

Quantification of Heat Map Data Displays for High-Throughput Analysis

Paul Juneau*

Statistical Services Group, Truven Health Analytics, Boyds, USA

Abstract

Heat maps have been used as a means to visualize high-density information in settings as diverse as astronomy, business analysis, and meteorology. Discovery biology research teams have also used heat maps to visualize gene clusters in genomics investigations or to study amino acid distribution in protein sequence analysis. Commercially available software packages, like *Spotfire*[®] or *SAS JMP*[®] afford scientific investigators the ability to construct heat maps and visualize information from studies, yet do not offer any form of summary statistic that would be useful in high-throughput investigations comparing the results of a large number of data visualizations simultaneously or viewing changes in the display longitudinally (over time).

Previously, Juneau suggested the usage of Plotnick's characterization of lacunarity (1996) for two-dimensional heat map data displays in two colors or shades. For c ($c > 2$) discrete shades (in a monochromatic map) or hues (in a full color display), the author will suggest a modification to Plotnick's approach using the underlying gliding box approach developed by Allain and Cloitre, but with an alteration in the means of counting features.

Keywords: Heat maps; Gliding box; Lacunarity; Quantification

Introduction

Background

Heat maps have been employed as a data visualization tool in the social sciences since the late nineteenth century [1-3], and more recently in arenas as diverse as astronomy [4-6], business analysis [7-9], meteorology [10-12] and quantum mechanics [13]. Changes in the ambient conditions of complicated systems like galaxies, the stock market, or large meteorological phenomena can be readily displayed via differences in color or grayscale to assist investigators in hypothesis generating or data interpretation activities. Heat maps afford investigators the ability to study high-density data sets in a single visualization, while maintaining measurement relationships and data integrity [14-16].

The popularity of heat maps has grown substantially with developments in the field of bioinformatics [17]. Numerous examples of published methods employing heat maps exist [18-31] however, the author has not, to date, seen an attempt to represent the information present in these very widely used data visualizations with a single summary statistic. Commercially available software packages, like *Spotfire*[®] or *SAS JMP*[®] afford scientific investigators the ability to construct heat maps and visualize information from studies, yet do not offer any form of summary statistic that would be useful in high-throughput investigations comparing the results of a large number of data visualizations simultaneously or viewing changes in the display longitudinally (over time). Previously, Weinstein (1997) had suggested using a "difference heat map" to compare visualizations pre and post-treatment. This approach seems tenable if only two time points are under consideration, but would result in several pre-post difference heat maps if one were interested in comparing change with baseline over time or all pair-wise comparisons of time responses.

One approach to numerically summarizing the content of a heat map might be to use a statistic like the percentage of the entire map colored or shaded by a specific hue or degree of brightness. Figure 1.1.1 illustrates three examples of tri-colored heat maps.

One could summarize each heat map by the percentage of tiles of a given color relative to the whole. The use of such a statistic would not differentiate the apparent geometry of the heat map depicted in Figure

1.1.1a from that of the other two. The geometries of Figures 1.1.1 b and 1.1.1 c might suggest an underlying block structure relationship between rows and columns, which possibly could be related to an underlying multivariate mechanism with a block diagonal covariance matrix [32]. Figure 1.1.1 c has a subtle difference in appearance to Figure 1.1.1 b. A black diagonal band bisects the block structure. Thus, a percentage approach does not account for the overall pattern presented in a heat map and would therefore not serve as a useful numerical summary of the relationships suggested.

A second approach could be to characterize the spacing-filling geometry of the tiles in the heat map via a Hausdorff-like dimension [1]. Consider a non-empty subset of R^n , say S , which may be covered with sets of diameter at most μ ($\mu > 0$), such that the diameter of the covering sets is normalized to S (i.e., the size of S is considered unity).

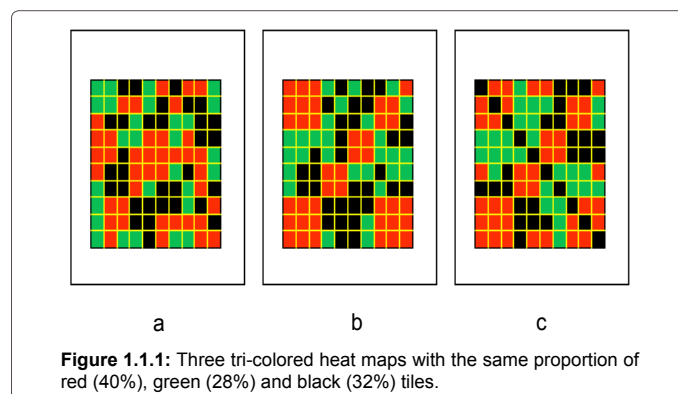


Figure 1.1.1: Three tri-colored heat maps with the same proportion of red (40%), green (28%) and black (32%) tiles.

*Corresponding author: Paul Juneau, Senior Statistician, Statistical Services Group, Truven Health Analytics, Boyds, USA, Tel: 240-246-7627; E-mail: paul.juneau@truvenhealth.com

Received May 07, 2015; Accepted May 28, 2015; Published June 09, 2015

Citation: Juneau P (2015) Quantification of Heat Map Data Displays for High-Throughput Analysis 6: 146. doi:10.4172/2153-0645.1000146

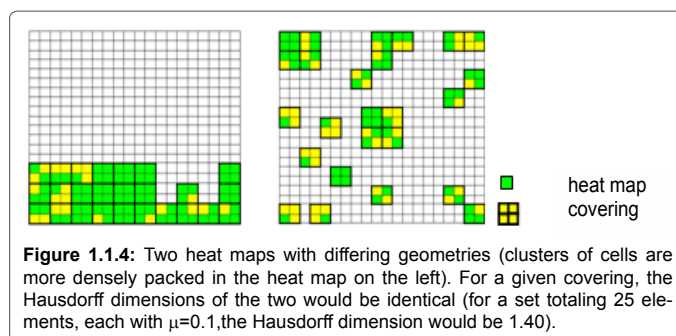
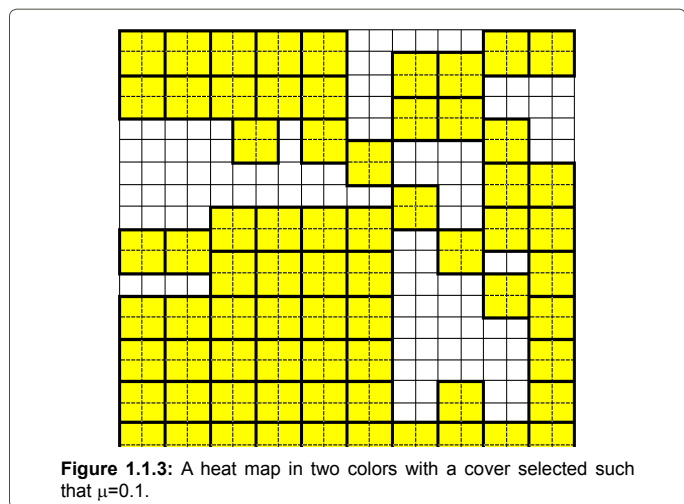
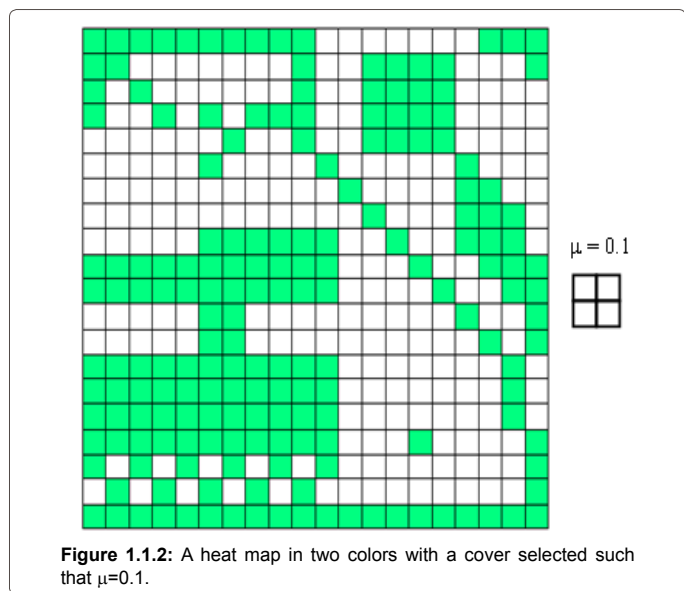
Copyright: © 2015 Juneau P. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The Hausdorff dimension is calculated in mathematics as a limiting procedure related to the logarithm of the number of sets of diameter at most, say μ , that cover S as the logarithm of the diameter of these sets approaches zero [33]. In practice, it might be the case that size of the features that form a pattern might be of interest, as well as the patterns themselves. Thus, usage of the Hausdorff dimension in a strict mathematical sense might not be practical because of the investigator's desire to consider features at least as large as μ ($\mu > 0$).

Consider the bi-colored heat map displayed in Figure 1.1.2 with a covering such that $\mu = 0.1$. The heat map in Figure 1.1.2 could be covered by a set as illustrated in Figure 1.1.3.

From the covering it is easy to calculate a Hausdorff dimension for $\mu = 0.1$. The major shortcoming of this approach is that, as was the cause with the percentage summary, the geometry of two heat maps may be markedly different; however, the Hausdorff dimension could be identical. An example of this phenomenon is illustrated in Figure 1.1.4.

This limitation of such dimension measurements was first recognized by Mandelbrot [34] in the context of numerically



summarizing fractals. Mandelbrot advocated the usage of a quantity called the *lacunarity* to numerically summarize fractals. The root of the word lacunarity is the Latin *lacunae*, meaning “space” or “hole”. Thus, as the Hausdorff dimension characterizes the “space-filling” properties of a set, the lacunarity measures the presence of gaps or holes in the set.

Juneau [1] suggested the usage of lacunarity to characterize heat maps in two colors or shades, based upon a method suggested by Plotnick [35]. Plotnick's method was based upon a more general case originally developed by Allain and Cloitre [2]. This approach will provide an investigator with a measure of the “gapiness” for one shade or color relative to a second. The balance of this paper will be based upon a suggested method for a setting with c color or shades, for $c > 2$. Section 1.2 will summarize the gliding box approach of Allain and Cloitre and illustrate its behavior for a heat map in 2 colors or shades. Section 2 will introduce a proposal for a form of modified lacunarity for the setting of c colors or shades ($c > 2$) and highlight the scaling feature of the approach. Section 3 will provide examples of three applications of the technique: cluster analysis, the summarization of longitudinal data in meteorology, and genomics.

The gliding box approach of Allain and Cloitre and the calculation of a heat map's lacunarity for heat maps in two colors or shades

For some heat map, $H \subset \mathbb{R}^2$, let P represent a *board* [36] that partitions H :

(1) Define p to be a polygon with s sides and diameter $(p) = \rho$, where $\text{diameter}(\cdot) = \text{maximum of the lengths of the } s \text{ sides of } p$;

(2) $P = \{ p \subseteq H \mid \bigcup p = H \text{ and } p_i \cap p_j = \emptyset \text{ for } i, j \in Z^+ \text{ and } i \neq j \}$;

(3) $\text{diameter}(p_i) = \text{diameter}(p_j) \forall i, j \in Z^+$. Without loss of generality, p can be defined to be a rectangle. A subset $p \subset P$ can be called a *feature* of the heat map. Figure 1.2.1 provides an illustration of H, P , and the partitioning of H by P .

Define a box, $B \subset P$, to consist of a set of contiguous features, p , whose union is similar to the polygon, p . Figure 3.2.2 illustrates the relationship between B & P .

The procedure for the gliding box approach is as follows. The box, B , traverses the length of H , beginning at the upper left corner of H . Define an indicator function, $\chi(p)$ as follows:

$$\chi(k) = \begin{cases} 1 & \text{if } p \text{ is gray (for } p \in B) \\ 0 & \text{otherwise} \end{cases} \tag{1.2.1}$$

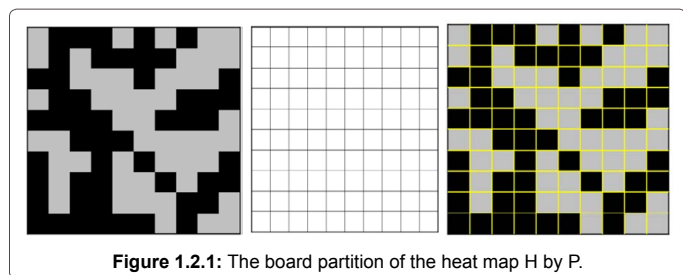


Figure 1.2.1: The board partition of the heat map H by P.

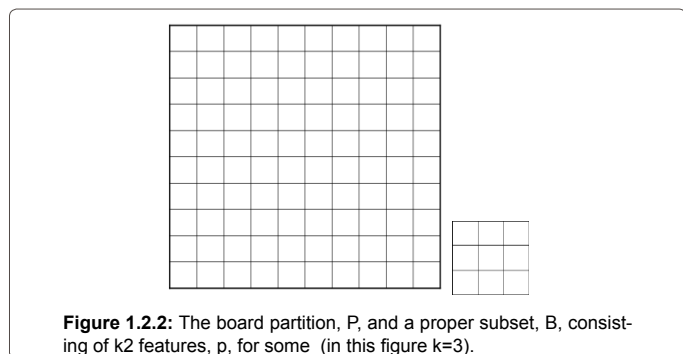


Figure 1.2.2: The board partition, P, and a proper subset, B, consisting of k2 features, p, for some (in this figure k=3).

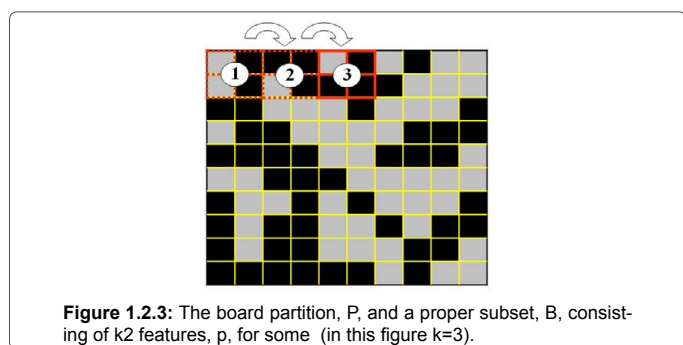


Figure 1.2.3: The board partition, P, and a proper subset, B, consisting of k2 features, p, for some (in this figure k=3).

Call the sum of the values of the indicator variable $\chi(k)$ over all features $p \subset B$ for the m^{th} movement of B across H, the score, ξ_m . The box is moved k features to the right for a box B of diameter, k . Figure 1.2.3 illustrates the movement of B across H. Thus, for each movement, m , of B across H, a single score is calculated by summing up all of the values of $\chi(k)$:

$$\xi_m = \sum_{p \subset B} \chi(k) \tag{1.2.2}$$

The total set of scores, say T, ($m=1$ to T) for the movement of B T times across H, may now be tallied to form a discrete probability mass function for all possible values from 0 to k^2 for a box of diameter, say, k . Call the probability mass function for all values from 0 to k^2 , Ψ .

The lacunarity of H (Mandelbrot), $\Gamma(k, T)$, may then be defined as:

$$\Gamma(k, T) = M_2(\xi) / (M_1(\xi))^2 \tag{1.2.3}$$

where

$$M_2 = (1/T) \sum_{m=1}^T \xi_m^2 \Psi \tag{1.2.4} \text{ and } M_1 = (1/T) \sum_{m=1}^T \xi_m \Psi \tag{1.2.5}$$

(i.e., M_1 and M_2 represent the first and second un-centered moments for the scores, $\xi_1, \xi_2, \dots, \xi_T$).

For two given sets, the one with the larger value of Γ will be a set of gray features more diffusely distributed throughout; i.e., the occurrence of black regions will be more frequent and their size relatively larger. Figure 3.2.4 provides an illustration of three sets H_1, H_2 and H_3 and their corresponding values of Γ .

How was Γ determined for the left-hand panel of Figure 1.2.4? If one employs a gliding box of size $k=2$, the first row of H_1 would have scores of 4, 1, 3, 1 and 3. In the second row, the scores would be 2, 3, 2, 3 and 2. Thus, if one were to proceed moving the gliding box over the entire heat map for the remaining three rows:

$$M_1 = (1/25) (0^2 + 1^7 + 2^7 + 3^8 + 4^1) = 49/25,$$

$$M_2 = (1/25) (0^2 + 2^2 + 1^2 + 7 + 2^2 + 7 + 3^2 + 8 + 4^2 + 1) = 123/25,$$

$$\Gamma(2) = M_2 / M_1^2 = 1.27.$$

Modification to the Allain and Cloitre approach

Development of a modified lacunarity using the Allain and Cloitre approach

Consider the heat map, H_c , depicted in Figure 2.1.1. As opposed to the heat map, H, depicted in Figure 2.1.1, this display is in more than two colors or shades. In Figure 2.2.3, recall that the procedure counted the number of features of a desired shade (gray). In essence, the procedure described in Section 1.2 describes the distribution of gray features relative to the distribution of the complementary features

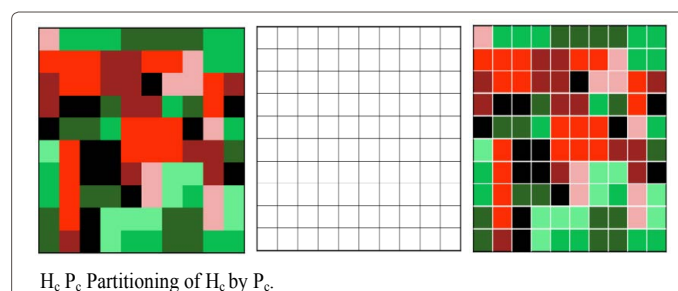


Figure 2.1.1: The board partition of the multicolored or multi-shaded heat map, H_c , by P_c .

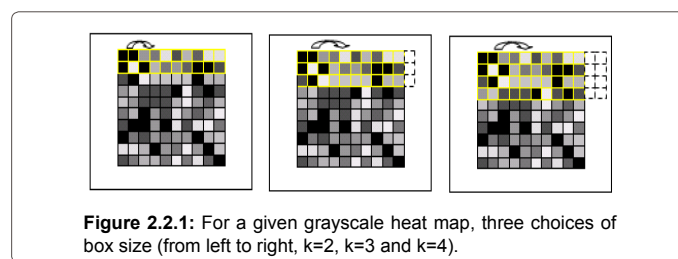


Figure 2.2.1: For a given grayscale heat map, three choices of box size (from left to right, $k=2, k=3$ and $k=4$).

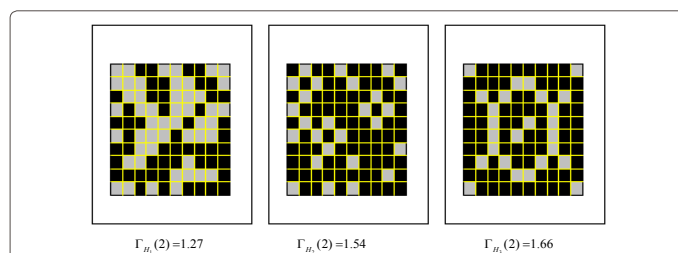


Figure 1.2.4: Three heat maps of varying amounts of gray shading and their corresponding lacunarity values.

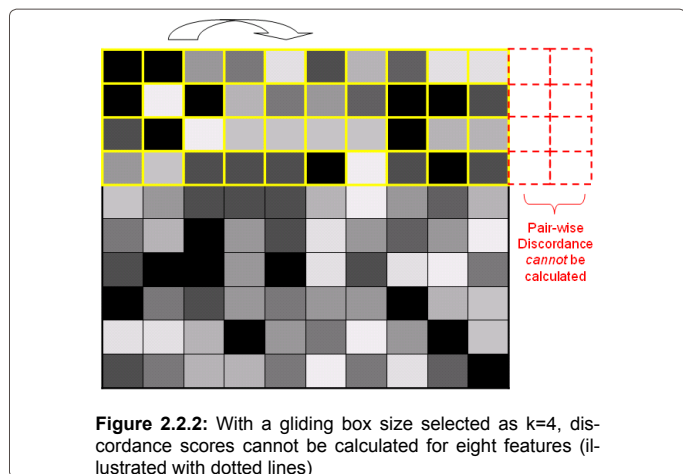


Figure 2.2.2: With a gliding box size selected as $k=4$, discordance scores cannot be calculated for eight features (illustrated with dotted lines)

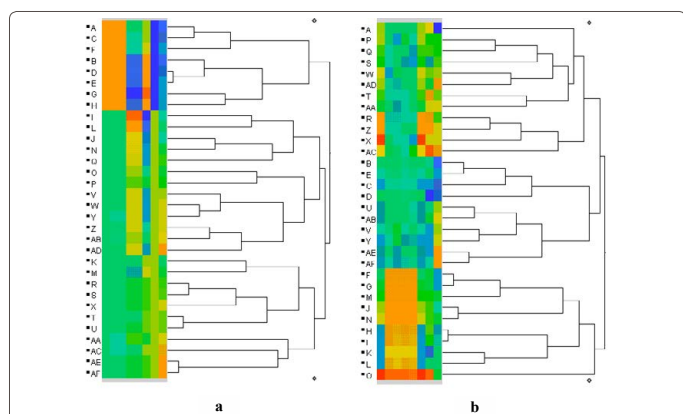


Figure 3.1.1: The results of two cluster analyses. Figure 3.1.1 (a) consists of 32 items (8-tuples) of simulated uniform (0,1) variates with a high degree of clustering induced by the constraint of perfect agreement between variates in 25% of the cases. Figure 3.1.1 (b) consists of 32 8-tuples simulated based upon functions of Gaussian variates and linear combinations of a subset of the components of each item.

(those that are not gray). The goal is to develop a form of measurement that simultaneously summarizes the clustering of colors or shades for feature subsets of H_c relative to the other colors or shades.

One possible approach to developing a modified Allain and Cloitre procedure for multi-shaded or multi-colored images or heat maps would be to count the number of neighboring discordant pairs within B as it traverses over H_c . In an intuitive sense, studying the distributional properties of the discordant pairs of features within the gliding box B provides a summary of the density of sets of features contained within H_c . Just as the lacunarity, Γ , defined in equation 1.2.3 for a two-shaded or bi-colored heat map increases as the number of the subsets with large clusters of features with the desired shade or color decreases, a lacunarity, Γ_{mac} , based upon modifying the gliding box algorithm of Allain and Cloitre that counts discordant pairs will increase as the number of subsets with large clusters of any color or shade decreases for a given heat map H_c .

Consider the gliding box, B_c , as defined in Figure 2.1.2. Label each feature p_{rs} to represent the feature in the r^{th} row and s^{th} column (i.e., in a form of matrix element notation). Define a feature, p_{rs} , to be adjacent to p_{rs} if p_{rs} and p_{rs} have sides with a common vertex. Now, define an indicator function, $\delta(p)$ as follows:

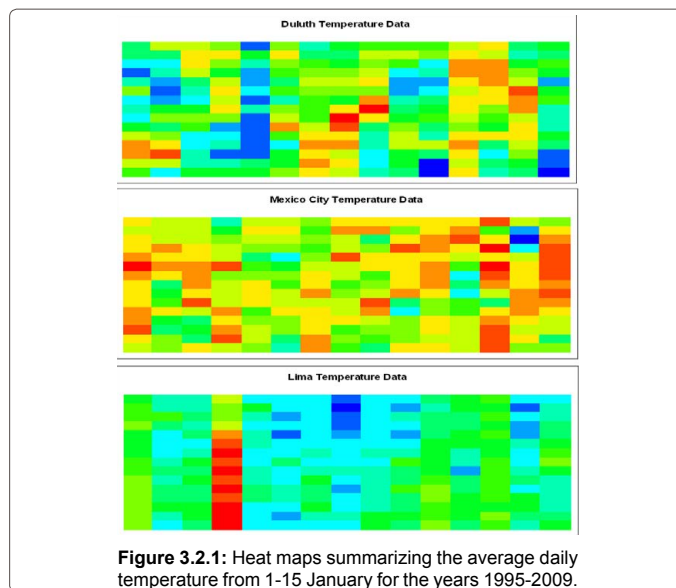
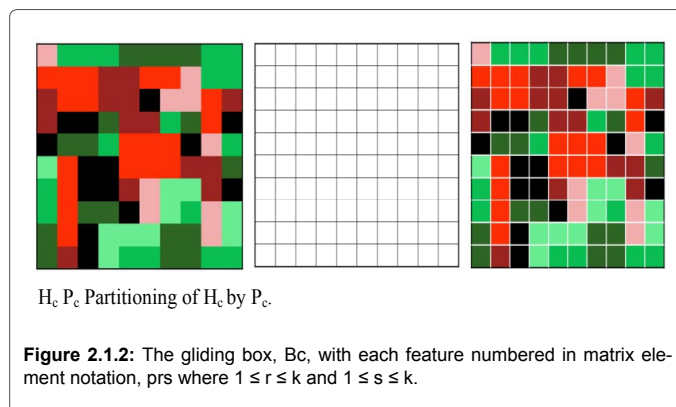


Figure 3.2.1: Heat maps summarizing the average daily temperature from 1-15 January for the years 1995-2009.



$H_c P_c$ Partitioning of H_c by P_c .

Figure 2.1.2: The gliding box, B_c , with each feature numbered in matrix element notation, p_{rs} where $1 \leq r \leq k$ and $1 \leq s \leq k$.

1 if the color or shade of p_{rs} does not agree with that of p_{a_r, a_s} for

$$p_{rs}, p_{a_r, a_s} \subset B_c$$

$$\delta_{rs}(k) = \quad (2.1.1)$$

0 otherwise

Call the sum of the values of the indicator variable $\delta_{rs}(k)$ over all features $p_{rs} \subset B$ the *discordance score*, Δ . The box is moved k features to the right for a box B of diameter, k . Thus, for each movement, m , of B across H , a single score is calculated by summing up all of the values of $\delta(k)$:

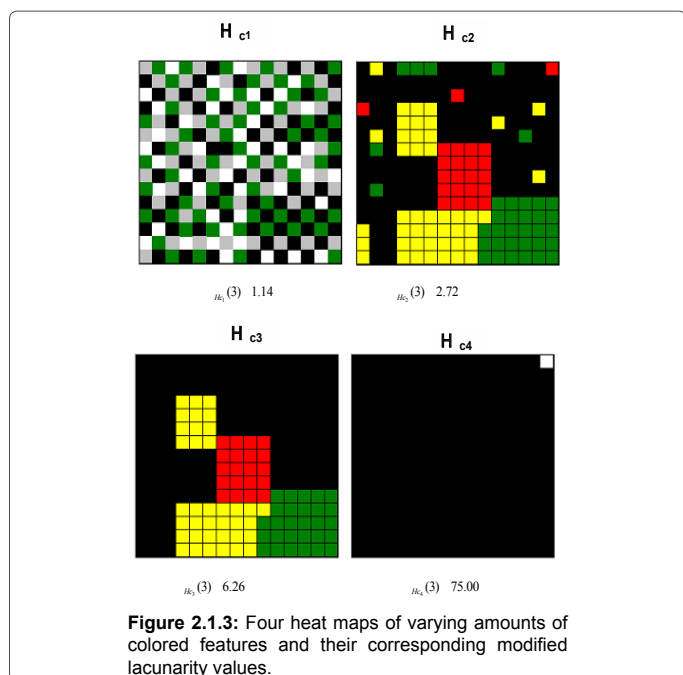
$$\Delta_m = \sum_{p_{rs} \subset B} \delta_{rs}(k) \quad (2.1.2)$$

Then a modified lacunarity, based upon the Alain and Cloitre approach Γ_{mac} , may be defined in a similar fashion as previously, in equation 1.2.3:

$$M_2(\Delta_m) / (M_1(\Delta_m))^2 \quad (2.1.3)$$

for all m movements of B_c across the heat map. The same definition of the first and second moments expressed by the notation in equations 1.2.4 and 1.2.5 would hold for equation 2.2.3. The modified lacunaritys of four heat map data displays with varying amounts of colored features are shown in Figure 2.1.3.

How was the value $\Gamma_{H_{c_1}}(3)$ determined? Using an approach similar



to the one used for the calculations of the left-hand panel in Figure 1.2.4, although, in this instance, with $k=3$:

$$M_1 = (1/225) \cdot (0^0 + 1^0 + 2^3 + 3^7 + 4^4 + 5^4 + 6^1 + 7^6 + 8^5) = 3.80,$$

$$M_2 = (1/225) \cdot (0^{2^2} + 1^{2^7} + 2^{2^3} + 3^{2^7} + 4^{2^4} + 5^{2^4} + 6^{2^1} + 7^{2^6} + 8^{2^5}) = 16.45$$

$$\Gamma(3) = M_2 / M_1^2 = 1.14.$$

The Choice of box size and its influence on the calculation of Γ_{mac}

Consider the situation illustrated in Figure 2.2.1. For a given choice of the gliding box's diameter, the coverage of the box over the heat map can result in a different number of features that may be contained within the gliding box as it completes its first row of coverage. Figure 2.2.1 illustrates three choices of box size. When $k=2$, note that for each row the movement of the box will result in the same number of partition pieces covered; however, this is not the case for $k=3$ and $k=4$. The algorithm suggested in Section 2.1 may still be employed in the circumstances illustrated in Figure 2.2.1 for $k=3$ and $k=4$ with minimal effect on the calculation of the modified lacunarity if the number of movements of the box is large relative to the size of the heat map. When a box spans a region outside of the heat map, the author recommends using the convention that the discordances be measured only on the portion of the box that is covering the heat map. An illustration of this suggested convention is illustrated in Figure 2.2.2.

Borys [37] derived the relationship between a reference lacunarity, say Γ_0 , with a gliding box with a diameter μ_0 (normalized to the heat map such that the heat map's size is considered unity) and the lacunarity determined for a box with a diameter μ (normalized to the same heat map), say, Γ :

$$\Gamma = \Gamma_0 (\mu_0 / \mu)^{D-2} \tag{2.2.1}$$

where D is the generalized fractal dimension reported in Ott [38]. For a covering of a set, it is possible to determine a fixed D , and a lacunarity calculated on a normalized diameter of μ can be compared

to that of one based on a normalized diameter of μ_0 . Thus, despite the fact that the user of a lacunarity-based technique is free to choose his or her value of scale, the lacunarity values for different choices of scale can be related via 2.2.1. If two users cannot agree on a common diameter for the box, one can easily transform the corresponding lacunarity values from one scale to another.

Illustrations of the Technique in Simulated and Real Data

An example of the calculation of the modified lacunarity in the arena of cluster analysis with simulated data

As mentioned in Section 1.1, heat maps are used to summarize results frequently in the field of bioinformatics, primarily after a cluster analysis is performed. Packages like SAS JMP[®] (version 9) allow users to examine the results of cluster analyses via a color map within the multivariate analysis module in the analysis platform of the product. Two data sets were simulated with SAS JMP[®]. The first data set consisted of 32 8-tuples (items with 8 attributes) of simulated uniform (0,1) variates. Perfect agreement between the components of 25% of the cases was artificially induced in the simulation to create large block structures in the SAS JMP[®] color map (i.e., to artificially create a very dense cluster of a small set of the items). A second data set consisted of 32 items. The first attribute or component of each item was simulated from a standard Gaussian distribution. For the next 5 of the components, half of the items were assigned linear combinations of the first attribute; the remaining half was assigned random Gaussian noise. The sixth and seventh components consisted of values simulated from a Gaussian distribution. The eighth component was simulated from a uniform (0,1) distribution. The two data sets were independently analyzed using the default options (hierarchical and Ward's method) in the multivariate analysis module within the analysis platform of SAS JMP[®]. The results of the cluster analyses for the two data sets are shown in Figure 3.1.1.

$\Gamma(2)$ values were calculated for the two heat maps illustrated in Figures 3.1.1 (a) and (b) as follows:

$$\Gamma_a(2): M_1 = (1/64) \cdot (0^1 + 6^6 + 8^2 + 10^2 + 12^8) = 6.6875$$

$$M_2 = (1/64) \cdot (0^{2^1} + 6^{2^6} + 8^{2^2} + 10^{2^2} + 12^{2^8}) = 64.6250$$

$$\Gamma_a(2) = M_2 / M_1^2 = 1.45.$$

$$\Gamma_b(2): M_1 = (1/64) \cdot (0^8 + 6^9 + 8^9 + 10^{22} + 12^{16}) = 6.6875$$

$$M_2 = (1/64) \cdot (0^{2^8} + 6^{2^9} + 8^{2^9} + 10^{2^22} + 12^{2^16}) = 64.6250$$

$$\Gamma_b(2) = M_2 / M_1^2 = 1.19.$$

The color map in Figure 3.1.1 (a) contains more regions with large blocks of a single color than the color map in Figure 3.1.1 (b). Figure 3.1.1 (a) contains large blocks of orange, green and blue, while Figure 3.1.1 (b) contains only a single large block of orange. If the two color maps were considered as landscapes, metaphorically, it would be as if Figure 3.1.1 (a) has several flat regions while Figure 3.1.1 (b) is more varied with only one planar region. The modified lacunarities describe this phenomenon: the modified lacunarity calculated for Figure 3.1.1 (a) is larger than the modified lacunarity of Figure 3.1.1 (b), suggesting that the color map of the first has more "holes" or regions of a single color (consistency) than that of the second.

An example of the calculation of the modified lacunarity with longitudinal temperature data from three cities

A real application of the modified lacunarity can be applied to the average daily temperatures measured in the first fifteen days of January from the years 1995-2009 for Duluth, Minnesota (USA), Mexico City, Mexico and Lima, Peru. These data can be found at <http://www.engr.udayton.edu/weather/>, the web site for the University of Dayton's *Temperature Data Archive*. Data were arranged in three 15x15 arrays. The columns of the arrays represent a single year (beginning with 1995 on the left); the rows days 1-15 of January for that year (beginning with 1 January in the first row).

The values were entered in MS Excel 2003. A MS Excel macro, available as freeware from <http://bitesizebio.com/2009/02/03/how-to-create-a-heatmap-in-excel>, was used to generate the heat map summaries illustrated in Figure 3.2.1.

A gross inspection of the data represented in the three heat maps allows the observer to glean some information. As expected, the average daily temperature of Lima, Peru is more consistent than the other two cities (because of its latitude and topology), as evident by the large blocks of color present in the corresponding heat map. Moreover, the average daily temperature of Duluth, Minnesota is more varied, as is evident by the relatively smaller blocks of color present in its corresponding heat map. These evaluations are highly subjective. If one needed to organize a large series of heat maps, of which these three are representative, an index of temperature consistency might be of value.

If one were interested in a single index quantifying the consistency of the temperatures for the three cities in the first 15 days of January for 15 years, he or she could use the modified lacunarity. Suppose that a meteorologist were interested in the agreement of temperatures between three days over three years as an estimate of consistency. He or she could use the modified lacunarity as a possible descriptive statistic. $\Gamma_{Duluth}(1,3)=1.13$, $\Gamma_{Mexico\ City}(1,3)=1.19$ and $\Gamma_{Lima}(1,3)=1.21$. Once again, the modified lacunarity summarizes the topology of the heat maps: larger values reflect the presence of more regions with a consistent temperature.

An example of the calculation of the modified lacunarity for a study of longitudinal system-based analysis of transcriptional responses to Type I Interferons

A third illustration of the modified calculation was made based on Figure 1.D in [39] a study of the transcriptional response to Type I Interferons. If one were to use a 2x2 gliding box, it would be possible to use the figure to estimate $\Gamma(2)$ for the portions of the heat map corresponding to APP x IFN- β 1a (upper left-hand 10x6 portion), APPxIFN- α 2b (right-hand side, adjacent 10x6 portion), JS x APP x IFN- β 1a (12x6 portion below APP x IFN- β 1a) and APPxIFN- α 2b (12x6 portion right-hand side, adjacent portion). The values would be 1.69, 1.13, 1.10 and 0.91. These values are consistent with the concept of the modified lacunarity: to find large homogeneous blocks within the heat map with little diversity in color. The figure with the most color change is APPxIFN- α 2b, suggesting a greater change in activity (varying from very dark blue to very bright yellow, or from -1 to 1, respectively).

Discussion

The concept of lacunarity introduced by Mandelbrot is used as a summary statistic in many applications [40-45]. To date, the lacunarity statistic has been applied only in settings with images of two colors or shades and has not been applied to summarize the content of heat

map data displays. The proposed modified lacunarity statistic affords investigators the option of summarizing or indexing large numbers of heat maps based upon the presence of large monochromatic blocks of features. The modified lacunarity proposed in this work is easy to compute and its interpretation intuitive in applied settings.

The limitation of this proposed statistic is its small variation relative to the larger variation perceived by the interpreter of a heat map. This feature is evident in Figure 3.2.1. The cause of this small variation in the value of the modified lacunarity is most likely due to the smaller correlations in yearly temperatures for a given day relative to the correlations between days within a given year. Where larger correlations exist between rows and columns in a heat map, larger blocks of uniform color can exist (see Figure 3.1.1). In the presence of larger blocks of color, values of the modified lacunarity can be more potentially more discriminating without the need for several decimal places (see Figure 2.1.3).

All of the lacunarity calculations performed in this work were calculated with MS Excel 2003. With the current state of the art in computer software, it seems plausible that one could output the color codes used to produce an image, transfer them to a simple spreadsheet and automate the gliding box process for several box sizes. Thus, due to the relatively simple implementation, ease of calculation and its intuitiveness, the modified lacunarity has the potential to be a new tool that can be used in many scientific arenas to aid in exploratory data analysis and subsequent hypothesis generation.

Acknowledgement

The author would like to thank Dr. Dianne Camp for her thorough review and subsequent helpful commentary that improved the overall quality of this work. The author would also like to recognize the excellent suggestion of one of the reviewers to add an example involving longitudinal analysis. This recommendation further added to aid in the intuitive understanding of the principle of the modified lacunarity in practice.

References

1. Juneau P (2005) Quantification of Heat Map Data Displays via Lacunarity, ASA Proceedings of the Joint Statistical Meetings. American Statistical Association 4122-4126 (Alexandria, VA).
2. Allain C, Cloitre M (1991) Characterizing the lacunarity of random and deterministic fractal sets. *Phys Rev A* 44: 3552-3558.
3. Wilkinson L, Friendly M (2009) The History of the Cluster Heat Map. *The American Statistician* 63: 179-184.
4. Briel UG, Henry JP (1996) An X-Ray Temperature Map of Abell 1795, A Galaxy Cluster in Hydrostatic Equilibrium. *The Astrophysical Journal* 472: 131-136.
5. Bourdin H, Sauvageot JL, Slezak E, Bijaoui A, Teyssier R (2004) Temperature Map Computation for X-ray Clusters of Galaxies. *Astronomy and Astrophysics* 414: 429-443.
6. Hindshaw G, Nolte MR, Bennett CL, Bean R, Doré O, et al. (2007) Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature Analysis. *The Astrophysical Journal Supplement Series* 170: 288-334.
7. Few S (2006) Multivariate Analysis Using Heatmaps. *BeyeNetwork: Global Coverage of the Business Intelligence Ecosystem*.
8. NASDAQ (2009) ETF Dynamic Heat map®.
9. Woolsey M (2009) Map: Where Americans are Living Well. *Forbes*.
10. Clifford D, Davenport I, Gurney R, Sandells M (2008) Snow: An Essential Climate Variable. National Centre for Earth Observation Annual Science Meeting: Exposing how the Earth system works.
11. Merchant CJ, Llewellyn-Jones D, Saunders RW, Rayner NA, Kent EC, et al. (2008) Deriving a Sea Surface Temperature Record Suitable for Climate Change Research from the Along-Track Scanning Radiometers. *Advances in Space Research* 41: 1-11.

12. National Oceanic and Atmospheric Administration (NOAA) Environmental Visualization Laboratory (2009) August 2009: One of the Warmest on Record.
13. Akagi H, Otobe T, Staudte A, Shiner A, Turner F, et al. (2009) Laser tunnel ionization from multiple orbitals in HCl. *Science* 325: 1364-1367.
14. Wu W, Noble WS (2004) Genomic data visualization on the Web. *Bioinformatics* 20: 1804-1805.
15. Auman JT, Boorman GA, Wilson RE, Travlos GS, Paules RS (2007) Heat map visualization of high-density clinical chemistry data. *Physiol Genomics* 31: 352-356.
16. Byrnes RW, Cotter D, Maer A, Li J, Nadeau D, et al. (2009) An editor for pathway drawing and data visualization in the Biopathways Workbench. *BMC Syst Biol* 3: 99.
17. Weinstein JN (2008) Biochemistry. A postgenomic visual icon. *Science* 319: 1772-1773.
18. Kim SK, Lund J, Kiraly M, Duke K, Jiang M, et al. (2001) A gene expression map for *Caenorhabditis elegans*. *Science* 293: 2087-2092.
19. Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, et al. (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318: 801-806.
20. Neely LA, Patel S, Garver J, Gallo M, Hackett M, et al. (2006) A single-molecule method for the quantitation of microRNA gene expression. *Nat Methods* 3: 41-46.
21. Ovaska K, Laakso M, Hautaniemi S (2008) Fast gene ontology based clustering for microarray experiments. *BioData Min* 1: 11.
22. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature* 460: 479-486.
23. Molina H, Yang Y, Ruch T, Kim JW, Mortensen P et al. (2009) Temporal Profiling of the Adipocyte Proteome during Differentiation using a Five-Plex SILAC Based Strategy. *Journal of Proteome Research* 8: 48-58.
24. Payton JE, Grieselhuber NR, Chang LW, Murakami M, Geiss GK, et al. (2009) High Throughput Digital Quantification of mRNA Abundance in Primary Human Acute Myeloid Leukemia Samples. *Journal of Clinical Investigation* 119: 1714-1726.
25. Trimboli AJ, Cantemir-Stone CZ, Li F, Wallace JA, Merchant A, et al. (2009) Pten in Stromal Fibroblasts Suppresses Mammary Epithelial Tumors. *Nature* 461: 1084-1093.
26. Zinzen RP, Girardot C, Gagneur J, Braun M, Furlong EE (2009) Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* 462: 65-70.
27. Chavez JD, Hoopmann MR, Weisbrod CR, Takara K, Bruce JE (2011) Quantitative proteomic and interaction network analysis of cisplatin resistance in HeLa cells. *PLoS One* 6: e19892.
28. Key M (2012) A tutorial in displaying mass spectrometry-based proteomic data using heat maps. *BMC Bioinformatics* 13 Suppl 16: S10.
29. Zhang Q, Cundiff JK, Maria SD, McMahon RJ, Woo JG, et al. (2013) Quantitative Analysis of the Human Milk Whey Proteome Reveals Developing Milk and Mammary-Gland Functions across the First Year of Lactation. *Proteomes* 1:128-158.
30. Avenir R, Geiger T, Elroy-Stein O (2014) Genome-Wide Identification and Quantification of Protein Synthesis in Cultured Cells and Whole Tissues by Puromycin-Associated Nascent Chain Proteomics (PUNCH-P). *Nature Protocols* 9: 751-760.
31. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, et al. (2015) Tissue-Based Map of the Human Proteome. *Science* 347.
32. Wilkinson L (2005) *The Grammar of Graphics*, New York: Springer.
33. Falconer K (1990) *Fractal Geometry: Mathematical Foundations and Applications*. New York: John Wiley and Sons.
34. Mandelbrot B (1993) A Fractal's Lacunarity, and how it can be Tuned and Measured. *Fractals in biology and medicine* 8-21.
35. Plotnick RE, Gardner RH, O'Neill RV (1993) Lacunarity Indices as Measures of Landscape Texture. *Landscape Ecology* 8: 201-211.
36. Weisstein EW (2009) Board Mathworld-A Wolfram Web Resource.
37. Borys P (2009) On the Relation between Lacunarity and Fractal Dimension. *Acta Physica Polonica* 40: 1485-1490.
38. Ott E (1994) *Chaos in Dynamical Systems*.
39. Pappas DJ, Coppola G, Gabatto PA, Gao F, Geschwind DH, et al. (2009) Longitudinal system-based analysis of transcriptional responses to Type I Interferons. *Physiol Genomics* 38: 362-371.
40. Dougherty G, Henebry GM (2002) Lacunarity analysis of spatial pattern in CT images of vertebral trabecular bone for assessing osteoporosis. *Med Eng Phys* 24: 129-138.
41. Saunders SC, Chen J, Drummer TD, Gustafson EJ, Broszofski KD (2005) Identifying Scales of Pattern in Ecological Data: A Comparison of Lacunarity, Spectral and Wavelet Analyses. *Ecological Complexity* 2: 87-105.
42. Yasar F, Akgünlü F (2005) Fractal dimension and lacunarity analysis of dental radiographs. *Dentomaxillofac Radiol* 34: 261-267.
43. Myint SE, Mesev V, Lam N (2006) Urban Textural Analysis from Remote Sensor Data: Lacunarity Measurements Based on the Differential Box Counting Method. *Geographical Analysis* 38: 371-390.
44. Chen X, Li BL, Zhang ZS (2008) Using Spatial Analysis to Monitor Tree Diversity at a Large Scale: A Case Study in Northeast China Transect. *Journal of Plant Ecology* 1: 137-141.
45. Karemore G, Nielsen M (2009) Fractal Dimension and Lacunarity Analysis of Mammographic Patterns in Assessing Breast Cancer Risk Related to HRT Treated Population: A Longitudinal and Cross-Sectional Study. *Medical Imaging 2009: Computer-Aided Diagnosis*.