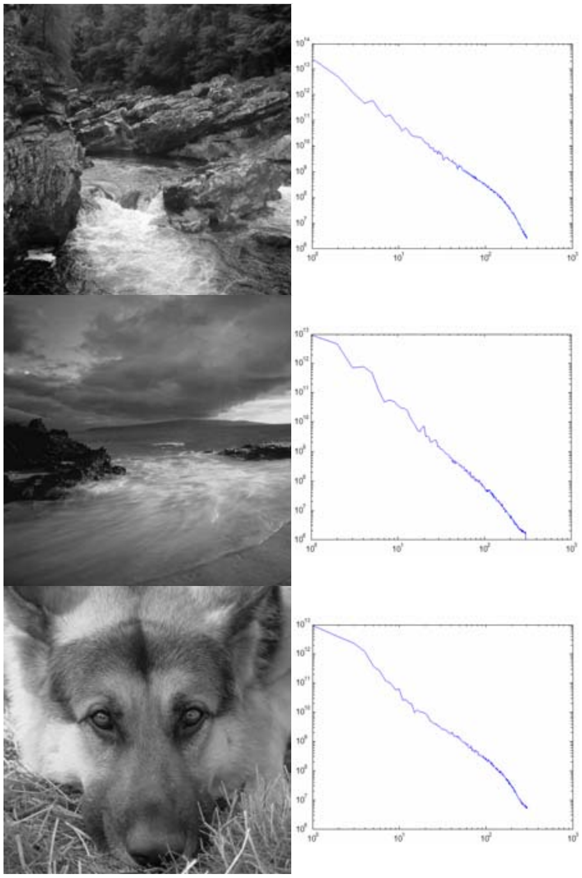

Figure 1 *Left: These are examples of natural images. Right: These plots correspond to the power spectra of the images on the left. The x-axis is the frequency on a log scale, and the y-axis is the amplitude which is also on a log scale.*

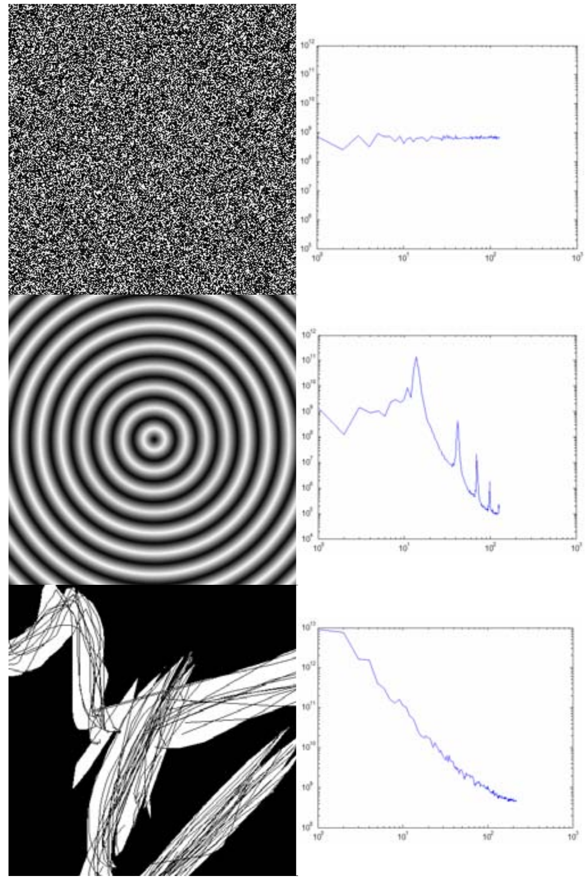

Figure 2 *Left: These are examples of unnatural images. The first is purely random noise. The second is a radial gradient repeated ten times. The third is just a scribbling. Right: These plots correspond to the power spectra of the images on the left. The x-axis is the frequency on a log scale, and the y-axis is the amplitude which is also on a log scale.*

2.2. Fourier transforms

One way to measure the spatial frequencies of a function is by using Fourier transforms. Essentially, sine and cosine waves of different amplitude and frequency are added together to form the intended function. These sine and cosine functions make up a Fourier series. For a two dimensional image, the Fourier transforms are performed over each line in the horizontal axis then over each line in the vertical axis or vice versa. In turn, to find the two-dimensional Fourier transform of an n-by-n image, one must find 2n one-dimensional Fourier transforms.

2.3. Natural and unnatural images

A natural image is any picture of nature [Fie87]. Pictures of a forest scene, a mountain, or a dog would be considered natural images. Three examples can be seen in Figure 1. This class of images constitutes an infinitely small fraction of all possible images [RST01], yet our visual system is precisely tuned to perceive them rather than some larger range of image types. To measure the spatial frequency distribution of these images, one begins by computing the Fourier transform. The rotational average of the two dimensional result yields a more manageable, one dimensional series also known as a power spectrum [Wei]. When the amplitude of this spectrum is plotted on a log-log scale as a function of frequency, the spatial frequency distribution can be visualized.

On the right of Figure 1, plots of the power spectra from the natural images are shown. These plots have nearly straight lines with slopes of approximately -2, which corresponds to an f^{-2} trend. The consistency between the plots is not trivial, as these images appear quite dissimilar. Unnatural images have very different power spectra. Figure 2 contains three unnatural images, and on the right their corresponding spatial frequency plots are shown. The distinctness of natural images becomes more evident in these plots, as the unnatural images do not show the f^{-2} trend. Many other papers [Fie87] have discussed exhaustive studies on large numbers of natural and unnatural images,

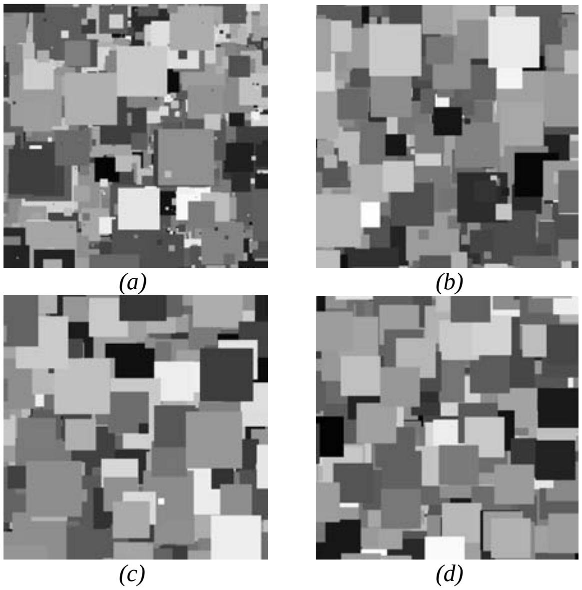

Figure 3 *These images were generated by randomly placing squares with random Gaussian-fit grayscale values. The size distributions are power (a), exponential (b), linear (c), and constant (d).*

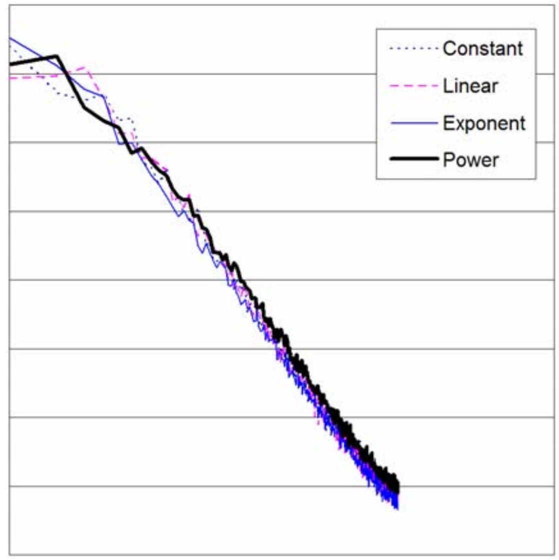

Figure 4 *The power spectra of images generated by occluding squares. Notice that the power function stays slightly above the others.*

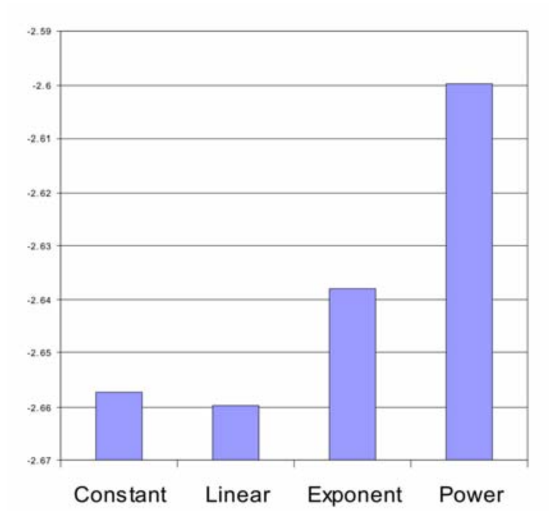

Figure 5 *The slopes of the linear trends fit to the power spectra in Figure 4.*

and all have repeatedly found the same results. The f^{-2} trend sets apart natural images.

Not all natural images will have a spatial frequency distribution of exactly f^{-2}. Image scaling as well as window size and shape can be used to explain why certain images deviate from the trend, yet whole groups of images such as sky scenes or images at a large scale have also been found to have unusual power spectra. Michael S. Langer showed that these anomalous groups make up only a small subset of all natural images and tend to cancel out each other when large numbers of images are measured [Lan00]. He also noted that these atypical natural images do not contain structure that is rich or interesting. For visualizations, images without interesting structure would be incapable of providing useful insight, and the corresponding visualization or its scale would be undesirable.

2.4. Size distribution

Daniel Ruderman investigated the cause of natural image characteristics being independent of calibration and visual environment [Rud97]. Many assume that natural image traits "result from edges, each with a power spectrum of $1/k^2$" [Rud97]. Ruderman disproves this belief using contradiction. Instead, he shows that statistically independent 'objects' are the cause. These objects are place randomly on top of each other and have a size distribution that follows a power function. The occlusion resulting from this collage of specifically sized objects produces the observed f^{-2} power spectrum.

Ruderman demonstrated this premise by producing images made up of a collage of squares. These squares are positioned randomly and given a random grayscale value that follows a Gaussian distribution. The size distribution of the squares is given by a power function or an exponential function. The images generated using the power function demonstrate a more natural power spectrum than those generated by the exponential function.

We have replicated Ruderman's natural image generation test and extended it to include linear and constant distributions along with the power and exponential distributions demonstrated in his experiment. Samples of the generated images can be seen in Figure 3. Their corresponding power spectra can be seen in Figure 4, and a linear fit was applied to the plots to obtain the slopes of the trends in Figure 5. The trends show that the power function size distribution has a slope (-2.5) that is closer to -2 than

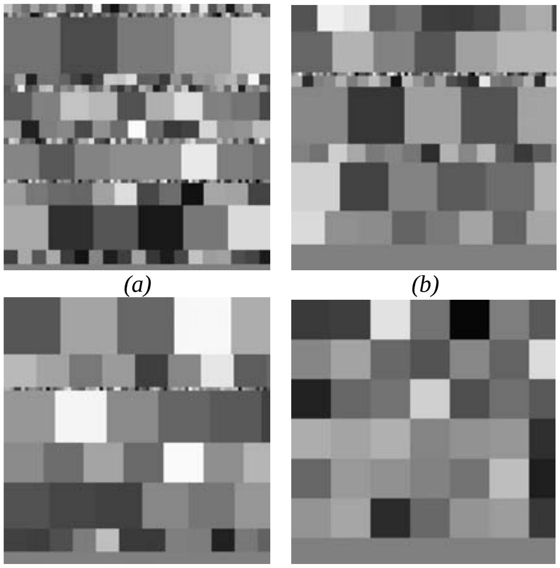

Figure 6 *These Treemap [JS91] resembling images were generated by creating rows of non-overlapping squares with random greyscale values. The size distributions are power (a), exponential (b), linear(c), and constant (d).*

those of the others (-2.6). In other words, the power function more closely emulates natural images.

3. Extending Existing Theories

In visualization, efforts are generally made to avoid occlusion. According to Ruderman, visualizations would therefore be incapable of having the characteristics of natural images. The implication is that we would have difficulty perceiving visualizations without occlusion because their power spectra would be different than that of our visual system. Accordingly, determining if the same size distribution rules apply to images without occlusion is crucial to knowing if these visualizations can exhibit natural traits.

3.1. Images without occlusion

We proceeded to revise our image generation program to prevent occlusion. Doing so turned out to be more difficult than creating the original program, as placement cannot simply be random. The squares need to be placed in such a way that no square occludes any other. The resulting image should form an artificial visualization with measurable characteristics. For a given size image, n rows were created, where n is dependant on the size distribution formula. This sizing applies to area, not width, making n reliant on an unexpectedly complex formula:

$$h \geq \sum_{i=0}^{n} \sqrt{f\left(f^{-1}(\min) + \frac{i \cdot \left(f^{-1}(\max) - f^{-1}(\min) \right)}{n} \right)}$$

In this equation, f is the size distribution function, h is the height of the image, min is the area of the smallest square, and max is the area of the largest square. The images of

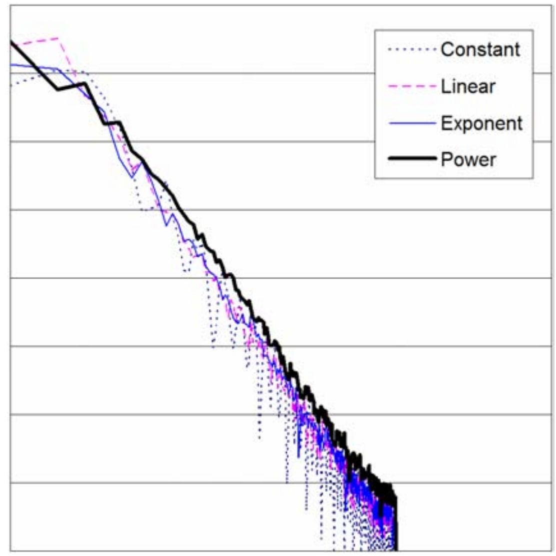

Figure 7 *The power spectra of images generated by rows of non-overlapping squares*

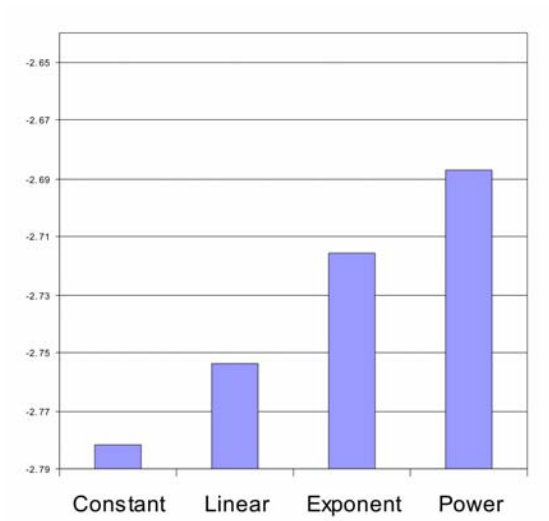

Figure 8 *The slope of the linear trends fit to the power spectra of the non-overlapping squares.*

non-overlapping squares can be seen in Figure 6. The power spectra and their trends (as seen in Figure 7 and Figure 8) are similar to those of the images with occlusion. The power function's slope is closest to -2 followed by the exponential, linear, and constant functions respectively. Clearly, these images show that images without overlapping objects can have natural characteristics.

The trends of these images have slopes that are all within a small range of less than 5%. Such a small range leaves too much room for these findings to result from a mere statistical anomaly. Many seemingly insignificant factors may have been the cause of one type of scaling seeming more natural than did another. We therefore addressed several potentially confounding factors in an attempt to deviate from the original outcome.

- Unusual run: This particular run of the image generating program could have resulted in a fluke, so the program was run several times.
- Row order: Each run had random row placement.
- Orientation: The rotational average of the two-dimensional Fourier transform treats all orientations equally.
- Extra space: The size of the image was set to precisely fit one scaling function for each run, thereby eliminating the black bar for that particular image.
- Image size: Image heights varying from 500 to 1200 pixels were used.
- Shape: Circles and randomized shapes were also used.

Despite the variation of all these factors, the power spectrum trends were not significantly affected. The order of the slopes always remained the same - power, exponential, linear, and then constant.

All of these slopes are well within the range of what could be considered a natural image [Rud97]. However, due to their nearly identical values, a means of differentiation besides slope needs to be used in tandem.

3.2. Average deviation

Natural images have power spectra that adhere very closely to a straight line, while the power spectra of unnatural images are more likely to have many spikes and steep dips. To determine how closely a spectrum follows its trend, its the average deviation was calculated. The average deviation is found by averaging the absolute value of each point's difference between the actual power spectrum and its corresponding trend. This technique will help determine a plot's linearity. The average deviations for two natural images as well as those for the non-overlapping squares can be seen in Figure 9. The natural images are similar to the power and exponential distributions not the linear or constant distributions.

4. Measuring Natural Visualizations

Neuroscientists have shown that our brains are wired to perceive natural images [Fie87][Fie93][KL03][OF96][RST01][SO01]. The distribution of photo receptors in the retina follows an f^{-2} pattern as does the distribution of ganglion cell receptors immediately behind the retina. This pattern is persistent throughout the visual cortex. Therefore, visualizations that are most like natural images should be the most cognitional. In order to determine the extent of a visualization's natural characteristics, we propose measuring the slope of its power spectrum on a loglog scale as well as finding the deviation from a linear-fit trend. A visualization with natural characteristics should have a slope near -2 with a minimal average deviation. The slope is the most important factor because the deviation of a spectrum with a slope far from -2 is unimportant. We have shown that this metric produces predictable, reproducible results for artificial visualizations, so the following examples will demonstrate that natural characteristics

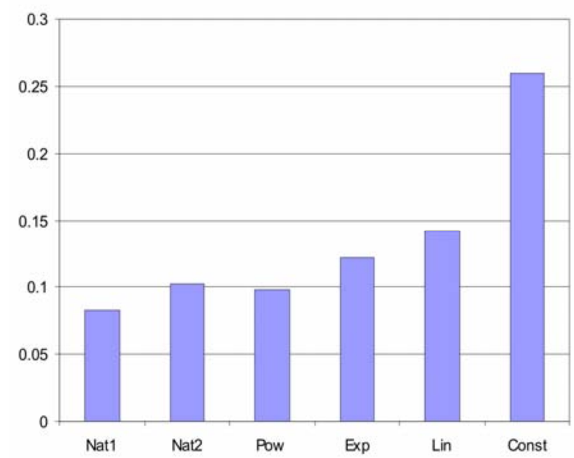

Figure 9 *The average deviation from the linear-fit trends of the power spectra from two natural images as well as the images from Figure 6.*

correlate closely with the preference and performance of actual visualizations.

4.1. Testing competition results

For one test, we looked at the InfoVis 2004 competition results [FGP04]. These visualizations all used the same dataset, which makes the comparison fairly objective. We analyzed an image for each of the first and second place winners (samples of which can be seen in Figure 10), and the results were better than even we expected. An image of each of the visualizations was taken from the Information Visualization Benchmarks Repository (http://www.cs.umd.edu/hcil/InfovisRepository/contest-2004/). The images were then converted to grayscale and truncated to be square in size, as shown in . Efforts were made to only truncate blank space around the sides. We then performed a spectral analysis of each of the images.

To compare the results, we found the distance of each slope from -2. Incredibly, all of the first place winners had slopes within .4 of -2, while the second place winners were mostly outside of that range. Figure 11 shows a clearly discernable distinction between the first and second place winners. The judges must have an unconscious preference for visualizations that are similar to natural images, as their evaluation accurately reflected the visualizations' natural measurements. The spatial frequencies have actually

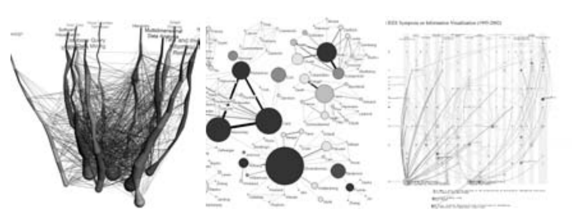

Figure 10 *Here are select thumbnails from the InfoVis 2004 contest. The left two visualizations received 1st place prizes. The right visualization received a 2nd place prize.*

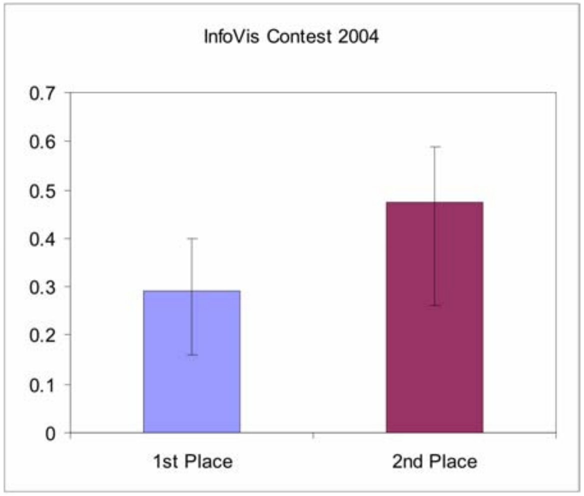

Figure 11 *This graph shows our analysis results from the 2004 Infovis contest. We measured the distance of the linear-fit trend from -2. We then took the average of those who came in first place and those who came in second place. The error bars show the total range for each rank.*

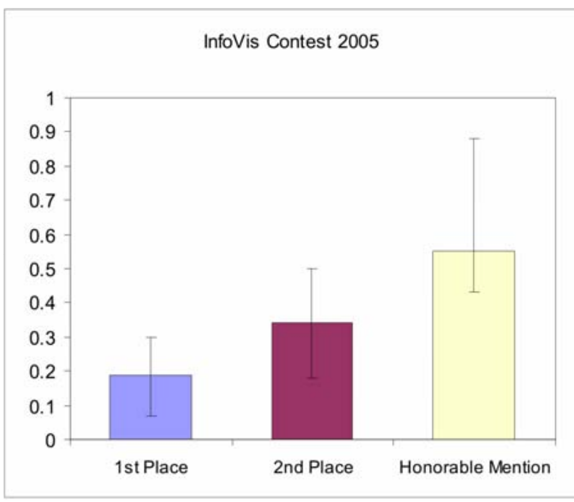

Figure 12 *In the results from the 2005 Infovis contest, a pattern between the first place, second place, and honorable mention averages is clearly prevalent. The error bars show the total range for each rank. All of the entrants' images can be found at http: //ivpr.cs.uml.edu/infovis05*

quantified the judges' preferences.

To show that this was not a fluke, we tested the results of the next year's competition [GCD*05] as well, and the results were even clearer due to the inclusion of an 'honorable mention' category (see Figure 12).

4.2. Testing user performance

Beyond simply predicting competition results, we can also show that a visualization's natural characteristics correlate with users' performance. To test the extent of this property, we analyzed the results of a user study of hierarchical visualizations .

The experiment's purpose was to time a user's ability to find structural similarities and differences within hierarchical data. The experiment looked at three interfaces that implement different hierarchical visualizations. Each user was assigned one interface which they could use to answer to six questions. The experimenters then recorded the time taken to answer each question. An important aspect of this experiment is that, as with the InfoVis competitions, everyone in the experiment used the same data. This aspect helps reduce unforeseen influences on the results.

After examining the average time taken by the users of each interface for each question, the experimenters found that Windows Explorer had the worst times, and the treemap and RINGS interfaces generally had similar times. We then used the information to correlate the time taken with the power spectra of screenshots. In this case, we not only looked for the relative order of the power spectra's slopes, but we also looked at the ratio of those slopes compared to the ratio of the times.

The results of the initial images analysis were predictable; Explorer has a power spectrum that is far from the natural standard, whereas the other two have more natural traits. We then plotted the correlation between the average response time and the distance of the power spectrum slope from -2 (Figure 13). Half of the questions had correlations with absolute values of around 90%. After examining the differences between the strongly and weakly correlated questions, we found that the strongly correlated questions required users to look for data that they were unlikely to have seen in a previous question. In other words, most of the weakly correlated questions were follow-up questions. For example, question one asked to

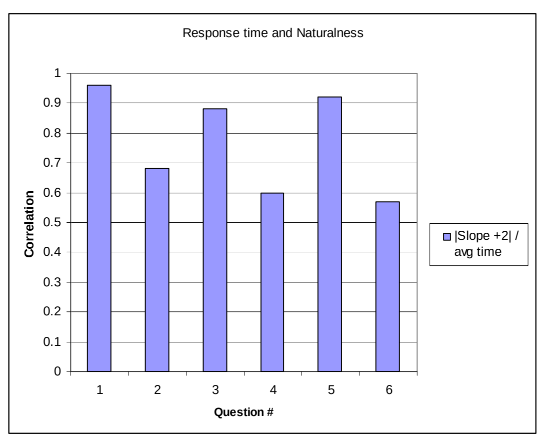

Figure 13 *The absolute value of the correlation between the user response time for each interface and the respective naturalness of a screenshot of that image*

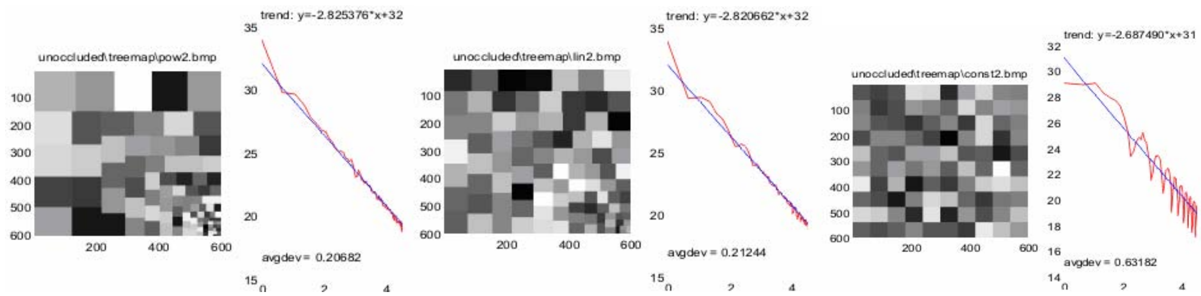

Figure 14 *These are treemaps generated using power (top), linear (middle), and constant (bottom) size distributions. Their corresponding power spectra are next to them. Notice the low average deviation for the power function.*

find similar folders and question two asked to find "very" similar folders. A probable cause of the distinction in correlation is that the influence of naturalness on an interface is strongest while the user is unfamiliar and still learning; whereas it is less influential once the user has memorized some of the information. Our general finding is that a user's ability to extract *new* information has a strong correlation with the power spectrum of the interface used.

5. Improving visualizations

We have demonstrated a technique for analyzing images to retrieve information about their natural characteristics. The resemblance of a visualization to a natural image has been shown to strongly relate with both competition results and user performance. A problem with using image analysis, however, is that it only gives the extent of naturalness, not the cause. A question thereby arises of whether this knowledge can be used as a guideline for the design of visualizations.

A likely culprit for an unnatural visualization is the underlying data. Fitting a visualization's data to a power function can make resulting visualizations more natural, and treemaps can be used to test that theory. Treemaps [BS02] are a variation of the unoccluding images from Section 3.1. When applying the color and size distribution methods from that section to treemaps, similar results are produced. Constant, linear, exponential, and power distributions produce power spectra near f^{-2}, and their average deviations decrease respectively. These images and their results can be seen in Figure 14. The implications of these findings mean that the size distribution of a visualization's data can help determine the visualization's natural qualities. Consequentially, distorting data to fit a power distribution may improve the resulting visualization.

6. Utility or art

Questioning what is being measured by a visualization's closeness to a natural image is essential. We have not established any causality between naturalness and visualization quality. They may both be caused by some other factor. A likely candidate for influencing naturalness and our preference of one visualization over another is aesthetic appeal. Art is usually found to have statistics similar to those of a natural image [RST01][Sch92]. This observation is made evident by our fascination with fractals [PS88] which have repeating shapes and frequently have a power spectrum just below f^{-2} [Sch92]. A possible implication is that visualizations are preferred due to their appearance rather than their ability to enhance cognition. On the other hand, these qualities may not necessarily be mutually exclusive, as the distinction between art and utility may not necessarily exist. Obviously this question leaves much room for research.

7. Limitations and future work

One should note that this measurement technique does have limitations. The metric only gauges the sizing and spacing in a visualization. It does not directly evaluate other aspects such as color or font variety. Moreover, it does not even provide appreciably helpful feedback as to the cause of a visualization's unnatural evaluation. The only appreciable feedback for improvement given by the power spectrum is a rough estimation of the underrepresented and overrepresented frequencies. To broaden the encompassment of the measurements, future work could be done to observe the power spectrum of each of the colors in a visualization or to study the influence of text sizing and distribution.

Another limitation is the inability to use power spectra as an all encompassing, exclusive means of visualization measurement. A correlation is not causality. Although natural characteristics show a non-random prevalence in preferred visualizations, some visualizations with very unnatural appearances are useful for certain tasks. A spreadsheet, for example, is probably the most effective visualization if the row and column of the desired information is known. Nevertheless, the spreadsheet has a highly unnatural power spectrum due to its regularity. The natural properties needed for an effective visualization might be determined by the task being performed. An interesting future study could compare the naturalness and performance of visualizations with the type of task being performed. If large numbers of similar visualizations were collected, one could also study the how the Fourier transform is affected by the type of visualization and its contents. Fourier transforms have been shown to be capable of categorizing natural images and determining their contents [TO03]. This approach might be also applicable to visualizations for the purpose of content recognition or automated feature extraction. We should also note that spatial frequency is not the only property of a natural image, and testing for other natural properties might provide more insight.

Further research of this metric's implications may help in understanding how insight is obtained from data. Since this measurement shows how well our visual system is adapted to perceiving an image, we may have a better ability to focus and extract information when that information is presented naturally. Past research has studied perception of natural stimulus and found that it has a very sparse representation in the brain [OF96]. If the interface for data is efficiently processed by the brain, more attention may be available to focus on the data that the interface is trying to convey. By naturalizing a visualization, we may be streamlining the process of perception. With more research, a more concrete neurological foundation for establishing visualization utility might result.

8. Conclusions

We have shown a strong correlation between the natural characteristics of a visualization and its preference and performance. This underlying principle has been built up from testing abstract nonrepresentational images and has been combined with work already done in the fields of neuroscience and computer vision. When applied to visualization, a metric based on natural image statistics has been shown to be consistent with assumed preference and competition results regardless of the use of occlusion. This connection has been demonstrated both theoretically and practically, and it shows that we can take advantage of our brain's enhanced receptivity to natural images. Hopefully, we can use this pattern for preferred visualizations to design future visualizations that, by their very nature, can better appeal to the human visual system.

Acknowledgements

We thank Bruno Olshausen for providing assistance and directing us to helpful sources. We would also like to thank Yue Wang for providing us with data from her user studies.

This work has been sponsored in part by the U.S. National Science Foundation under contracts CCF 9983641, CCF 0222991, OCI 0325934, and IIS 0552334, and the U.S. Department of Energy under Lawrence Livermore National Laboratory Agreement No. B537770, No. B548210 and No. B550194.

Reference

[Bra97] Brath R.: Metrics for effective information visualization. *Proceedings of Information Visualization* 1997 Symposium, 1997.

[BS02] Bederson B., Shneiderman B.: Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies. *ACM Trans-actions on Graphics*, 21(4), October 2002.

[FGP04] Fekete J. D., Grinstein G., Plaisant C.: IEEE InfoVis 2004 Contest, the history of InfoVis, (2004) http://www.cs.umd.edu/hcil/iv04contest

[Fie87] Field D. J.: Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America* A, 4(12):2379–2394, December 1987.

[Fie93] Field D. J.: Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems. M. Farge, J. Hunt, and J. Vassilicos, editors. *Wavelets, Fractals and Fourier transforms: New Development and New Applications*, pages 151–193. Oxford University Press, 1993.

[GCD*05] Grinstein G., Cvek U., Derthick M., Trutschl M.: IEEE InfoVis 2005 Contest, Technology Data in the US, http://ivpr.cs.uml.edu/infovis05

[JS91] Johnson B., Shneiderman B.: Tree-maps: A space-filling approach to the visualization of hierarchical information structures. *Proceedings of IEEE Visualization* 1991 Conference, pages 284–291, 1991.

[KL03] Karklin Y., Lewicki M.: Learning higher order structures in natural images. *Network: Computation in Neural Systems*, 14:483–499, 2003.

[Lan00] Langer M. S.: Large-scale failures of f^{-a} scaling in natural image spectra. *Journal of the Optical Society of America*, 17:28-33, 2000.

[OF96] Olshausen B., Field D. J.: Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7:333–339, 1996.

[PS88] Peitgen H., Saupe D. editors. The Science of Fractal Images. Springer-Verlag, New York, 1988.

[RST01] Reinhard E., Shirley P., Troscianko T.: Natural image statistics for computer graphics. Technical Report UUCS-01-002, School of Computing, University of Utah, March 2001.

[Rud97] Ruderman D.: Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398, 1997.

[Sch92] Schroeder M.: Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise. W.H. Freeman & Company, 1992.

[Sea95] Sears A.: Aide: A step toward metric-based development tools. *ACM Symposium on User Interface Software and Technology*, pages 101–110, 1995.

[Shn96] Shneiderman B.: The eyes have it: a task by data type taxonomy for information visualization. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[SO01] Simoncelli E., Olshausen B.: Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24:1193–1216, 2001.

[TO03] Torralba A., Oliva A.: Statistics of natural image categories. *Network: Computation in Neural Systems*, 14:391–412, 2003.

[Tuf90] Tufte E.: Envisioning Information. Graphics Press, Cheshire, CT, 1990.

[Wei] Weisstein E. W.: Power spectrum. MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/PowerSpectrum.html

[WTM05] Wang Y., Teoh S. T., Ma K. L.: Can Tree Visualization Really Discover Knowledge?. Tech. Report No. 2005-31, Department of Computer Science, University of California at Davis, December 2005