Spatial Autocorrelation-Based Information Visualization Evaluation

Joseph A. Cottam and Andrew Lumsdaine
Center for Research in Extreme Scale Technologies (CREST)
Indiana University
Bloomington, IN
jcottam, lums@indiana.edu

ABSTRACT

A data set can be represented in any number of ways. For example, hierarchical data can be presented as a radial node-link diagram, dendrogram, force-directed layout, or tree map. Alternatively, point-observations can be shown with scatterplots, parallel coordinates, or bar charts. Each technique has different capabilities for representing relationships. These capabilities are further modified by projection and presentation decisions within the technique category. Evaluating the many options is an essential task in visualization development. Currently, evaluation is largely based on heuristics, prior experience, and indefinable aesthetic considerations. This paper presents initial work towards an evaluation technique based in spatial autocorrelation. We find that spatial autocorrelation can be used to construct a separator between visualizations and other image types. Furthermore, this can be done with parameters amenable to interactive use and in a fashion that does not need to take plot schema characteristics as parameters.

Keywords

Automated Evaluation, Spatial Autocorrelation

1. INTRODUCTION

Representing information requires making decisions. Not all decisions are value-neutral. For example, significantly different trends can appear when a variable is log-transformed before projection. When working with complex data, there may literally be thousands of such potentially significant decisions to make. It is infeasible for visualization designers to explore the entire design space for all visualizations they create. Instead, they rely on heuristics and domain practices to guide them to acceptable representations.

Relationship evaluation is one of the principle applications of statistical analysis. Many statistical models are, at their core, a means of capturing and discussing relationships found in data. Spatial autocorrelation is one way of statistically evaluating a spatially-distributed relationship. Visualizations are 2D projections of data sets. In many ways, a useful visualization is one that visually presents relationships that are actually present in the underlying data. This paper presents an initial investigation of spatial autocorrelation as a means of evaluating visualizations.

In the abstract, spatial autocorrelation attempts to measure the amount of "pattern" found in a 2D space by comparing a given region to its neighbors and summarizing the amount of variation discovered. We hypothesize that visualizations will exhibit a different characteristic amount of pattern than other types of visual communication. This paper directly tests that hypothesis. Furthermore, this paper provides evidence that certain levels of patterning are associated with "better" visualizations. This evidence leads us to believe that spatial autocorrelation is a useful metric for evaluating visualizations, though it is probably not a complete solution in its own right.

2. RELATED WORK

Tukey and Tukey coined the term Scagnostics (*Sca*tterplot Dia*gnostics*) for automatically evaluating the relationship between variable pairs based on the image-space characteristics of a scatterplot [22]. Wilkinson, Anand, and Grossman [24] provide details on Tukey and Tukey's original work and extend it with graph-theoretic measures. This original work considers measures such as convex hull length plot density, convex hull perimeter length, convex hull area, and the relationship between iso-level contours. The extended work considers the individual points of a scatterplot as entities in a graph and evaluates the plots based on measures related to non-convex hulls and minimum spanning trees. These measures are used to characterize scatter plots for features (such as "stringiness") that may be useful for interpretation or detecting errors. Scagnostics can be used to reduce the number of variable-pairs that an analyst needs to consider by pruning to only those that have interesting features.

Pargnostics (*Par*allel coordinates Dia*gnostics*) was introduced by Dasgupta and Kosara [8] to evaluate parallel coordinates plots based on image-space metrics. Pargnostics is built around histograms of the start-position, slope, line crossings, and over-plotting. It also includes entropy and mutual-information measures.

Both Pargnostics and Scagnostics are concerned with specific types of visualizations. More general measures that

BELIV 2012, Seattle, WA, USA
Copyright 2012 ACM 978-1-4503-1791-7...$10.00

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

Figure 1: Moran's I – N is the number of regions; i and j range from 0 to N; X is the dependent variable; $\bar{X}$ is the mean of X; w_{ij} is an element in a weights matrix.

$$C = \frac{(N-1)\sum_i \sum_j w_{ij}(X_i - X_j)^2}{2W \sum_i (X_i - \bar{X})^2}$$

Figure 2: Geary's C – W is the sum of all w_{ij}; other values the same as in Figure 1

are image-type agnostic have also been developed. Pixnostics [19] is one such framework. In image-space, Pixnostics filters possible projections by examining the histogram of the distribution of "gray values" in the image. The distribution of values is used to compute entropy and spread of the image, and prioritize the presentation of high-information images. This process is also repeated in subregions to identify interesting regions.

Information theory factors heavily into Pargnostics and Scagnostics. Several aspects of information theory were also investigated by Chen and Jaenicke [7]. They consider entropy, joint entropy, conditional entropy, and mutual information to construct measures on images and source data. Specifically, they quantify the information loss and the relationship between different views (under conditions such as zooming, panning, and filtering). They find several effective measures, but indicate some pitfalls as well, since the measures they consider do not take perceptual characteristics or fine-grained co-locality into account. As with Pixnostics, the measures of Chen and Jaenicke can be used in subregions as well as whole images.

3. SPATIAL AUTOCORRELATION

In general, measures of spatial autocorrelation capture the amount of "pattern" found in 2D space by comparing a given region to its neighbors and summarizing the amount of variation discovered. This results in a single measure describing the entire 2D data set. Intuitively, an image of pure static has no spatial autocorrelation, and a solid field has perfect spatial autocorrelation. Furthermore, an ordered arrangement of regions has higher spatial autocorrelation than a random rearrangement of the same regions.

There are several different measures of spatial autocorrelation. Moran's I [17] is the oldest and most common. The calculation for Moran's I is shown in Figure 1 and gives values between -1 (anti-correlated) and 1 (fully correlated). We also consider Geary's C [9], which is inversely related to Moran's I, but more sensitive to local correlation. Geary's C is computed per Figure 2 and ranges from 0 (fully correlated) to 2 (anti-correlated).

All measures of spatial autocorrelation have a notion of "neighborhood." Neighborhoods establish the regions that are to be compared to each other. The spatial autocorrelation measures establish neighborhoods by a "weights matrix," where the value at position (i, j) in the matrix w determines how to weight the value in region j when regarding region i. The matrix is normalized for calculations but abstractly may contain arbitrary positive values. The practical implication of using a weights matrix is that neighborhoods may be arbitrarily constructed. Contiguity, distance, and reflexivity may or may not be present. The weights matrix and its implicit description of neighborhoods have a large impact on the spatial autocorrelation measures. For example, a black/white checkerboard pattern has perfect spatial autocorrelation if considering the diagonals, perfect anti-correlation if considering the cardinal directions, and neutral spatial autocorrelation if considering all eight directions.

In contrast to spatial autocorrelation, non-spatial methods for evaluating 2D distributions often rely on histograms of the values. However, such methods cannot distinguish (for example) an image from a scrambling of its parts. To determine if a specific image has *significant* spatial structure, a reference distribution of values must be established. There are three main methods for establishing this distribution: assume a normal distribution, assume a uniform distribution, and generate a distribution though permutations of random scrambling. There is no clear consensus in spatial analysis literature on deciding which reference distribution to use.

Many measures have been adapted to a per-region evaluation as well as to the global measures. Per-region measurements are also called "Local Indicators of Spatial Association" (LISA) [2]. In LISA calculations, each region being considered is given a measure of spatial autocorrelation. Intuitively, a local measure indicates how much variation to expect between a specific point and the points of its neighborhood, scaled by the deviation of the current position from the expected value. This local measure is analogous to a derivative in calculus in interpretation, helping to identify inflection points, but its derivation is significantly different.

Our investigation principally used the PySAL library [18]. PySAL includes implementations of Moran's I, Geary's C, and a LISA adaptation of Moran's I. PySAL computes all of the major methods for determining significance. Though PySAL includes utilities for creating weight matrices, the regularity of the required matrices made a custom method more practical. GeoDa [3] was also used to double-check specific results and in LISA investigation.

3.1 Data Set & Preparation

We acquired visualizations prepared as examples for the D3 visualization tool [5]. This set was selected because it contains a variety of different visualizations, including different visualizations of the same data, and many standard visualizations. Visualizations that were primarily for artistic purposes were excluded, as were visualizations that *required* interaction to observe significant portions of the underlying data set. Visualizations that included interaction were collected in their default state for our primary analysis. In addition to the D3 example images, a number of patterned images with known correlation values and photographs were used as controls.

$$L = R * \frac{299}{1000} + G * \frac{587}{1000} + B * \frac{114}{1000}$$

Figure 3: The ITU-R 601-2 "Luma" transform [1].

Some visualizations with interaction were collected in multiple states, but primary analysis focused on the default states (see Sections 4.6 for one treatment of multiple states reached through interaction). Visualizations were acquired through screen-capture at 72dpi. When necessary, multiple screen captures were performed and combined to ensure the entire visualization was used in the analysis.

For all of our analysis, each pixel in an image was considered a region and all neighborhoods were established using Manhattan distance. All regions within the neighborhood were given equal weight.

To facilitate investigation, images were down-sampled using the Python imaging library (PIL) [1] to a variety of resolutions. Downsampling was performed by first converting the images to the RGBA color space (to prevent issues with indexed colors). Second, the PIL "Image.thumbnail" method was employed for downsampling, with the "ANTIALIAS" option set. In a down-sampled set, all images had the same maximum dimension, with the alternate dimensions proportionally scaled (rounded up to the nearest whole pixel).

Spatial autocorrelation is well-defined for continuous values. The PySAL library only handles linear continuous values. Since color belongs to a three dimensional space, it is not directly approachable in PySAL. Our investigation proceeded on grayscale images derived from the originals (similar to the reduction done in Pixnostics [19]) using the ITU-R 601-2 "Luma" transform (see Figure 3), provided through the PIL "Image.convert" method.

4. RESULTS

Our eventual goal is to develop a visualization effectiveness measure. This paper lays the groundwork for evaluating spatial autocorrelation. To these ends, this paper investigates the effects of several different data treatments on spatial autocorrelation in visualizations and reference images. Specifically, we investigated: (1) the impact of different reference distributions for determining significance; (2) the impact of different spatial autocorrelation measures; (3) the impact of image scaling; (4) the impact of neighborhood size; and (5) local autocorrelation. We also compared the characteristics of visualizations to our reference images. These investigations provide a baseline for future investigations and reduce the combinations and treatments that future investigations need to consider.

4.1 Measures and Distributions

As mentioned, there are several standard metrics for spatial autocorrelation and several ways of determining the significance (e.g., p-value) of a result. The three major distribution types are (1) normal, (2) uniform, and (3) simulated by permuting the presented values. We performed initial analysis on the full-size images using both 2 and 250 permutations for the simulated distributions, as well as the normal and

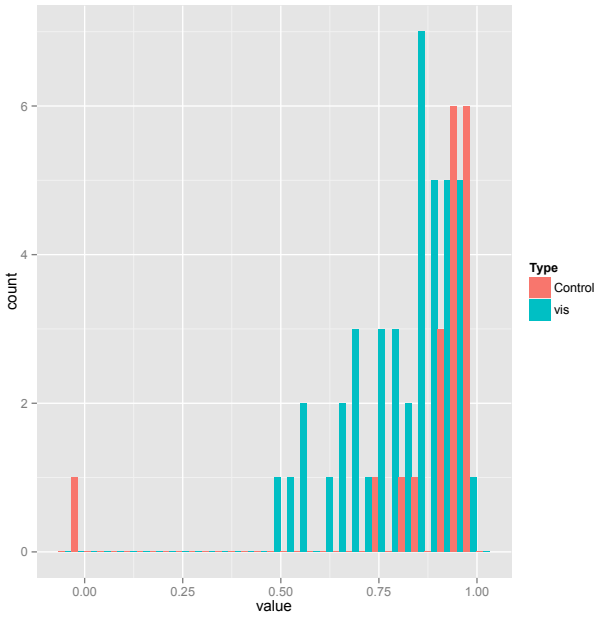


Figure 4: Distribution of Moran's I for full-sized images and normal distribution. Other distributions were not significantly different.

uniform distributions. Results for the normal distribution are shown in Figure 4. Our first finding is that there is no significant difference between the spatial autocorrelation statistics returned when using any of the reference distributions ($p<0.005$). Therefore, we conclude that the normal distribution is suitable for our analysis and all future p-values are with respect to it. All of the spatial autocorrelation statistics acquired for visualizations were significant ($p<0.001$). Some of the reference images did not have significant spatial autocorrelation values under some distributions (p values ranging from <0.001 to .73). Due to intermediate value formulations in PySAL, the spatial autocorrelation statistic for a solid image (e.g., fully correlated) resulted in "not-a-number." Therefore, the solid image does not appear in much of the analysis.

Moran's I and Geary's C are the two most popular measures of spatial autocorrelation. They are inversely related to each other; high values of Moran's I correspond to low values of Geary's C. However, they are sensitive to different features. Geary's C is known to be more sensitive to local deviations from spatial autocorrelation. Throughout our analysis, we compared the two statistics. However, the results returned from analysis using Geary's C did not significantly differ from those returned using Moran's I. Therefore, we conclude that either measure is effective, and we use Moran's I (the more common measure).

Having established the stability of the measures over different distributions, we investigated whether spatial autocorrelation can be used as a classifier. We constructed a generalized linear model in R with binomial variance distribution and logit link where type was dependent on Moran's I. We found that spatial autocorrelation can be used to distin-

guish visualizations from the reference images (a collection of images corresponding to extreme cases and photos).The parameter estimate was 1.5 and p=0.05, indicating that typical Moran's I can be used to (statistically) distinguish the visualizations from the reference images. Classifying visualizations as having a Moran's I between .4 and .92 correctly categorizes 70% of the images and 85% of the visualizations. From a practical standpoint, a model such as this can provide a "sanity check" for a visualization designer. Misclassification according to this model indicates that a visualization's design has deviated from standard practices. This is not the same as establishing the "goodness" of a visualization, but it is a step in that direction.

Informally, we observe that several known good general-purpose visualizations (such as scatterplots) are found in the upper-values of the observed range. Conversely, visualizations with issues that significantly restrict their utility (such as node-link diagrams, which perform more poorly than matrix-based representations as graph size increases [10,13]) in the lower region. This leads to a provision "goodness" rule that indicates that higher spatial autocorrelation is preferred to lower spatial autocorrelation, up to a limit.

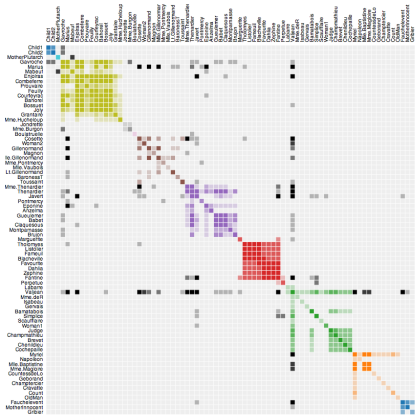
4.2 Image Size

Image size directly impacts both the apparent proximity of items and the statistics used to measure proximity (and therefore, correlations between proximate items). Generally speaking, larger images exhibited a wider spread in the correlation metrics. Standardizing the image size to 50 pixels reduces the range of values that visualization spatial autocorrelations returned. Conversely, the range of values for the control photographs expanded when image size was changed. Non-photographic reference images, such as static and the checkerboard, were constructed at their original resolutions to have per-pixel properties; the downsampling significantly altered those properties, so the down-sampled versions of those images and related results are invalid. Statistically, using smaller images *increased* the discriminating ability of the statistical model, giving a parameter estimate of 9.15 and p=0.05. Images scaled down to to just 250 pixels using the same methodology were statistically indistinguishable (using student's t-test) from the full-size image.

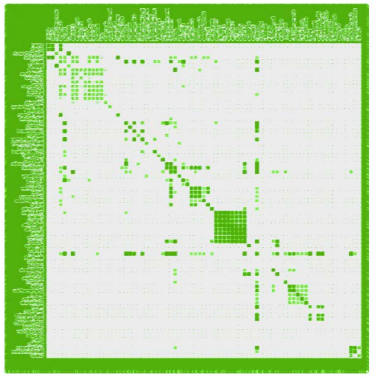
The analysis presented in Section 4.1 was repeated for the down-sampled images and the conclusions about reference distribution and correlation between measure values holds at all three investigated image sizes.

4.3 Neighborhood Size

Two neighborhood sizes were investigated, Manhattan distances of 1 and 5. Manhattan distance of 1 is equivalent to a "rook" neighborhood of size 1. Using a distance of 5, a square region with diagonals of size 10 and rotated 45-degrees is created. Working with thumbnails of size 50 and 250, no significant difference was detected in the spatial autocorrelation measures of the visualization images. Full-size images with neighborhood size 5 were not fully investigated due to computation expense, but the three images that completed processing in under 40 hours were within .0001 of the values reported for neighborhood of size 1. Therefore, we conclude that neighborhood size does not significantly impact spatial autocorrelation for the visualizations investigated.



(a) Original



(b) LISA Map

Figure 5: Local spatial autocorrelation (LISA) map. Darker green indicates higher LISA values.

4.4 Local Spatial Autocorrelation

Spatial autocorrelation can be applied to the single regions of an image as well as the image as a whole. When used in this way, global spatial autocorrelation measures are modified into LISA. Many global measures have a corresponding local measure. Significant LISA values indicate a change in the spatial autocorrelation measure surrounding a specific region. We hypothesized that local spatial autocorrelation may be used to guide investigation to particular regions of a visualization. We considered both the full-sized images and thumbnails with neighborhoods of size 1 and 5. The LISA measurements did identify locally-interesting features in some plot types (such as clusters and outliers, see Figure 5). However, these results were often obvious from the original plot.

4.5 Case Study: Directory Structure

Several of the D3 examples used the Flare directory structure as their underlying data set (results are summarized in Figure 6). File size and count were used in addition to the hierarchy of the directories. The representations used spanned the range of spatial autocorrelation measurements.

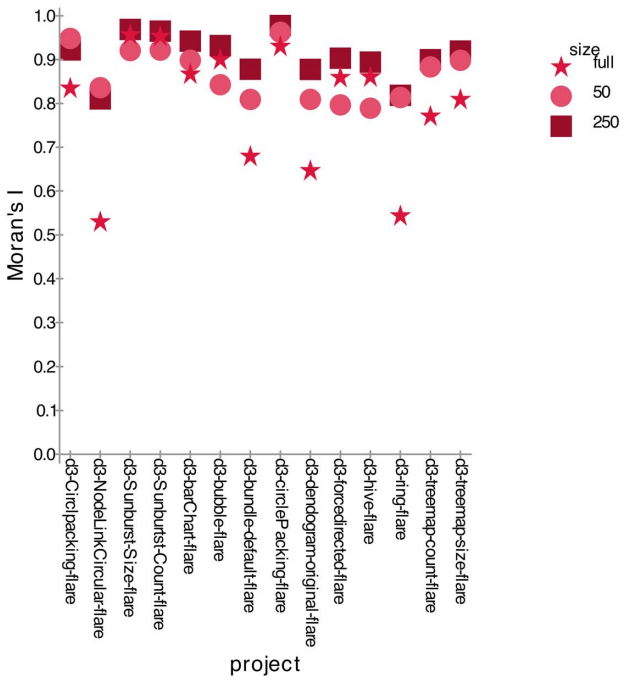
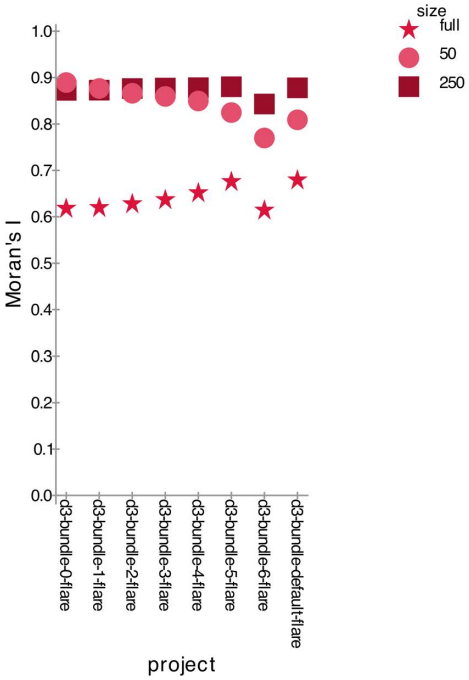
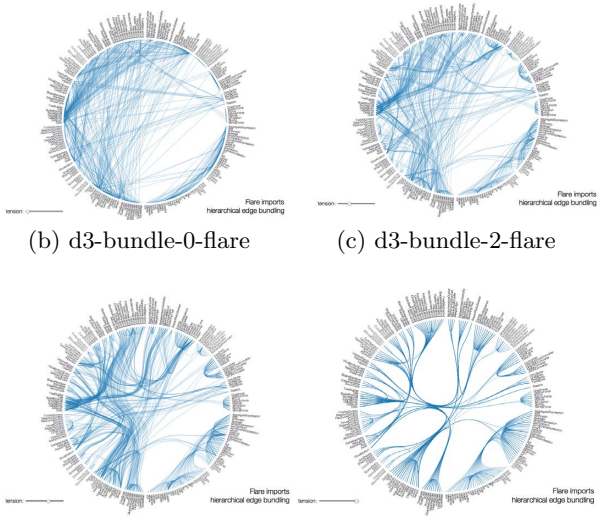


Figure 6: Comparison of different Flare directory structure representations.



(a) All Tensions



(b) d3-bundle-0-flare (c) d3-bundle-2-flare

(d) d3-bundle-4-flare (e) d3-bundle-6-flare

Figure 7: Chord diagrams with increasing bundling tension as the index number increases.

The most highly correlated were the space-filling techniques (tree-map and circle packing). When applicable, representing file-size vs. file-count did not significantly impact measures, indicating that the principal thing being represented in the visualization is the hierarchy. The least-correlated representations were the more node-link style visualizations. This includes loosely bundled chord diagrams, the concentric circle node-link tree and the force-directed layout.

4.6 Case Study: Chord diagrams

The Chord diagram with hierarchical edge bundling (see Figure 7) exhibits unusual behavior. Spatial autocorrelation metrics significantly change as the bundling tension changes (the example data set is the Flare directory tree). Curiously, image size significantly impacts the measure in this case (though neighborhood size did not). For full-sized images, higher tension leads to higher spatial autocorrelation. The converse is true for thumbnails. These two sequences approach each other (see Figure 7), with higher tension leading to closer agreement. Given earlier analysis, it appears higher tension is preferable as it leads to spatial autocorrelation values that are more distinct from those found in photographs.

4.7 Additional Visualizations

Visualizations were acquired from the ProtoVis and Prefuse galleries [6, 12]. These visualizations fell within the same range as the D3 examples. Notable exceptions are the "Data Mountain" (Moran's I=.93, see Figure 8) and other demos which relied heavily on photos. These had higher spatial autocorrelation values than the majority of other visualizations.

Using the rule that "good" visualizations have spatial autocorrelation near the center or high-end of the observed range, several classic visualizations are confirmed to be "good." This includes Minard's map of Napoleon's march (Moran's I = .72), Florence Nightingale's representation of causes of death in Crimean War (Moran's I = .91), and a scatterplot matrix of Anderson's Iris dataset (Moran's I=.73)

Earlier work on information-theory based metrics presented evidence for entropy as a basis of visualization evaluation. While known-effective visualizations were derived using entropy, so were highly ineffective visualizations (see Figure 9). Using the spatial autocorrelation measures, the encoding

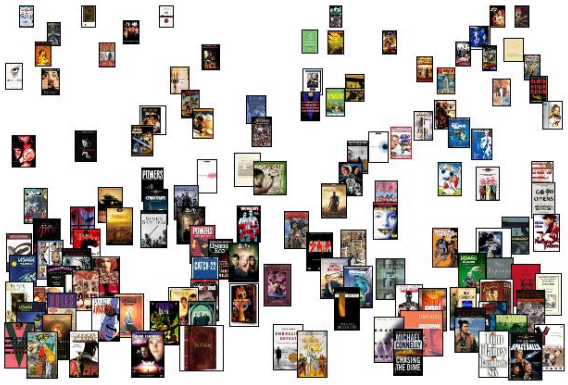


Figure 8: Data Mountain example from Prefuse [12].

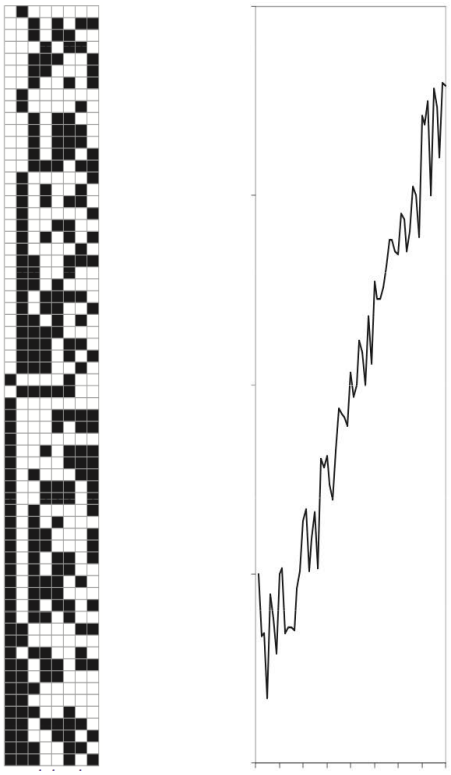


Figure 9: Sample time-series images from Chen and Jaenicke [7].

in Figure 9a falls outside of the provisional "good" range (Moran's I=.01), while the more interpretable visualization falls inside of the "good" range (Moran's I=.77, see Figure 9b). The poor encoding is noted by Chen and Jaenicke to "...require a huge perceptual and cognitive load to decode if not impossible" [7]. This result provides evidence that spatial autocorrelation can be used in concert with other measures to provide an effective automatic guide.

5. PRACTICALITY

The amount of work required to compute spatial autocorrelation measures is directly tied to the number of regions and the number of neighbors in a data set. Even modest-sized images can have hundreds of thousands of pixels, and therefore regions, in them. The underlying spatial autocorrelation measures are amenable to parallelization based on region. However, PySAL and GeoDa do not appear to use parallelization in their implementations. Therefore, these tools do not provide a practical basis for full-image evaluation, though the underlying spatial autocorrelation measures may still be adapted.

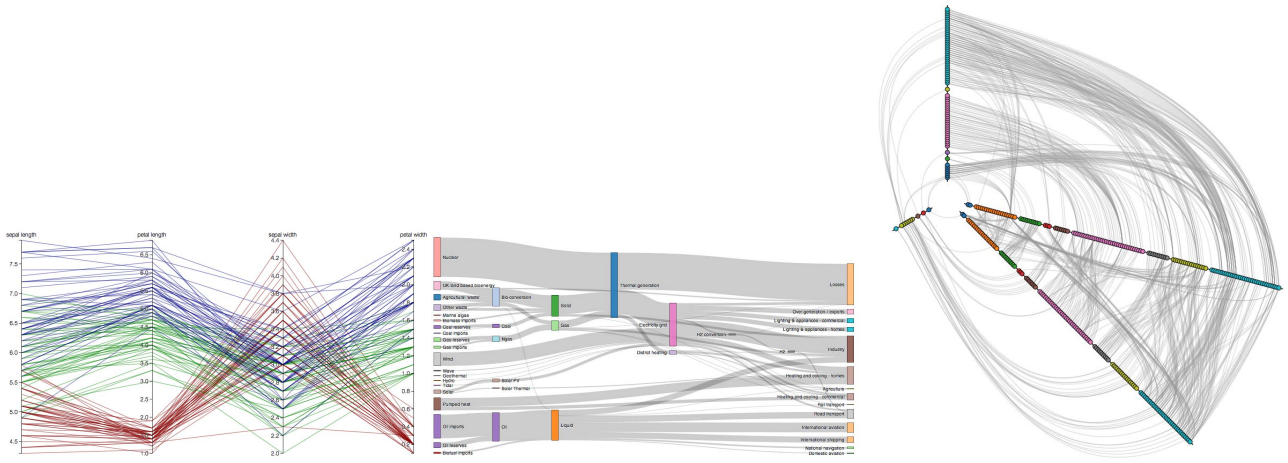
For example, using a neighborhood of Manhattan-distance 5 and full-sized images results in global spatial autocorrelation runtimes in excess of 40 hours (tests were aborted at 40hrs, only three images completed). Using just a neighborhood of distance 1, the runtime was reduced to just over eight hours. Local spatial autocorrelation times exceeded 80hrs before being stopped. Thumbnails of size 50 and 250 were also investigated to see if smaller images were valid approximations for the larger images. Thumbnails with a maximum dimension of 50 had near-interactive execution times in PySAL (analysis average .6 seconds), even for LISA calculations. Analysis on thumbnails with a maximum dimension of 250 required an average of one minute to complete. Given the analysis presented earlier, using thumbnails of size 50 or 250 appear to be acceptable proxies for full-sized images. Therefore, spatial autocorrelation may be valid as on-the-fly evaluation tool, though smaller image sizes are preferred.

6. FUTURE WORK

Informally, we have observed that visualizations with low spatial autocorrelation correspond to those with significant issues for data sets like those used in this study (such as node-link diagrams for networks of more than 100 nodes [10]). Similarly, classic visualization examples have spatial autocorrelation in the middle or upper region of the observed range. This classification rule needs to be confirmed with more detailed study. It is, however, the most promising general rule for using spatial autocorrelation as a guide for visualization design.

We have demonstrated that spatial autocorrelation can separate a set of curated visualizations from photographs. Additionally, we have shown that multiple autocorrelation measures do not significantly differ in their results. However, we have not rigorously addressed the question of identifying "good" visualizations. We have only provided an informal observation for a "goodness" rule. Characterizing "good" visualizations in general is still an open question in the field of information visualization. A canonical test for "goodness" over a specific data set is to encode the same data in multiple ways and perform task analysis on each. Comparing a ranking derived in this manner to spatial autocorrelation would provide significant insight into the applicability of spatial autocorrelation to the general problem of visualization evaluation. Small-scale curated sets currently exist for specific data sets (for example [20, 21]). Collecting/creating a broad data set of this type would be a valuable resource for all visualization analysis research.

Spatial autocorrelation was selected as our measure of investigation because it is appropriate for image analysis, it is a mature tool, and it yields a measure that can be directly compared across images. It specifically considers locality (as



(a) Parallel Coordinates (b) Sankey Diagram (c) HIVE Plot

Figure 10: Plots with distinct regions representing different types of data.

defined by the weights matrix), thereby differing from other image-based evaluation metrics (such as minimum description length [16], entropy and mutual information [7, 23]). Preliminary investigation (see Section 4.7) shows an ensemble of approaches may yield an effective tool for automated evaluation. In addition to existing information-theory based evaluation tools, minimum entropy [15] and synergy [11] may be effective because they can include locality (much like spatial autocorrelation does). Bertini et. al. [4] provide a taxonomy that includes many other candidates for an ensemble approach.

Correlations based on color are an obvious next step. Working with color presents interesting and complex problems, stemming from the need for a continuous dependent variable in spatial autocorrelation. All the spatial autocorrelation metrics rely on addition, subtraction, and division in the dependent variable's space. Though addition and subtraction have analogous operations in color spaces, their statistical interpretation is unclear. Furthermore, an analog for division in color is not apparent. Working with color differences in CIELab space may provide a means of avoiding this difficulty. Since the distance between colors in CIELab space is based on perceptual differences, the color differences have a direct interpretation. However, spatial autocorrelation metrics use a mean of the dependent variable to essentially discuss the difference of a value from the expected value. This expected value calculation requires significant consideration for color. Several values are reasonable, including 50% gray, the CIELab Centroid of the colors used in the image, or a centroid from a local neighborhood. Using a multi-variate measure on the individual components of an alternative color encoding is another avenue.

Visualizations naturally have multiple regions. Even in simple visualizations, there is a data region surrounded by axes and legends. Parallel coordinates, Sankey diagrams, and HIVE plots [14] have interleaved data axis regions and connecting regions, (see Figure 10). These different regions may deserve different treatments, but our analysis did not separate the different region types. At least in the LISA analysis, contextualizing information (such as labels and gridlines) had high significance, though their actual function is of contextualization (see Figure 5). These regions may be treated differently either by producing plots that exclude them or by building a weights matrix that encodes the relationships between regions.

7. CONCLUSIONS

Spatial autocorrelation is a well-developed tool for evaluating patterns distributed in two dimensions. This paper has demonstrated that it provides a plausible basis for automatic visualization evaluation over a broad range of visualization types *without* special treatment for different visualization types. We have shown that downsampling images and using limited neighborhood sizes still leads to useful results and greatly increases practicality. We have also shown evidence that spatial autocorrelation can be combined with other visualization metrics to create an effective ensemble.

8. REFERENCES

[1] Python Imaging Library. http://www.pythonware.com/products/pil/index.htm, November 2012.

[2] L. Anselin. Local indicators of spatial association-LISA. *Geographical analysis*, 27(2):93–115, 1995.

[3] L. Anselin, I. Syabri, and Y. Kho. GeoDa: An Introduction to Spatial Data Analysis. *Geographical Analysis*, 38(1):5–22, 2006.

[4] E. Bertini and G. Santucci. Visual quality metrics. In *Proceedings of the AVI workshop on Beyond time and errors: novel evaluation methods for information visualization (BELIV)*, pages 1–5, New York, NY, USA, 2006. ACM.

[5] M. Bostock. d3.js (gallery). http://github.com/mbostock/d3/wiki/Gallery, November 2012.

[6] M. Bostock. Protovis (gallery). http://mbostock.github.com/protovis/ex/, November 2012.

[7] M. Chen and H. Jaenicke. An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16:1206–1215, 2010.

[8] A. Dasgupta and R. Kosara. Pargnostics: Screen-space metrics for parallel coordinates. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1017–1026, Nov. 2010.

[9] R. C. Geary. The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician*, 5(3), 1954.

[10] M. Ghoniem, J.-D. Fekete, and P. Castagliola. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. *Information Visualization*, 4(2):114–135, July 2005.

[11] V. Griffith and C. Koch. Quantifying synergistic mutual information. *CoRR*, abs/1205.4265, 2012.

[12] J. Heer. Prefuse (gallery). http://prefuse.org/gallery/, November 2012.

[13] R. Keller, C. M. Eckert, and P. J. Clarkson. Matrices or node-link diagrams: which visual representation is better for visualising connectivity models? *Information Visualization*, 5(1):62–76, Mar. 2006.

[14] M. Krzywinski, I. Birol, S. J. M. Jones, and M. A. Marra. Hive plots—rational approach to visualizing networks. *Briefings in Bioinformatics*, 13(5):627–644, Sept. 2012.

[15] H. Li, K. Zhang, and T. Jiang. Minimum entropy clustering and applications to gene expression analysis. In *Proceedings of the IEEE Computational Systems Bioinformatics Conference (CSB)*, pages 142–151, Los Alamitos, CA, USA, 2004. IEEE Computer Society.

[16] I. Liiv. Towards information-theoretic visualization evaluation measure: a practical example for bertin's matrices. In *Proceedings of the 3rd workshop on Beyond time and errors: novel evaluation methods for information visualization (BELIV)*, pages 24–28, New York, NY, USA, 2010. ACM.

[17] P. A. P. Moran. Notes on Continuous Stochastic Phenomena. *Biometrika*, 37(1/2):17–23, June 1950.

[18] S. J. Rey and L. Anselin. Pysal: A python library of spatial analytical methods. *The Review of Regional Studies*, 37(1):5–27, 2007.

[19] J. Schneidewind, M. Sips, and D. Keim. Pixnostics: Towards measuring the value of visualization. In *IEEE Symposium On Visual Analytics Science And Technology (VAST)*, pages 199 –206, 31 2006-nov. 2 2006.

[20] M. Tory, D. Sprague, F. Wu, W. Y. So, and T. Munzner. Spatialization design: Comparing points and landscapes. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1262–1269, Nov. 2007.

[21] M. Tory, C. Swindells, and R. Dreezer. Comparing dot and landscape spatializations for visual memory differences. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1033–1040, Nov. 2009.

[22] J. W. Tukey and P. A. Tukey. Computer Graphics and Exploratory Data Analysis: An Introduction. In *Proceedings of the Sixth Annual Conference and Exposition: Computer Graphics*, pages 773–785, Fairfax, VA, 1985. Nat. Computer Graphics Association.

[23] C. Wang and H.-W. Shen. Information theory in scientific visualization. *Entropy*, 13(1):254–273, 2011.

[24] L. Wilkinson, A. Anand, and R. Grossman. Graph-theoretic scagnostics. In *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS)*, pages 157– 164, Los Alamitos, CA, USA, 2005. IEEE Computer Society.