



**HAL**  
open science

## Bristle Maps: A Multivariate Abstraction Technique for Geovisualization

Sungye Kim, Ross Maciejewski, Abish Malik, Yun Jang, David S. Ebert,  
Tobias Isenberg

► **To cite this version:**

Sungye Kim, Ross Maciejewski, Abish Malik, Yun Jang, David S. Ebert, et al.. Bristle Maps: A Multivariate Abstraction Technique for Geovisualization. IEEE Transactions on Visualization and Computer Graphics, 2013, 19 (9), pp.1438-1454. 10.1109/TVCG.2013.66 . hal-00796336

**HAL Id: hal-00796336**

**<https://inria.hal.science/hal-00796336v1>**

Submitted on 29 Aug 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Bristle Maps: A Multivariate Abstraction Technique for Geovisualization

SungYe Kim, Ross Maciejewski, *Member, IEEE*, Abish Malik, Yun Jang, *Member, IEEE*, David S. Ebert, *Fellow, IEEE*, and Tobias Isenberg, *Member, IEEE*

**Abstract**—We present Bristle Maps, a novel method for the aggregation, abstraction, and stylization of spatio-temporal data that enables multi-attribute visualization, exploration, and analysis. This visualization technique supports the display of multi-dimensional data by providing users with a multi-parameter encoding scheme within a single visual encoding paradigm. Given a set of geographically located spatio-temporal events, we approximate the data as a continuous function using kernel density estimation. The density estimation encodes the probability that an event will occur within the space over a given temporal aggregation. These probability values, for one or more set of events, are then encoded into a bristle map. A bristle map consists of a series of straight lines that extend from, and are connected to, linear map elements such as roads, train, subway lines, etc. These lines vary in length, density, color, orientation, and transparency—creating the multivariate attribute encoding scheme where event magnitude, change, and uncertainty can be mapped as various bristle parameters. This approach increases the amount of information displayed in a single plot and allows for unique designs for various information schemes. We show the application of our bristle map encoding scheme using categorical spatio-temporal police reports. Our examples demonstrate the use of our technique for visualizing data magnitude, variable comparisons, and a variety of multivariate attribute combinations. To evaluate the effectiveness of our bristle map, we have conducted quantitative and qualitative evaluations in which we compare our bristle map to conventional geovisualization techniques. Our results show that bristle maps are competitive in completion time and accuracy of tasks with various levels of complexity.

**Index Terms**—Data transformation and representation, data abstraction, illustrative visualization, geovisualization.

## 1 INTRODUCTION

As data dimensionality increases, the encoding of variables and their relationships is often abstracted down to a representative subset for analysis in a single display, or dispersed across a series of coordinated multiple views [1–3]. Moreover, many techniques have been developed to visually encode multiple data attributes/variables for each data sample to enable interactive analysis, ranging from discrete glyph attribute encoding [4] to more spatially continuous color, transparency, and shading encodings [5–7]. As the number of visualized variables increases, the amount of information that can be effectively displayed becomes limited due to over-plotting and cluttering [8]. This is especially a problem in geographical visualization as a key attribute of the data is the location within the two-dimensional map space.

In geographical visualization, data can be described at any given location on a map. The data being described can come from an aggregated measurement, a direct event occurrence, or various other means. In dense data sets, plot-

ting events as symbols on the map (e. g., Fig. 1(a)) leads to cluttering and is often unable to convey a meaningful sense of event magnitude within the data. Aggregation of the data by defined boundaries, such as county or census tract boundaries (e. g., Fig. 1(b)), leads to a loss of specificity in data location and runs afoul of the Modifiable Areal Unit Problem [9]. Furthermore, it is known that the level of data aggregation can affect aspects of task complexity such as information load and the user’s ability to recognize patterns within the data [10]. In order to combat problems associated with areal aggregation, dasymetric mapping focuses on using zonal boundaries that are based on sharp changes in the statistical surface being mapped [11]. However, even when grouping data into small spatial quadrats, data can either be over-aggregated or under-aggregated. A third option is to estimate the discrete event points as a continuous function (e. g., Fig. 1(c)); such a mapping, however, only allows for the use of color as a means of representing data variables. As an encoding based on underlying network data, Fig. 1(d) shows a traditional line map. However, its representation is still restrained by the color and thickness of the lines.

In order to increase the amount of information that can be visualized within the constraints of a thematic map, this paper explores a novel method of multivariate encoding. Inspired by ideas of symbolic encoding from Spence [12] and choices of visual encodings by Wilkinson [13], we have developed the bristle map (Fig. 1(e)), a novel method for the aggregation, abstraction, and stylization of geographically located spatio-temporal data. The bristle map consists of a series of straight lines extended from and connected to

- S. Kim, A. Malik and D. S. Ebert are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. E-mail: {inside|amalik|ebertd}@purdue.edu.
- R. Maciejewski is with Arizona State University, Tempe, AZ, USA. E-mail: rmacieje@asu.edu.
- Y. Jang is with Sejong University, Seoul, South Korea. E-mail: jangy@sejong.edu.
- T. Isenberg is with INRIA, Saclay, France. E-mail: tobias.isenberg@inria.fr.

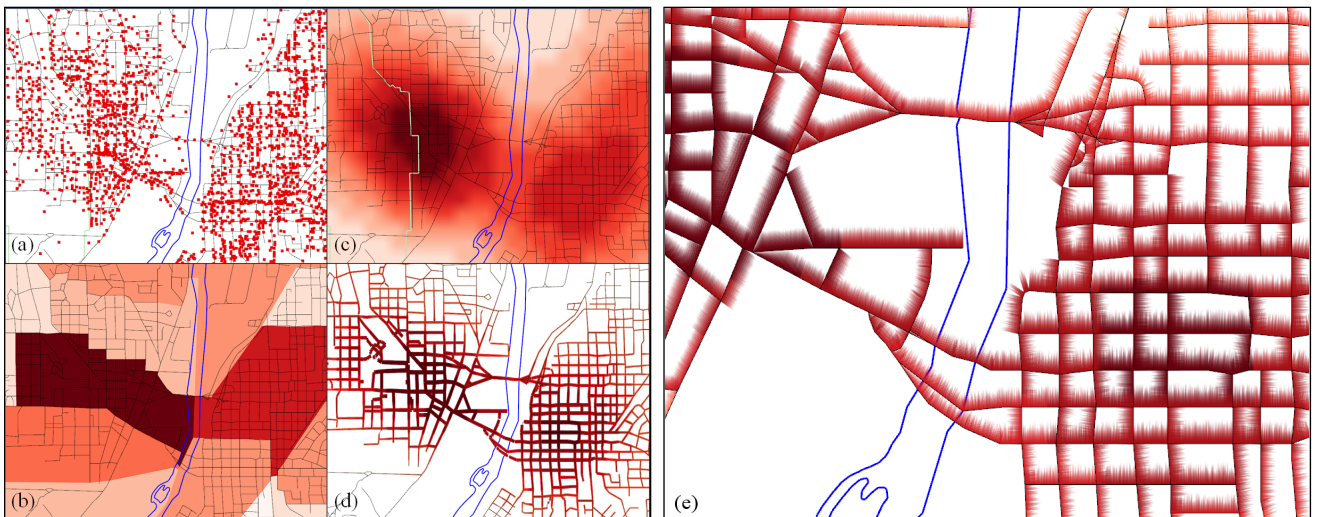


Fig. 1: Data abstraction in geovisualization. In this image, we show crimes in West Lafayette and Lafayette, Indiana where the blue line represents the Wabash River. (a) Plotting events as points. (b) Aggregation of points by areal units. (c) Approximation of a continuous domain from point sampling. (d) Approximation of a continuous domain using solid lines applied to roads. (e) Our abstraction using a series of bristle lines applied to roads.

linear map elements (roads, train lines, subway lines, etc.) that have some contextual relationship with the data being visualized. We vary these lines with respect to their color, length, density, and orientation to allow for a unique encoding scheme that can be used to create informative maps. With respect to the other representations shown in Fig. 1, our technique utilizes the underlying geographical context as a part of its symbology, thereby directly incorporating geographical elements within its encoding scheme. One of the major advantages of the bristle map technique is that the basis domain of the data (e. g., street network) remains highly visible regardless of the color scale being used. If one compares Fig. 1(c) and (e), the street network in Fig. 1(e) is clearly visible because the lines only ‘bristle off’ to one side, whereas in Fig. 1(c) some streets are hardly discernible due to the dark colors.

To demonstrate our technique, we focus on categorical spatio-temporal event data (e. g., emergency department logs, crime reports). In such data, events consist of locations in time and space where each event fits into a hierarchical categorization structure. These categories are typically processed as time series and snapshots of time are aggregated and typically visualized on a choropleth map [14]. Past work [6, 15] has shown that the use of kernel density estimation [16] is highly suitable in the spatial analysis of such data. Thus, our approach incorporates kernel density estimation as a means of estimating the underlying distribution of spatio-temporal events. Using the estimated distribution in an area for a given category (or categories) and temporal unit, we incorporate the underlying geographical network structure into the visual encoding. Bristles are extended from this underlying structure, and the color, length, density, transparency, and orientation of each bristle is mapped to a particular variable (or set of variables). Schemes presented in this paper include combinations of the following mappings:

- length, density, and color as data magnitude,
- orientation and coloring for bivariate mapping,
- color and length for bivariate mapping,
- color and density for bivariate mapping, and
- length and transparency for temporal variance.

Given the available parameters for visual encoding within the bristle map, other encodings also exist, which illustrate the flexibility and power of our technique. Our work focuses on showing how bristle maps can be used to show spatial and temporal correlations between variables, encode uncertainty in a unique way, and maintain geographical context through linking our visual encoding directly to geographical components. As such, the bristle map is a powerful multivariate encoding scheme that is adaptable to various attribute encodings to create richly informative visualizations.

## 2 RELATED WORK

Many techniques in multivariate data visualization focus on a means of reducing clutter and highlighting information through a variety of approaches including filtering (e. g., [17]), clustering (e. g., [18]), and sampling (e. g., [19]). In this section, we focus particularly on techniques within geographical visualization for improving the understanding of thematic/statistical maps, as Wilkinson [13] noted that the problem of multivariate thematic symbology for maps is that they are not only challenging to make, but also challenging to read.

In geographical visualization, the most common means of data representation is the choropleth map in which areas are shaded or patterned in proportion to a measured variable. Such maps are typically used to display only one variable, which is mapped to a given color scale. Other research has focused on encoding multivariate information into choropleth maps (such as uncertainty) with textures

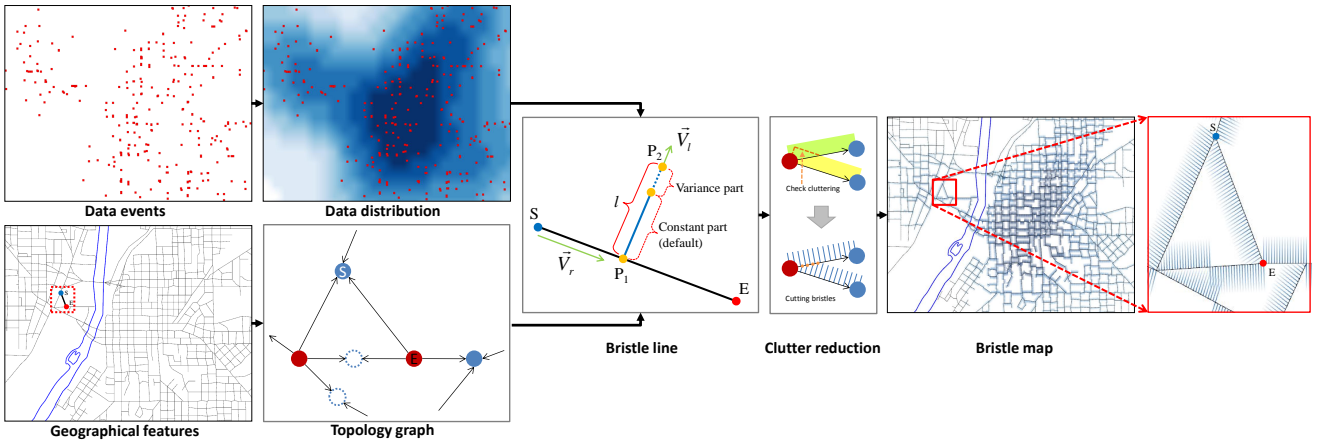


Fig. 2: The bristle map generation pipeline. Beginning with data events, a continuous abstraction is created. We also create a topology graph from contextually important linear features (in this case roads). Next, bristles are extended from these features based on the continuous abstraction and the topology. Clutter reduction is performed when generating each bristle, and finally the resultant bristle map is generated.

and patterns [20], creating bivariate color schemes for visualizing interactions between two variables [21, 22], or animating choropleth maps to enhance the exploration of temporal patterns and changes [23]. We present bristle maps as a robust alternative to these schemes in which multivariate attributes are instead mapped to a variety of graphical properties of a line (length, density, color and orientation), as opposed to utilizing a bivariate color scheme, texture overlays, or animation.

More recent geographical visualization techniques have included extensions to choropleth mapping ideas. Hagh-Shenas et al. [24] compared the effectiveness of visualizing geographically referenced data through the use color blending (in which a single composite color conveys the values of multiple color encoded quantities) and color weaving methods (in which colors of multiple variables are separately woven to form a fine grained texture pattern). The results from their study indicate color weaving to be more effective than color blending for conveying individual distributions in a multivariate setting. Saito et al. [25] proposed a two-tone pseudo coloring method for visualizing precise details in an overview display. Under this scheme, each scalar value is represented by two discrete colors. Sips et al. [26] focused on revealing clusters and other relationships between geo-spatial data points by their statistical values through the over-plotting of points. This work was later extended [27] to combine a cartogram-based layout to provide users with insight to the relative geo-spatial positioning of the dataset while preserving cluster information and avoiding over-plotting. Other cartogram techniques include the WorldMapper Project [28] which is used to represent social and economic data of the countries of the world. In each of these, novel data visualization techniques are created; however, the distortion of spatial features (country boundaries, roads) is often undesirable. While these techniques focus on displaying large amounts of aggregate data on small screens, our technique focuses on enhancing details of geographical context within the data. A similar concept of preserving data context is found

in Wong et al.’s [29] GreenGrid in which they visualize both the physics of the power grids in conjunction with the geographical relationships using graph based techniques.

Along with the previously described map schemes and cartogram distortions, there has been work in the use of heatmaps based on spatial data. Fisher [30] applied heatmaps to visualize the trends of the interactions of users with interactive maps that are based on their view of the geographic areas. Maciejewski et al. [6] used heatmaps as one of the tools to find aberrations or hotspots that facilitate the exploration of geo-spatial temporal datasets. Work by Chainey et al. [15] illustrated a number of different mapping techniques for identifying hotspots of crime and demonstrated that kernel density estimation provides analysts with an excellent means of predicting future criminal activities.

In conjunction with previous visualizations, other research has focused on expanding the dimensionality of the data being displayed by utilizing three-dimensional visuals. Van Wijk and Telea [7] utilized color and heightfields to visualize scalar functions of two variables. Tominski et al. [31] explored embedding 3D icons into a map display as a means of representing spatio-temporal data. In contrast, our work focuses on a two-dimensional encoding scheme that incorporates a variety of the visual variables described by Bertin [32] and Wilkinson [13] as a means of representing multivariate data.

Finally, it is important to note that our technique is akin to traditional traffic flow maps (e. g., Fig. 1(d)) seen in a variety of atlases; however, provides more generalized schemes. In traffic flow maps, the amount of data that can be displayed is restrained by the color and the width of the line representing linear elements (i. e., roads) on the map. Our work is similar to that of the traffic flow maps in that we utilize width (specifically, matched to the length in our bristle maps) and color as underlying visual variables of our encoding. However, our work also incorporates bristle density as a means of further encoding parameters. In the following sections, we compare our encodings to a variety of methods including the point, color, and flow line maps.

### 3 BRISTLE MAP GENERATION

In Fig. 1, we developed our motivation for the need to directly incorporate geographic features to the underlying data in order to better preserve contextual information. It is clear that the aggregation of data into arbitrary geographical areas obscures data, while the continuous approximation of an underlying data source can lead to incongruent mappings with respect to geographic features. Furthermore, both these mappings are limited in the fact that only color and texture are available for variable encoding, limiting the amount of data that can be displayed to either a single variable or possibly two variables in the case of a bivariate color map. The goal of this work is to create visual encodings for higher order structures.

The bristle map was inspired by the Substrate simulation of Tarbell [33] and abstract renderings of map scenes in work by Isenberg [34]. Given these images, our work focuses on using the underlying visual properties to intelligently encode information for display. In *The Grammar of Graphics* [13], Wilkinson discusses the combination of several perceptual scales into a single display. Here, he notes the idea of separable dimensions of the data is a key issue, where discriminations between stimuli are of key importance in the visualization. The Substrate aesthetic directly lends itself to this approach as color, line length, and orientation are distinct classes within Wilkinson’s table of aesthetic attributes and each of these visual parameters directly contributes to the substrate aesthetic.

Fig. 2 illustrates the bristle map generation pipeline. Given underlying data events, we compute a continuous distribution. We also create a topology graph from given geographically relevant linear content for clutter reduction described in Section 5. As an example of geographical content, if the underlying data was water pollution we could use a city sewage map for the geographic components, for our crime data examples we use roadways. Each linear geographic component consists of a series of line segments, and we extend *bristle lines* from these line segments. These bristle lines emerge perpendicularly from the underlying geographical line segment and are allowed to vary in length, density, color, transparency, and orientation, to facilitate multivariate data encoding. The third stage of the bristle map generation pipeline (Fig. 2) illustrates the bristle line concept for each geographical line segment,  $\overline{SE}$ , and  $\overline{P_1P_2}$  defines our generated bristle line. Each bristle line is created using the vector equation of a line as shown in Equation 1.

$$P_2 = P_1 + \vec{V}_1 L_l = P_1 + \vec{V}_1 (t \times L_{lmax}) \quad (1)$$

Here,  $P_1$  is a point on the contextually relevant geographic line segment,  $\overline{SE}$ ,  $\vec{V}_1$  is a unit vector perpendicular to the line  $\overline{SE}$ , and  $L_{lmax}$  is the maximum length of the  $\overline{P_1P_2}$ .  $L_l$  is the length of the  $\overline{P_1P_2}$  determined by a parameter  $t$ .

Each line from  $P_1$  to  $P_2$  is drawn in such a manner that it will either encode different properties of a multivariate data set, or use a data reinforcement technique where properties are encoded to the same variable to provide redundant cues. We utilize three encoding properties for each bristle; length,

TABLE 1: Parameters, corresponding variables, and ranges.

Parameters	Potential variables	Range
Base position ( $P_1$ )	Geographic location	(Double, Double)
Length 1 (constant portion)	Data magnitude	Double
Length 2 (variance portion)	Temporal variance, accuracy	e. g., monthly/yearly
Color	Data magnitude	Discrete, Continuous
Transparency	Temporal variance, accuracy	Double [0.0, 1.0]
Orientation ( $\vec{V}_1$ )	Temporal difference, data type	Clock-wise, Counter clock-wise
Density	Average data magnitude on an area ( $\overline{SE}$ )	Double

color, and orientation. The length of a line  $\overline{P_1P_2}$  is separated into two portions; a constant component, which is proportional to the magnitude of the variable being encoded, and a variance component. It captures temporal variance or other properties such as level of certainty. The color of a bristle  $\overline{P_1P_2}$  is proportional to the underlying variable distribution to be encoded at point  $P_1$ . When the variance component is used, its transparency is adjusted as a means of visually distinguishing it from the constant component. Orientation of the bristle line is always perpendicular to  $\overline{SE}$  and is utilized for bivariate comparison (i. e., day/night, two data types) and/or clutter reduction. To summarize, length and color represent a local data magnitude property at point  $P_1$ . We also choose to encode redundant information into the density of the number of bristles placed on a given line segment where the density of the bristles along  $\overline{SE}$  is decided by an average data value on a line segment  $\overline{SE}$ .

For each visual encoding, the underlying data is assumed to be continuous over a given geographical segment, such that for all points between any two nodes on the underlying contextual geographic structure, a data distribution value is associated with the point. In the case of a discrete data set (e. g., crime locations), the choice of an appropriate means of data interpolation with regards to the underlying geographic information is dependent on the data analysis being performed. Based on the recommendations of Chainey et al. [15], we apply a kernel density estimation (KDE) [16] to approximate the underlying distribution of crimes over the geographic features. The kernel density estimation procedure used is defined by the following equation:

$$\hat{f}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{h} K\left(\frac{\mathbf{x} - X_i}{h}\right) \quad (2)$$

Here, the window width of the kernel placed on point  $\mathbf{x}$  is proportional to a window bandwidth,  $h$ , and the total number of samples,  $N$ . We utilize the Epanechnikov kernel [16], Equation 3:

$$K(\mathbf{u}) = \frac{3}{4} (1 - \mathbf{u}^2) 1_{(\|\mathbf{u}\| \leq 1)} \quad (3)$$

where the function  $1_{(\|\mathbf{u}\| \leq 1)}$  evaluates to 1 if the inequality is true and zero for all other cases.

Thus, given a multivariate dataset where locations in space and time correspond to a series of categorized events,

we can create bristle maps that encode various properties of the data. Note that this technique relies on the data being contextually relevant to an underlying geographical network. For example, crime event data with its 2D geographical coordinates is recorded and hence defined by addresses on streets; thus it is contextually relevant to a street network. Data sets in which this contextual relationship does not exist should utilize other visual encoding schemes. Table 1 shows the parameters in our bristle map and their corresponding potential variables being encoded to each parameter. In the following section, we present a series of potential parameter combinations for various bristle map encodings and discuss the various results.

## 4 ENCODING SCHEMES

The bristle map is a powerful visual encoding scheme that lends itself to a variety of data encodings, examples of which we present next. For demonstration purposes we employ categorical spatio-temporal police reports collected in Tippecanoe County (specifically West Lafayette and Lafayette, IN, USA), from 1999 to 2010. The data set contains the date, time, crime type (e.g., armed/unarmed aggravated assault, armed robbery, burglary, homicide, noise, other assaults, rape, rape attempted, residential entry, robbery, theft, vandalism, and vehicle theft), and the address of each recorded criminal event. Note that other datasets can be easily encoded with bristle maps, and our choice of data was only made to illustrate the technique.

Utilizing this multivariate crime data set we discuss potential encoding schemes for multivariate spatio-temporal data. We then provide illustrations of each described encoding scheme with respect to our crime data set. Encoding schemes presented in this section include the use of bristle color, length, and density to encode data magnitude, the use of bristle orientation to inform temporal comparison, and the encoding of temporal variance in the bristle lengths.

### 4.1 Color, Length, and Density as Data Magnitude

Here, we discuss our technique for encoding the color, length, and density of the bristles into two separate variable groups. As both color and length (size) fall into two distinct categories of aesthetics according to Wilkinson [13], the use of separate variables for both categories allows for a distinguishable visual data encoding. In both cases, we assign data magnitude to both a color scale and a length scale. We note that such an encoding scheme has the potential to portray data more effectively than visualizations that map each data variable to a single display parameter. As noted in the arguments for the use of redundant color scales by Rheingans [35], the use of different display parameters is able to convey different types of information. Furthermore, by combining encodings in a redundant manner, it is possible to reinforce the encoding scheme. The utility of redundant color scales was confirmed by Ware [36].

In our encoding scheme, each bristle line's length,  $L_l$ , is calculated using Equation 4 based on a parameter,  $t$ , and

the maximum length,  $L_{lmax}$ .

$$L_l = t \times L_{lmax} = (\alpha \times \kappa_{P_1} + \beta \times \nu_{P_1}) \times L_{lmax} \quad (4)$$

For this visual encoding of the bristles, the parameter  $t$  is defined by the ratio of the data value at  $P_1$ , which we call  $\kappa_{P_1}$ , the ratio of the temporal variance at  $P_1$ ,  $\nu_{P_1}$ , and a set of tuning parameters ( $\alpha$  and  $\beta$ ) which provide weights to the constant and variance components as shown in Fig. 2. In this work, we use  $\alpha=1.0$  and  $\beta=0.3$ . Note that the choice of encoding the variance at a 30% value was chosen through trial and error by generating visualizations that the authors found to be the most useful and aesthetically pleasing. For problems where determining exact data values from the visual encoding is required (as opposed to approximating high and low rates), the variance portion is removed from the equation entirely by using  $\beta=0.0$ . As such, by creating the encoding scheme with diverse parameters, we are able to generate more aesthetic choices and visualizations. It is important to note that not all encodings will be appropriate and are most likely task dependent.

The  $L_{lmax}$  portion of Equation 4 is defined in Equation 5.

$$L_{lmax} = \rho \times \log_b \left( \frac{1}{N_r} \sum_{i=0}^{N_r-1} L_{SE} \right) \quad (5)$$

In this equation, we take the average length of all line segments (where  $N_r$  is the total number of line segments in the map) and calculate  $L_{lmax}$  using a non-linear function such that the length of bristle lines does not grow in an unbounded manner when zooming in. Moreover,  $L_{lmax}$  is modified by the parameter  $\rho$ , where  $\rho$  is the ratio of the current zoom level to the initial zoom level, to decouple our technique with the zoom level. In this work, we use  $b=15$  for the base of a log function.

Next, we determine the number (or density) of bristles,  $N_l$ , to be drawn on each line segment  $\overline{SE}$  using Equation 6.

$$N_l = \rho \left( \frac{\zeta}{\lambda \overline{L_{SE}}} \right) \kappa_{\overline{SE}} \quad (6)$$

Here,  $N_l$  is calculated using two user-defined constants  $\lambda$  and  $\zeta$ , where  $\lambda$  is the unit geographical length (distance) and  $\zeta$  is the number of bristle lines per unit geographical length. We use  $\lambda=0.0009$  and  $\zeta=3-15$  in our current visualization. As the bristle density may also be used to encode data magnitude parameters in bristle map generation,  $N_l$  should be proportional to the ratio of average data value on  $\overline{SE}$ ,  $\kappa_{\overline{SE}}$ . Moreover, we also apply  $\rho$  such that  $N_l$  will be independent of the zoom level to preserve the extent of density.

For color, we allow users to choose either a continuous or a sequential color scheme from Color Brewer [37]. Then, data is linearly mapped to a probability that a crime of type A will occur at geographic point B, where the probability is estimated from the underlying data distribution using kernel density estimation as described in Section 3.

Fig. 3 illustrates our length, density, and color encoding using the previously described crime data set. Burglary is

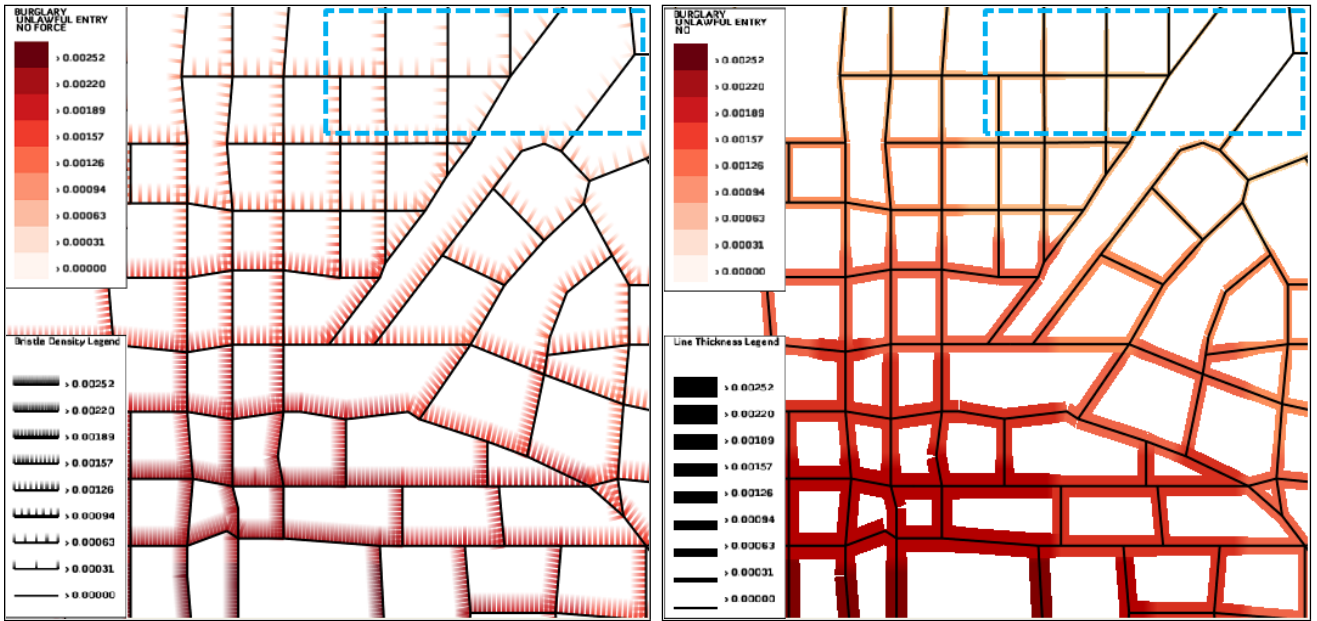


Fig. 3: (Left) Our bristle map encodes burglary rates with both bristle color and bristle density. (Right) A line map encoding burglary rates as both line color and line thickness. Compared to the line map, our bristle map provides a distinguishable visualization by incorporating bristle density. For example, bristle lines on the right top area are easily identified, whereas thickness in the line map on the same area is too small to clearly be perceived.

encoded with the red color scheme, and color is proportional to the probability (calculated from the underlying point distribution using kernel density estimation) that a burglary occurred at a given location. Fig. 3 (left) shows our bristle map encoding for burglary rates with a color scheme and bristle density, and Fig. 3 (right) shows a line map encoding the same information with a color scheme and line thickness for comparison to our bristle map. Compared to this line map, our bristle map provides the advantages of additional dimensionality through the density of bristle lines. In this scheme, one is able to easily encode two variables in different combinations of bristle map parameters (i.e., color and density with a constant length, color and length with a constant density), and provide users with distinguishable visual parameters that seem to focus attention to various details.

#### 4.2 Multivariate Encoding: Separating Length, Density and Color, and using Orientation

In the previous section, we illustrated how our method can be utilized for univariate encoding by using a redundant encoding scheme. However, a major benefit of bristle maps is the ability to encode multivariate attributes. One example of this is seen in day versus night time comparison.

Here, one can utilize the orientation to separate two temporal components of a single variable by mapping the temporal components to different orientations of the geographic feature. For instance, it is likely that the rates of data variable will be different with respect to day and night occurrences. We illustrate this visual encoding in Fig. 4. We separate the events into day (6:00 am–6:00 pm) and night (6:01 pm–5:59 am) and map the daytime rates

to red and one orientation, and nighttime rates to blue and the other orientation. In Fig. 4 (right) we illustrate a bristle map encoding of one variable (burglary) during 2009 where length, density and color represent the magnitude of the burglary as well as the encoding of day and night parameters is explored as line orientation.

In Fig. 4 (right), we show areas of high/low nighttime crime, high/low daytime crime, and combinations there within. In contrast, a traditional heatmap using a univariate color scheme can only show either daytime crime (Fig. 4 (left top)), or nighttime crime (Fig. 4 (left middle)). Hence, several heatmaps are needed to see day and night variations as shown in Fig. 4 (left column). Viewers must mentally combine the images to locate regions of the map that have high crime levels at daytime and nighttime, thereby increasing their cognitive load.

Another means of reducing the cognitive burden would be to create a heatmap of the difference between night and day. Fig. 4 (left bottom) shows the difference of day and night data, and the divergent color scheme shows where high daytime or high nighttime crimes occur. For instance, in Fig. 4 (left bottom) the right area indicates higher rates during day, the left area shows higher rates during night, and the border area between the blue and red color schemes only indicates that day and night rates were approximately equal, regardless of them being low or high. Moreover, you need other color maps to explore areas where one occurs similarly high or low during day and night time.

Bristle map encodings have benefits in this situation. When we explore a daytime versus nighttime bristle map in Fig. 4 (right), we see that there exists distinct temporal profiles along the road lines where we see exclusively dominant areas during either day or night. For instance,

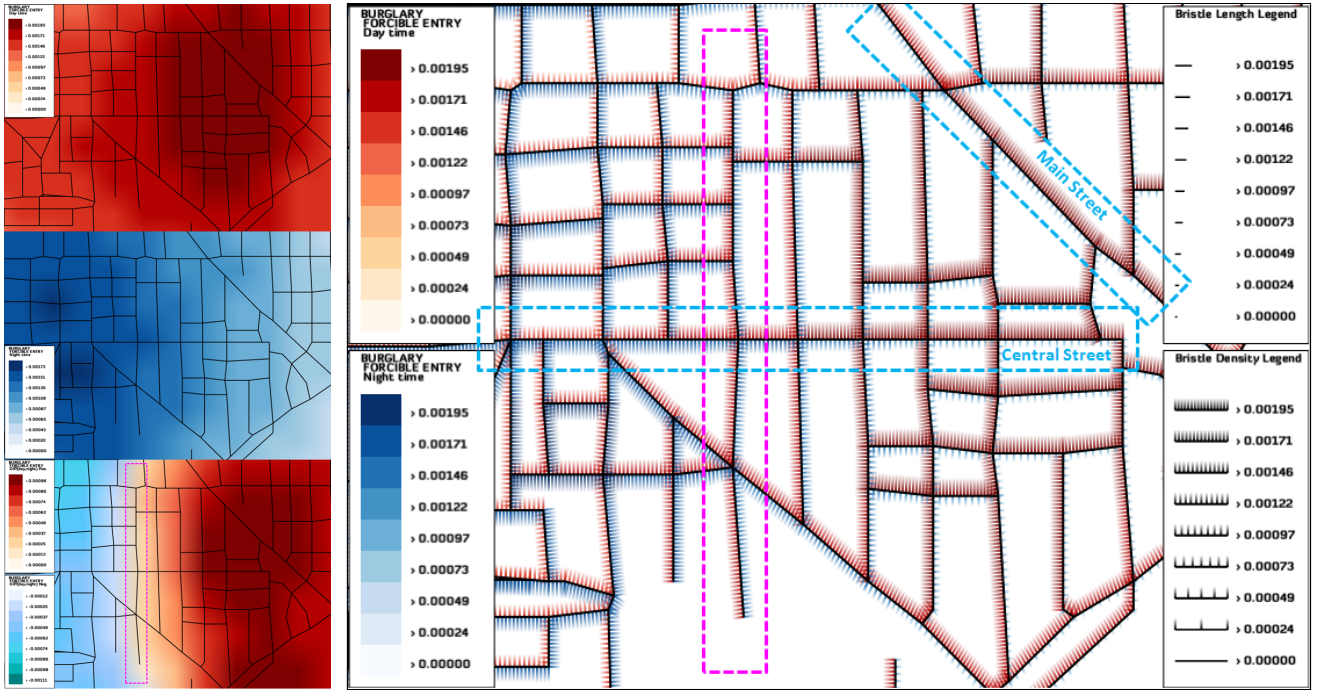


Fig. 4: Encoding daytime versus nighttime variations. (Left column) From top to bottom, color maps showing day, night and the difference of day and night burglary rates. (Right) Our bristle map separating the burglary rates into their day and night components with opposite orientation along roads. Note that a color map cannot present two components (i. e., both daytime and nighttime burglary rates) at the same location, hence three color maps are needed to see day and night variations simultaneously. Our bristle map can present such information within one bristle map by using different orientations of bristle lines.

see the diagonal road from the top center to the right center (Main Street, Lafayette, IN) showing that daytime burglary dominates along this road. Another observation is made on the horizontal road at the center of the map (Central Street, Lafayette, IN). Along this road, daytime burglary rates increase from left (west) to right (east), whereas nighttime burglary rates decrease from left to right. For the center area in Fig. 4 (left bottom), where the blue and red color schemes meet, we also see in Fig. 4 (right) that it has relatively equally high rates during both day and night. Such a comparison allows people to understand the differences between the data; however, when subtracting, areas of nearly equal daytime and nighttime crimes will be colored the same. Thus, areas that are safe during both day and night, and areas that are highly dangerous during both day and night will appear the same in the difference color map. In contrast, bristle maps allow viewers to quickly observe trends related to both day and night.

Another example of multivariate encoding using our bristle map is done by separating and/or combining bristle parameters. For instance, bristle density (or length) encodes a variable A, and color encodes a variable B while being presented on one orientation. Similarly, another two variables (C and D) could be encoded and presented on the other orientation. However, this type of parameter combination should be determined carefully so as not to increase viewers' cognitive load. Its effectiveness would depend on several factors such as data type and analysis

purpose. In Section 6, we conduct experiments to explore the effectiveness of different parameter combinations.

### 4.3 Encoding Data Variance

As introduced in Fig. 2, each bristle can include a portion generated for temporal variance of data, see Equation 4. To present the temporal variance of the data over time, we compute both the monthly and yearly mean and variance values. For a given discrete data set during time periods  $N_T$ , we first calculate continuous distributions over time. Then, we determine mean and standard deviation values with respect to the underlying data distribution for the entire data set over a given temporal aggregation. Thus, we calculate the mean  $\mu$  and variance  $\sigma$  values from time varying data  $K_i$ , where  $i \in [0, N_T - 1]$ . Note that  $\mu$  and  $\sigma$  are computed only once as they represent constant values for a given dataset. Mean and variance values for each grid point  $j$  are calculated using Equation 7 and 8, respectively. Variance is then used to weight the parameter  $\beta$  in Equation 4 such that given the data magnitude at the current time  $K_{cur}$ , we compute the ratio of variance at the current time,  $\tilde{\sigma}$  as shown in Equation 9. As such, the parameter  $t$  in Equation 4 can be detailed as shown in Equation 10 to represent the length of bristle lines with respect to temporal variance.

$$\mu[j] = \frac{1}{N_T} \sum_{i=0}^{N_T-1} K_i[j] \quad (7)$$



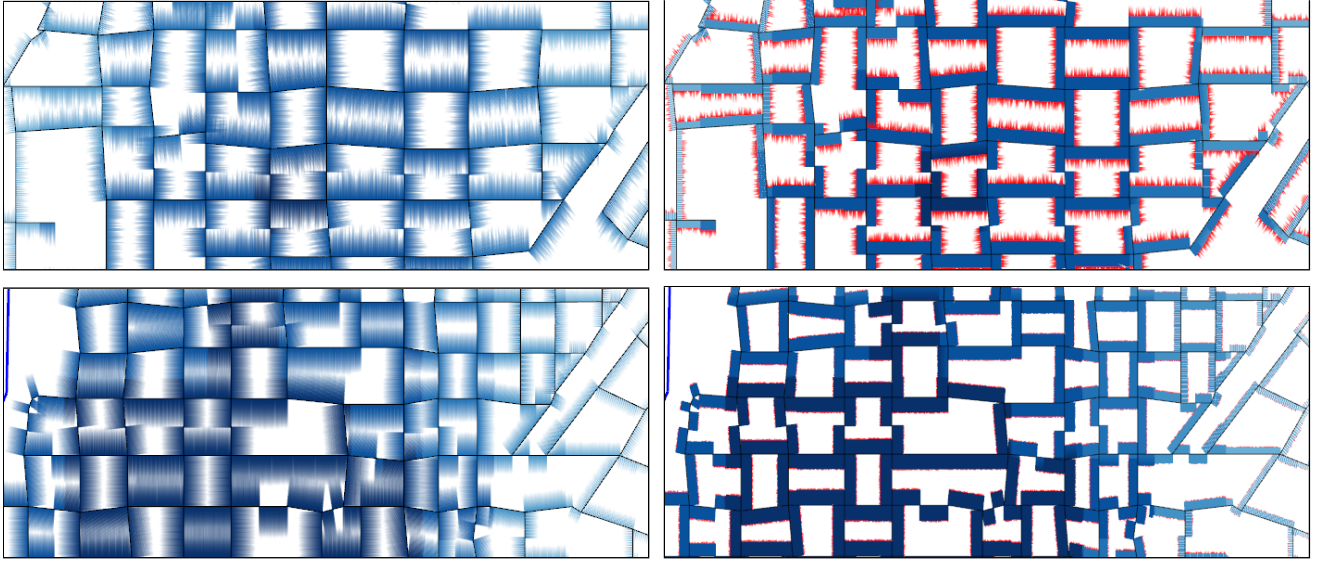


Fig. 5: Encoding data variance of vandalism-graffiti in Lafayette, IN, USA, in 2010 creating an uncertainty aesthetic. Yearly variance of vandalism-graffiti is represented in (a) a residential area and (c) a commercial area without distinguishing the variance component in the bristle length. (b) and (d) show the results using a highlight color for the variance portion and full alpha values for the constant portion of bristle lines. Here, we clearly see that our bristle map can encode the temporal variance and create an uncertainty aesthetic using the variance component.

$$\sigma[j] = \sqrt{\frac{1}{N_T} \sum_{i=0}^{N_T-1} (\mu[j] - K_i[j])^2} \quad (8)$$

$$\tilde{\sigma}[j] = \frac{1}{\sigma[j]} |\mu[j] - K_{cur}[j]| \quad (9)$$

$$t = \alpha \times \kappa_{P_1} + \beta \times \nu_{P_1} = \alpha \times \kappa_{P_1} + \beta \times \left( \frac{\tilde{\sigma}_{P_1}}{\tilde{\sigma}_{max}} \right) \quad (10)$$

Furthermore, the variance term,  $\nu_{P_1}$ , in parameter  $t$  in Equation 4 can also be revised to encode an uncertainty factor by using randomness. We may also encode an uncertainty factor by using color and transparency to enhance the variance component. When using color and transparency, we use a highlight color for the variance component, and then fade out the variance component over the bristle length with a full alpha value for one end point and an alpha value weighted by the variance for the other end point. The constant portion of the bristle is assigned an alpha value of 1 to both end points as it represents an exact data value. Hence, according to the data type and analysis purpose, the encoding of parameter  $t$  and the use of the variance portion can be different and should be assessed with respect to the visual message trying to be conveyed. Fig. 5 illustrates the application of encoding the data variance of vandalism with the uncertainty factor. In Fig. 5(a, c), we use the same color scheme for the constant and variance portions of bristle lines. To enhance the variance component in Fig. 5(b, d) we highlighted the variance portion in a different color and assigned full alpha values for the constant portion of bristle lines. Fig. 5(a, b) show the same area. In this area, the bristle length shows large fluctuations, indicating a high

yearly variance Fig. 5(c, d) show another area. In this area, the bristles are of a nearly constant length, indicating low yearly variance. When considering that the area in Fig. 5(a, b) includes residential areas, while the area in Fig. 5(c, d) includes the downtown Main street, an art theater, and the City Hall in Lafayette, IN, our bristle map shows that the residential areas have higher yearly variance of vandalism (graffiti) when compared to commercial areas.

## 5 BRISTLE CLUTTER REDUCTION

Although our bristle map can encode various characteristics from multivariate data, it often suffers from clutter around the intersections of road lines. In order to minimize cluttering, we employ two strategies in our bristle map generation pipeline (Fig. 2); 1) using topology among road lines to determine bristle orientation to minimize clutter and 2) cutting bristle lines crossing neighbor road lines.

### 5.1 Using Topology

Each bristle map contains an underlying topology of the contextual geographic network that the data is mapped to. In the topology graph, each node is defined as either ‘outward’ or ‘inward’ as illustrated in Fig. 6. Using the topology graph, we choose each segment’s bristle line orientation such that the overlap of the bristles at intersections will be minimized, thereby reducing the clutter. If the encoding scheme requires both sides of the edge to contain bristles, then clutter at each intersection is inevitable. However, in cases where bristles map to only one side of an edge, we use the right-hand rule to decide the orientation. Hence bristle lines on edges connected to neighboring outward and inward nodes are generated in a manner that provides a reasonable reduction in clutter (Fig. 6).

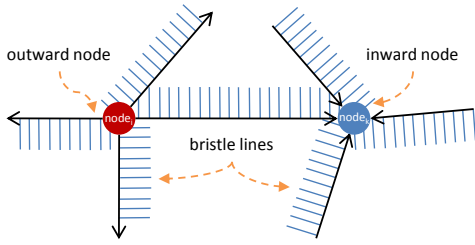


Fig. 6: To minimize clutter, a topology graph consisting of directed edges as road lines and outward (red) and inward (blue) nodes on the intersection of lines is used to decide the bristle line orientation.

Choosing the orientation of bristle lines in order to minimize overlap can be considered as a 2-coloring problem in vertex coloring; one color presents ‘outward’ while the other presents ‘inward.’ Vertex coloring is a well known graph problem, where no two adjacent nodes share the same color. Moreover, coloring a general graph with the minimum number of colors is known to be an NP-complete problem. In our case, the minimum number of colors should always be 2 but such 2-colorability is not guaranteed for general road lines. While deciding the orientation of bristle lines, we often have undesirable topology generating inevitable overlap of bristle lines. Fig. 7 (upper row) shows such a bad topology example and our strategy to solve this issue. In Fig. 7(a), we see two clutter areas caused by an undesirable configuration of neighbor nodes which guarantee bristle overlap. To solve this, we consider the addition of a virtual node in a topology graph as shown in Fig. 7(b), thereby allowing for an orientation switch midway across the edge and reducing the clutter. For neighboring two inward nodes (blue), we add a virtual outward node (red dotted circle) at the road line connecting two inward nodes resulting in splitting bristle lines on the road line. Similarly, a virtual inward node (blue dotted circle) is added for neighboring two outward nodes (red).

## 5.2 Avoid Crossing Neighbors

Another cluttering case is illustrated in Fig. 7(c). When two road lines intersecting with less than a  $90^\circ$  angle have bristle lines, some of the bristle lines overlap as illustrated in Fig. 7(c). For this case, we forbid bristle lines to cross neighbor road lines by placing the end point of a bristle line on the neighbor road line as shown in Fig. 7(d). We first check the intersection of bristle blocks (colored boxes in Fig. 7) for the current road line on which we are generating bristle lines and its neighboring road lines by using the topology graph. If the blocks are intersected, we then check if a bristle line crosses the neighbor road lines by utilizing the intersection algorithm of 2D line segments [38]. This idea is based on the theory of amodal completion (or amodal perception) [39] in psychology that describes how the human visual system completes parts of an object even when it is only partially visible. Although the length of a bristle line represents data magnitude, benefits from

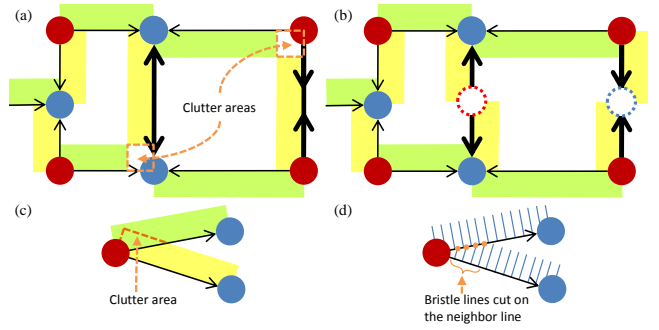


Fig. 7: Two pairs of the cluttering cases and our methods to minimize clutter. Colored box areas on a side of each edge line indicate the orientation for bristle lines. (a) Case 1: bad topology, where two inward nodes (blue) share a line and two outward nodes (red) share a line, generates inevitable clutter. (b) Virtual nodes (dotted circles) are added to split an edge line. (c) Case 2: a small angle between edge lines causes a clutter area. (d) Bristle lines crossing a neighbor edge line are cut on the neighbor line.

cutting the length to avoid clutter dominate the side effects from data misunderstanding that could be caused by clutter. Moreover, when using redundant encoding utilizing bristle length and density as data magnitude, bristle density could help viewers complete parts of the bristle lines. Fig. 8 shows four image pairs before and after applying our clutter reduction strategies. Some improvements could also be considered in the future. For instance, our strategies still generate cluttered bristle lines in cases where road lines are very dense or close to others. We perform experiments in Section 6 to see how people understand the differences before and after clutter reduction. Here we note that the experiments performed were for comparison and identification tasks. In these task types, line direction (as will be shown in the experiments) had little impact on the user results. However, in a cluster/delineate task in which users are asked to segment the data, the splitting of direction may influence the user’s perception of cluster boundaries. As such, we recommend that map designers take caution in employing this scheme and use it only in appropriate map contexts. Future work will explore other schemes and design issues to handle neighbor crossings and influence on map design.

## 6 EVALUATION

To evaluate the effectiveness of our bristle maps, we conducted two quantitative controlled experiments. These studies are both comprised of an introductory session, and a training session. In the first study, five tasks were conducted to evaluate the efficiency of bristle maps compared to existing visualization methods (point, color (kernel density estimated, KDE), and line maps as shown in Fig. 1(a), (c), and (d)) and post-task questionnaires for qualitative feedback. In the second study, two tasks were conducted to evaluate the accuracy of users in estimating values from each of the map types (point, KDE, bristle and line) as well as evaluating the perceived aesthetics of each image. Prior to each study, a pilot study was also conducted to

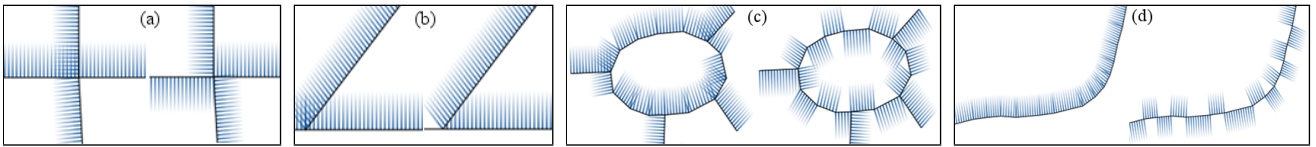


Fig. 8: Before/after image pairs of our clutter reduction. Each pair shows a case of (a) changing bristle orientation using topology, (b) cutting bristle lines crossing neighbor road lines, (c) circular roads, and (d) curved roads.

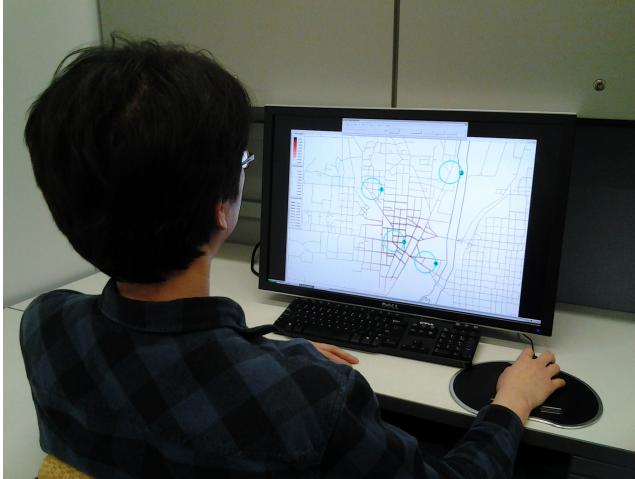


Fig. 9: Example setup for our experiment.

ensure that each task contains a fair comparison among the techniques.

**Participants:** In the first study, thirty graduate students (23 males, 7 females) in engineering, science, and statistics from our university participated in the study. All participants reported that they had experience in visualizing data on geographical maps using colors or icons (e.g., paper maps, online map services). The experience varied from almost daily (11 participants), 1–2 times a week (17 participants) to 1–3 times a month (2 participants). For the identification/accuracy tasks and aesthetic comparisons (Tasks 6 and 7), a secondary study was run on twenty-six undergraduate students in engineering from our university.

**Apparatus:** The experiment was performed on a 30" monitor using our experimental application running on Windows XP, as shown in Fig. 9, where all visualizations were generated with  $2228 \times 1478$  resolution. Each visualization was overlaid with numbered circles as shown in Fig. 9. Participants selected one of the numbers to answer the question in each trial using buttons in the interface panel on the top of the screen. Criminal incident reports collected in West Lafayette and Lafayette, Indiana, USA from 1999 to 2010 were used in each trial, but different types of crimes were selected to generate visualizations in the training phase and in the actual study.

**Design:** We employed a repeated measure design of tasks incorporating variations of the images shown in Fig. 1(a), (c), and (d) and line maps similar to those of Fig. 3 (right). Table 2 shows the number of data sets, techniques (cases as shown in Fig. 8 for Task 5), and trials in each task. For example, in Task 1 we utilized 18 different data sets to

compare 5 different techniques (i.e., point map, color map, line map, and bristle maps using two different encoding schemes). Hence, each participant performed  $18 \times 5 = 90$  trials in Task 1. In Task 3, we compared 6 different techniques (i.e., point map, bivariate color map, line maps in two different encoding schemes, and bristle maps in two different encoding schemes) with 15 data sets, resulting in 90 trials. Due to the difficulty of creating good examples to be used from our real crime data, we used fewer data sets in Tasks 4 and 5. In summary, each participant performed a total of 374 trials in Tasks 1 to 5, and it generally took 90 minutes.

Since the design of Tasks 1–5 focused on questions of comparing regions, a secondary study was also conducted. This study was again a repeated measure design of tasks incorporating variations of the images shown in Fig. 1(a), (c), and (d) and line maps similar to those of Fig. 3 (right). However, here the subjects were asked to identify the values of regions in the image. Areas of homogeneous visual variables were circled in each image and the subjects were asked to approximate the amount of crime per region. As a final task, the subjects were simultaneously presented with a point map, color map, bristle map and line map and asked to rank order the images based on their aesthetic values.

For all Tasks, trial order was varied using a magic square method [40] in each task. Completion time and participants' answers were recorded for a quantitative metric. The collected data from each task was subjected to an analysis of variance (ANOVA) test to determine if the average time and accuracy of task completion were significantly different among techniques. A post-hoc Tukey HSD test was then performed to determine significance between the techniques. P-values reported in this study come from the resultant Tukey HSD test. Before the study, participants were introduced to our experiment application and the techniques through an introductory session and a training session. During the training session, participants could ask questions and receive guidance in the use of the experiment application and analysis of each visualization. Once the training was completed, participants moved to the actual study. After completing each task (Tasks 1 to 4) participants were asked to answer the questionnaire to rate the efficiency of the techniques using a five-point Likert scale [41]. After completing Task 5, participants were also asked to describe their impression with regards to visual complexity for before and after image pairs applying our clutter reduction. In the questionnaire, we stated that the visual complexity is high if a participant felt any kind of difficulty or confusion in understanding the density, length

TABLE 2: The number of data sets, techniques (cases in Fig. 8 for Task 5), and trials.

	Data sets	Techniques (or cases)	Trials
Task 1	18	5	90
Task 2	18	5	90
Task 3	15	6	90
Task 4	12	4	48
Task 5	7	8	56
Task 6	2	4	24
Task 7	2	4	2

and color of bristle lines that encode the underlying data. Finally, after finishing all tasks, participants were asked to rate the overall efficiency among techniques.

**Hypotheses:** In this experiment, we hypothesized that our bristle maps would be better than or equally as good when compared to the other techniques in terms of task completion time and accuracy. Specifically, we hypothesized that our bristle maps would be better than other techniques as the complexity level of tasks increased from univariate to multivariate. The rationale of this assumption is that the line map and bivariate color map use at most two variables, whereas the several encoding parameters in our bristle map have the potential to create effective encoding combinations. We also hypothesized that our clutter reduction strategies would be useful to minimize cluttering on areas where a large number of bristle lines are created. In our follow-on experiment exploring identification of values, we hypothesized that bristle maps would be as accurate as all other representations in determining values. We also hypothesized that bristle maps would be ranked higher in terms of their aesthetics.

**Tasks:** We tested seven tasks: three for univariate, bivariate, and multivariate data encoding, respectively, one for temporal variance encoding, one for the clutter reduction, one for accuracy comparisons among the rendering styles and one for aesthetic comparisons.

In Task 1, when given four regions highlighted in circles on the map, participants were asked to “find the region with the highest crime rate” in different visualizations representing spatio-temporal crime data using point, color, line-T (data encoded in the line (T)hickness), bristle-CLD (a redundant data encoding using (C)olor, (L)ength, and (D)ensity), and bristle-LD (a redundant data encoding using (L)ength and (D)ensity).

In Task 2, four regions were highlighted in circles on the map. Participants were asked to “find the region with the highest crime rates at both (or either) day and night time,” using point (encoding day/night time crime rates in different colors), color, line-TO (data encoded as line (T)hickness and using (O)rientation for day/night crime rates), bristle-CLDO (redundant data encoding using (C)olor, (L)ength and (D)ensity, and using (O)rientation to indicate day/night crime rates), and bristle-LDO (data encoded using (L)ength and (D)ensity, but in a constant color, using (O)rientation to indicate day/night crime rates). The point map had differently colored points for day and night time crime rates, and two maps (day and night time color maps) were

given in different colors for the color map.

In Task 3, four regions were highlighted in circles on the map. Participants were asked to “find the region with the highest crime rates for both (or either) two crimes (crime 1 and 2),” using point map (encoding two crimes in different colors), bivariate color map (Color-B), line-TO (a data encoding using (T)hickness in different colors, and using (O)rientation to indicate crime types), line-CT (encoding crime 1 using (C)olor and crime 2 using (T)hickness), bristle-LDO (a redundant data encoding using (L)ength and (D)ensity, and using (O)rientation to indicate crime types), and bristle-CD (an encoding using (C)olor to indicate crime 1 and (D)ensity to indicate crime 2, with constant length).

In Task 4, participants were given two regions highlighted in circles on the map. Then, they were asked to “find the region with the highest temporal variance” in different visualizations using point maps, color maps, line maps, and bristle-LDV (a redundant data encoding using (L)ength and (D)ensity, and representing (V)ariance in the variance part of a bristle line). For the point, color, and line maps, multiple images were displayed on the screen to provide visualizations during several years. Our bristle map embedded the variance in the variance part of the bristle length as shown in Fig. 2 (third stage) and 5 (right column).

In Task 5, given two regions predefined in circles on bristle maps, participants were asked to “answer if crime rates on this given two regions look either different or the same as each other.” Fig. 8 shows representative image pairs before and after applying our clutter reduction method. In trials, participants compared each case in Fig. 8 to a base case (i. e., bristle lines on a single straight road).

In Task 6, subjects were presented with a series of images with a single predefined circle which covered an area consisting of homogeneous visual variables (i. e., identical color, bristle length, thickness, etc.). A univariate encoding was explored, and the Bristle-CLD settings were utilized for the bristle map. Participants were asked to estimate the amount of crime in the area using the provided scale (or scales in the case of bristle and line maps). Time and accuracy of the results were measured.

In Task 7, subjects were presented simultaneously with four images representing the same data set. These images consisted of a point map, a color map, a bristle map and a line map. Subjects were asked to rank order the images in order of most to least aesthetically pleasing.

## 7 RESULTS AND DISCUSSION

After all tasks were completed, times and answers collected during the study were analyzed using a single-factor ANOVA. A post-hoc Tukey HSD test was then performed to determine significance between the techniques. P-values reported in this study come from the resultant Tukey HSD test. For accuracy, the percentage of correct answers was computed.

**Task 1:** A one-way between-subjects ANOVA was conducted to compare the effect of different map visualizations on a subject’s time and accuracy in determining

TABLE 3: Tukey HSD results for Task 1.

	$p$ -value <	Point map	KDE map	Line-T
Time	Bristle-CLD	<b>.00001</b>	<b>.00001</b>	<b>.00042</b>
	Bristle-LD	<b>.00001</b>	<b>.00001</b>	<b>.01811</b>
	$p$ -value <	Point map	KDE map	Line-T
Accuracy	Bristle-CLD	<b>.00001</b>	.1554	<b>.01851</b>
	Bristle-LD	<b>.00001</b>	.3214	<b>.00602</b>

areas with highest crime rates within a given visualization. Conditions varied based on the given visualization, point maps, kernel density estimated color maps, line maps and bristle maps. There was a significant effect of visualization type on time at the  $p < .05$  level for the conditions [ $F(4, 145) = 35.366, p = .0000001$ ] and a significant effect of visualization type on accuracy at the  $p < .05$  level for the conditions [ $F(4, 145) = 3266.782, p = .000000006$ ]. Because statistically significant results were found, we computed a Tukey post-hoc test with results reported in Table 3. In Table 3  $p$ -values  $< .05$  indicate that groups were statistically different from one another.

The result showed that the bristle maps groups were both significantly different than the point, color and line maps in terms of speed (at the  $p < .05$  level). Specifically, the bristle map groups average times were 50.7 seconds and 56.6 seconds for the CLD and LD conditions respectively, which was slightly faster than the Line-T condition at 69 seconds and much faster than the point map condition at 102.6 seconds. However, the color map group was the fastest at 34.6 seconds.

For accuracy, the bristle maps groups were both significantly different than the point map group in terms of accuracy (at the  $p < .05$  level). Specifically, the bristle map groups accuracy ratings were 99.6% and 99.8% for the CLD and LD conditions respectively, which was much higher than the point map condition with accuracy of 41.4%. No accuracy differences were found when compared to the other groups. See Table 8 for more specific results.

The comparison between color maps and bristle maps showed that color maps were better than the bristle map in terms of average time, and were not significantly different in terms of accuracy. This shows that bristle maps as a redundant encoding scheme has the same potential to convey data as single parameter encoding schemes; however, traditional schemes such as color maps may allow for a quicker comparison in the univariate case.

Comparing Bristle-LD and Line-T, we saw that the length of the bristle map matches the thickness of the line map. Hence, the bristle density was useful to find answers in Task 1 in terms of completion time and accuracy. Some participants also mentioned bristle density in their qualitative feedback as “*Bristle map is especially good when density of the bristles is also used*” and “*In bristle map, length and density were more noticeable than color difference.*” In this univariate encoding test, the point map showed the worst results and the color map was the best results in terms of time and accuracy as shown in Table 8.

TABLE 4: Tukey HSD results for Task 2.

	$p$ -value <	Point map	KDE map	Line-TO
Time	Bristle-CLDO	<b>.01713</b>	.70091	.05943
	Bristle-LDO	<b>.02024</b>	.81621	.07166
	$p$ -value <	Point map	KDE map	Line-TO
Accuracy	Bristle-CLDO	<b>.00001</b>	.07062	.36692
	Bristle-LDO	<b>.00001</b>	.99999	<b>.01283</b>

**Task 2:** A one-way between-subjects ANOVA was conducted to compare the effect of different map visualizations on a subject’s time and accuracy in determining areas with highest crime rates at both day and nighttime within a given visualization. Conditions varied based on the given visualization, point maps, kernel density estimated color maps, line maps and bristle maps. There was a significant effect of visualization type on time at the  $p < .05$  level for the conditions [ $F(4, 145) = 2.717, p = .032$ ] and a significant effect of visualization type on accuracy at the  $p < .05$  level for the conditions [ $F(4, 145) = 89.89, p = .0000002$ ]. Because statistically significant results were found, we computed a Tukey post-hoc test with results reported in Table 4. In Table 4  $p$ -values  $< .05$  indicate that groups were statistically different from one another.

As we hypothesized, the result showed that the bristle maps groups were both significantly different than the point maps in terms of speed (at the  $p < .05$  level). Specifically, the bristle map groups average times were 86.3 seconds and 87.2 seconds for the CLDO and LDO conditions respectively, which was slightly faster than the point map condition at 106.2 seconds.

For accuracy, the bristle maps groups were both significantly different than the point map group in terms of accuracy (at the  $p < .05$  level). Specifically, the bristle map groups accuracy ratings were 90.5% and 93.3% for the CLDO and LDO conditions respectively, which was much higher than the point map condition with accuracy of 63.1%. See Table 8 for more specific results.

The comparison between color maps and bristle maps showed that color maps were better than the bristle map in terms of average time, and were not significantly different in terms of accuracy. This shows that bristle maps as a redundant encoding scheme has the same potential to convey data as single parameter encoding schemes; however, traditional schemes such as color maps may allow for a quicker comparison in the univariate case.

Findings also indicated that Bristle-LDO was better than Line-TO in terms of accuracy, whereas Bristle-CLDO was not significantly different from Line-TO in terms of accuracy. This indicated that the bristle density seems to be useful in finding correct answers in Bristle-LDO, but it was not in Bristle-CLDO. Further testing in combinations of visual variables and the ability to determine levels of sparseness will be done in the future.

**Task 3:** A one-way between-subjects ANOVA was conducted to compare the effect of different map visualizations on a subject’s time and accuracy in determining areas

TABLE 5: Tukey HSD results for Task 3.

	$p$ -value <	Point map	KDE-B	Line-TO	Line-CT
Time	Bristle-LDO	<b>.00009</b>	<b>.00515</b>	.58128	.73239
	Bristle-CD	<b>.01131</b>	<b>.03469</b>	.15506	.20693
Accuracy	Bristle-LDO	<b>.00001</b>	<b>.00001</b>	.27189	<b>.02771</b>
	Bristle-CD	<b>.00001</b>	<b>.00001</b>	.07002	.41194

with highest crime rates in two types of crimes within a given visualization. Conditions varied based on the given visualization, point maps, kernel density estimated color maps, line maps and bristle maps. There was a significant effect of visualization type on time at the  $p < .05$  level for the conditions [ $F(5, 174) = 6.655, p = .00001$ ] and a significant effect of visualization type on accuracy at the  $p < .05$  level for the conditions [ $F(5, 175) = 144.24, p = .00000001$ ]. Because statistically significant results were found, we computed a Tukey post-hoc test with results reported in Table 5. In Table 5  $p$ -values  $< .05$  indicate that groups were statistically different from one another.

The result showed that the bristle maps groups were both significantly different than the point maps and color maps in terms of speed (at the  $p < .05$  level). Specifically, the bristle map groups average times were 88.2 seconds and 94.5 seconds for the LDO and CD conditions respectively, which was faster than the point map condition at 118.3 seconds and the color map condition at 115.3 seconds.

For accuracy, the bristle maps groups were both significantly different than the point map group and the color map group in terms of accuracy (at the  $p < .05$  level). Specifically, the bristle map groups accuracy ratings were 94.4% and 90.4% for the LDO and CD conditions respectively, which was much higher than the point map condition with accuracy of 26.6% and the color map condition with accuracy of 72.6%. See Table 8 for more specific results.

Note that we separated parameters for different crime types in Bristle-CD; (C)olor encodes crime 1 and (D)ensity encodes crime 2. Bristle-CD showed a significant effect compared to the bivariate color map as shown in Table 5. However, generation on this type of bristle maps should be selected carefully since one parameter could dominate the other. For instance, when we use color and length to separate two crime data, short bristle length for low crime rates in crime 2 removes bristle lines in dark color for high crime rates in crime 1. In our experiment, we selected color and density for two crimes, with constant length of bristles.

**Task 4:** A one-way between-subjects ANOVA was conducted to compare the effect of different map visualizations on a subject’s time and accuracy in determining areas with high temporal variance within a given visualization. Conditions varied based on the given visualization, point maps, kernel density estimated color maps, line maps and bristle maps. There was a significant effect of visualization type on time at the  $p < .05$  level for the conditions [ $F(3, 116) = 42.051, p = .00001$ ] and a significant effect of visualization type on accuracy at the  $p < .05$  level for the

TABLE 6: Tukey HSD results for Task 4.

	$p$ -value <	Point maps	KDE maps	Line maps
Time	Bristle-LDV	<b>.00001</b>	<b>.00001</b>	<b>.00001</b>
Accuracy	Bristle-LDV	<b>.00001</b>	<b>.00001</b>	<b>.00001</b>

conditions [ $F(3, 116) = 42.33, p = .00001$ ]. Because statistically significant results were found, we computed a Tukey post-hoc test with results reported in Table 6. In Table 6  $p$ -values  $< .05$  indicate that groups were statistically different from one another.

The result showed that the bristle maps groups were both significantly different than the point maps, line maps and color maps in terms of speed (at the  $p < .05$  level). Specifically, the bristle map groups average time was 48.4 seconds for the LDV condition, which was faster than the point map condition at 194 seconds, the color map condition at 171.8 seconds, and the line map condition at 178.9 seconds.

For accuracy, the bristle maps groups were both significantly different than the point maps, line maps and color maps in terms of speed (at the  $p < .05$  level). Specifically, the bristle map groups accuracy rating was 94.7% for the LDV condition, which was much higher than the point map condition with accuracy of 53.6%, the color map condition with accuracy of 72.6% and the line map condition with accuracy of 75.5%. See Table 8 for more specific results.

As we hypothesized, we found that the representation of temporal variance in bristle maps was significantly faster and accurate in terms of both average time and accuracy compared to providing several images of the point, color and line maps. Moreover, we found that techniques showed the increasing pattern from the point maps to Bristle-LDV as shown in Table 8. This indicates that changes among several images would be better perceived in line patterns than in points or colors.

**Task 5:** A one-way between-subjects ANOVA was conducted to compare the effect of different map visualizations on a subject’s time and accuracy in determining areas with high temporal variance within a given visualization. Conditions varied based on the given visualization, point maps, kernel density estimated color maps, line maps and bristle maps. There was no significant effect of visualization type on time at the  $p < .05$  level for the conditions [ $F(1, 56) = .328, p = .569$ ] and no significant effect of visualization type on accuracy at the  $p < .05$  level for the conditions [ $F(1, 56) = .315, p = .315$ ]. In Task 5, we found that bristle lines with and without clutter reduction did not differ significantly w.r.t. both average time and accuracy for all cases (Fig. 8). This means that the base bristle lines and bristle lines before applying clutter reduction and the base and bristle lines after applying our clutter reduction are perceived similarly by participants. Moreover, when told that the bristle line orientation does not encode data, the opposite orientations of bristle lines on a single straight

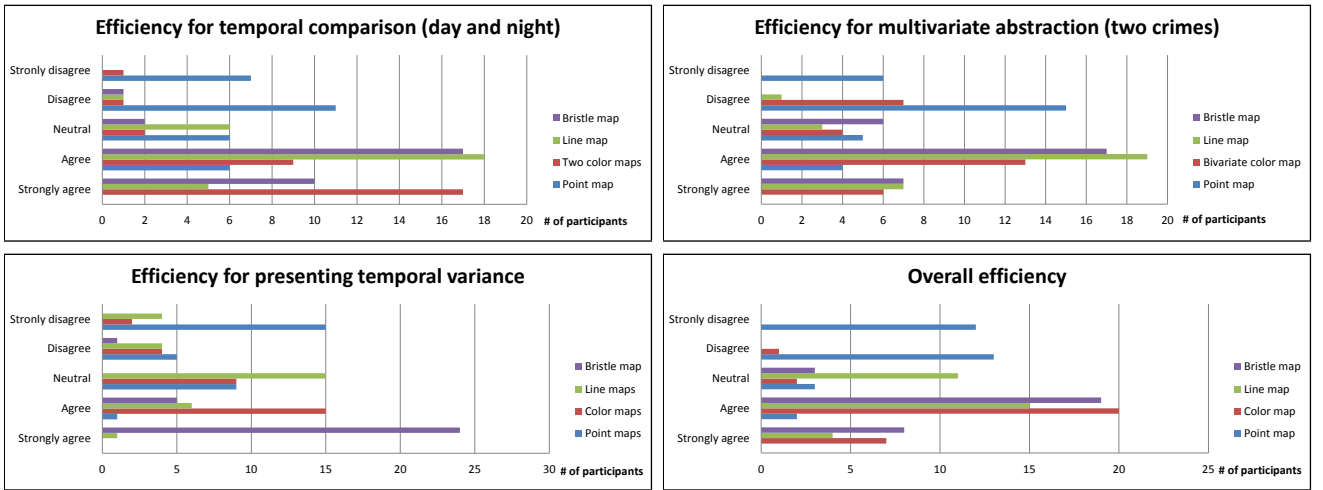


Fig. 10: Results from qualitative feedback for Tasks 2, 3, and 4 as well as overall efficiency.

TABLE 7: Average Rank Ordering by Aesthetics

	Point	KDE Map	Bristle	Line
Average	2.26	2.6	2.79	2.34
Std Dev	1.18	.97	1.19	1.10

road caused by virtual nodes (Fig. 7(b)) did not affect accuracy (87.7%). Other cases showed 42-58% of accuracy.

**Task 6:** For Task 6, we hypothesized that subjects would be as accurate as all other representations in determining values. In Task 6, we found that bristle maps did not differ significantly w.r.t. accuracy when compared with point map, color map and line map identification (ANOVA results of  $p$ -value=.18093,  $F=1.63$ ). However, we found that bristle maps did differ significantly w.r.t. time when compared with point map, color map and line map (ANOVA results of  $p$ -value=.0314,  $F=2.622$ ). Particularly, we found line maps and heat maps to both be significantly faster than point maps and bristle maps in identifying values (Tukey HSD test value of  $p < .05$ ). Overall, these results indicate that in terms of accuracy, all geographical representations were equally useful; however, participants were (on average) over 1 second quicker in value judgments on both line maps and colors maps. This is most likely due to the fact that participants were quicker at making color judgments as compared to counting points and mentally linking multiple variables for the bristle maps.

**Task 7:** In Task 7, we found that users had a highly variable rating of which image appeared to be more aesthetically pleasing. The average positions and standard deviations are summarized in Table 7. Here we find that while bristle maps have a slightly higher average ranking, there is no significant difference between the aesthetic ordering. A one-way between-subjects ANOVA was conducted to compare the rankings of map visualizations by subject in determining which visualization was ranked highest in aesthetics. There was no significant effect of visualization type on aesthetics at the  $p < .05$  level for the conditions

$$[F(3, 183) = 1.79, p = .149].$$

**Qualitative evaluation:** Fig. 10 shows the results from qualitative feedback. Among the 30 participants, 27 participants (90%) agreed or strongly agreed that the bristle map was efficient for day and night time comparison in Task 2, 26 for two color maps and 23 for line map. 24 participants (80%) agreed or strongly agreed that the bristle map was efficient for the comparison of two crimes in Task 3, 26 for the line map and 19 for the bivariate map. 29 participants (96.6%) agreed or strongly agreed that the bristle map was efficient for temporal variance representation. In the question for overall efficiency, 27 participants (90%) agreed or strongly agreed that bristle maps and color maps were overall efficient, and 19 (63.3%) for line maps. For point maps, 25 participants (83.3%) disagreed or strongly disagreed.

Participants were also asked to answer visual complexity and preference questions regarding the before (NCR) and after (CR) image pairs applying our clutter reduction. For the circular case (Fig. 8(c)), 96.5% of participants felt that NCR has higher visual complexity and 78.5% preferred CR. For the curved case (Fig. 8(d)), 65.5% of participants answered that CR has a higher visual complexity and 64% preferred NCR. While both cases use a technically identical clutter reduction algorithm, participants reported different visual complexity and preference for them. This indicates that our clutter reduction could be improved by considering the complexity of the underlying network structure.

**Summary and Limitations:** As a univariate encoding, the bristle maps were significantly different (in terms of speed and accuracy) than the point, color and line maps. In the case of the point and line maps, bristle maps use resulted in a higher average correctness and speed; however, the color map for the univariate case had the fastest response and accuracy totals. This seems to indicate that the redundant encoding scheme is actually not beneficial in these cases. As such, use of bristle maps for single variable encoding is not recommended.

With regards to bivariate and multivariate encoding,

TABLE 8: Average time and accuracy.

	Technique	Average time (seconds)	Accuracy (%)
Task 1	Point map	102.6 ± 40.9	41.4 ± 4.3
	KDE map	<b>34.6 ± 7.8</b>	<b>100 ± 0</b>
	Line-T	69 ± 21.9	98.1 ± 3
	Bristle-CLD	50.7 ± 15.3	99.6 ± 1.4
	Bristle-LD	56.6 ± 17.1	99.8 ± 1
Task 2	Point map	106.2 ± 34.7	63.1 ± 6.7
	KDE map	90 ± 30.9	93.1 ± 5.15
	Line-TO	100.5 ± 28.5	87.9 ± 10.4
	Bristle-CLDO	<b>86.3 ± 30.6</b>	90.5 ± 8.5
	Bristle-LDO	87.2 ± 28	<b>93.3 ± 5.2</b>
Task 3	Point map	118.3 ± 41.1	26.6 ± 12.5
	KDE-B	115.3 ± 47.9	72.6 ± 7.9
	Line-TO	<b>84 ± 22.8</b>	<b>96.4 ± 6.2</b>
	Line-CT	86.1 ± 22	86.8 ± 16.7
	Bristle-LDO	88.2 ± 23.7	94.4 ± 7.6
	Bristle-CD	94.5 ± 28.3	90.4 ± 16.7
Task 4	Point maps	194 ± 73.5	53.6 ± 17.6
	KDE maps	171.8 ± 58.8	61.9 ± 15.6
	Line maps	178.9 ± 61.8	75.5 ± 13.5
	Bristle-LDV	<b>48.4 ± 14.8</b>	<b>94.7 ± 13.4</b>

bristle maps and line maps outperformed color and point maps. This is not surprising as bristle and line maps are able to combine variables into a single image, where as in the case of point and color maps, the user must mentally combine the two images together. Bristle-(C)LD also showed a significant effect of the bristle density compared to Line-T. As a bivariate encoding, using orientation in bristle maps was not significant compared to two color maps. However, in the comparison with the bivariate color map, Bristle-LDO showed a significant effect in terms of average time and accuracy. As such, we have that Bristle-(C)LDO as a bivariate encoding scheme created a middle level of cognitive load in-between two color maps and a bivariate color map. Bristle maps also showed potential as a multivariate encoding technique in a single view. Based on the results in Task 3, a point map using various colors and a multivariate color map would considerably increase users' cognitive load. In Tasks 1-3, we also observed that there is no significant effect between the bristle maps using the different encodings. The representation of temporal variance in the bristle map was significantly different from other methods. Our results also showed the differences among point, color and line maps. Participants could better find the region with higher temporal variance when using line maps than using point and color maps. In the qualitative evaluation, 90% of the participants agreed or strongly agreed the overall efficiency of bristle maps to find answers. However, users also strongly preferred the color map in these cases as well.

Finally, we found that with regards to accuracy in identifying values, no technique outperformed any others. However, users were significantly faster in identifying values in both the color and line map scenarios. We hypothesize that in both cases the user focused only on the color, where as in the point map case they needed to count the points and in the bristle map case they needed to reconfirm the univariate

value by double checking several of the encoding legends.

Overall, this technique would be recommended when encoding large amounts of multivariate spatio-temporal point data. As the number of point samples increase, aggregation techniques are need to allow for quick summaries of the data, and, as is evidenced by our studies, pure spatial location representation by glyphs results in too much overlap for accurate measurement and evaluation. As the number of variables increase, color map representations allow for the encoding of variables only along a single visual variable (resulting in bivariate color maps or small multiple plots).

In using multivariate encodings, it is extremely important to understand the interaction effects that the visual variables will introduce in one another. Research into the perceptual interactions among different visual variables was performed by Acevedo and Laidlaw [42]. They measured the perceptual interference of icon size, spacing and brightness, noting that brightness outperforms spacing and size while being subject to interferences from both spacing and size. Acevedo and Laidlaw also noted that spacing also outperformed size, which contradicted some previous results; however, this result seems to align with our participants noting that the bristle spacing was a useful cue. Their results were reportedly due to the spacing sampling along a sinusoidal curve. The sampling of our bristles follow a uniform pattern within classification bins. Thus, there seems to be sufficient scientific evidence to justify using sparsity as a discriminating variable in the case of the bristle maps; however, further studies on this are warranted. Stone [43] has also studies the effect of size in color perception, noting that color appearance changed dramatically with the size being viewed. As such, it may be better to utilize fewer map classifications (color bins) when using bristle maps in order to increase the perceptual distance between each color being visualized.

The main limitations of the bristle map technique is that the combinations of data encoding can potentially prove overwhelming for the designer, and a poor choice on variable encoding can result in a suboptimal visualization. In particular, previous studies have provided results that can be used to predict that certain combinations of visual variables will either enhance or impede map reading. For example, the combination of length and density form an emergent property akin to Bertin's definition of grain. Such effects cannot be ignored; however, bristle maps can be encoded to take advantage of such combinations, as shown in Tasks 3 and 4.

Finally, with regards to scalability of the bristle map technique, in areas of dense roadways, different aggregation methods would need to be considered. As the roads become dense, the ability to plot lines of perceptually different length would become untenable. However, a solution to this would be to draw only the most important roads, thereby removing smaller roads from the analysis, or utilizing bristle maps in a focus+context manner.

To summarize, we posited several different hypotheses. First, we hypothesized that our bristle maps would be



better than other techniques as the complexity level of tasks increased from univariate to multivariate. As the number of variables under analysis increased, the bristle maps outperformed more traditional analyses in terms of speed. However, at lower levels of complexity, traditional techniques such as density estimated heatmaps were found to be best in terms of speed. Second, we hypothesized that our bristle maps would be as accurate as all other representations in determining values. This hypothesis was verified through our user study where subjects estimated the magnitude of crimes on a map using points, density estimated heat maps and bristle maps. Results showed that there were no significant differences between map styles. Finally, we also hypothesized that bristle maps would be ranked higher in terms of their aesthetics. This hypothesis was refuted as our results did not demonstrate any significantly different rating for any of the different mapping techniques.

## 8 CONCLUSIONS

In this work, we have described our novel multivariate data encoding scheme, the Bristle Map. This scheme provides a novel approach for encoding color, length, density, and orientation as data variables and allowing the user to explore correlations within and between variables on a single view. Given the number of parameters available within this encoding, this article has presented only a subset of potential encodings and examples. Here, we have shown the use of encoding bristle lines with redundant information, multivariate attributes for variable comparison, and temporal variance. We also showed a means of potentially encoding data uncertainty. To minimize overlap of bristle lines, we generated a topology graph from underlying geographical line features and employed strategies for clutter reduction. Then, to evaluate the effectiveness of bristle maps, we performed an evaluation study, where we explored different visual encoding combinations within the bristle maps and compared with existing techniques in several tasks. Based on our experiment results, we believe that our bristle map technique has much potential to increase the amount of information that can be visualized on a single map for geovisualization.

## ACKNOWLEDGMENTS

The authors would like to thank Ahmad M. Razip for his help in setting up a web-based user study environment. This work was supported by the US Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001. Jang's work was supported in part by the Industrial Strategic technology development program, 10041772, funded by the Ministry of Knowledge Economy (MKE, Korea). Isenberg's work was supported in part by a French DIGITEO chair of excellence.

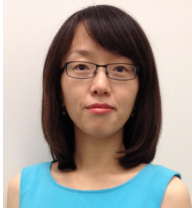
## REFERENCES

- [1] A. MacEachren, D. Xiping, F. Hardisty, D. Guo, and G. Lengerich, "Exploring High-D Spaces with Multi-form Matrices and Small Multiples," in *Proc. InfoVis*. Los Alamitos: IEEE Computer Society, 2003, pp. 31–38. doi> 10.1109/INFVIS.2003.1249006
- [2] C. North and B. Shneiderman, "Snap-Together Visualization: A User Interface for Coordinating Visualizations via Relational Schemata," in *Proc. AVI*. New York: ACM, 2000, pp. 128–135. doi> 10.1145/345513.345282
- [3] C. Weaver, "Cross-Filtered Views for Multidimensional Visual Analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 2, pp. 192–204, Mar./Apr. 2010. doi> 10.1109/TVCG.2009.94
- [4] D. S. Ebert, R. M. Rohrer, C. D. Shaw, P. Panda, J. M. Kukla, and D. A. Roberts, "Procedural Shape Generation for Multi-Dimensional Data Visualization," *Computers & Graphics*, vol. 24, no. 3, pp. 375–384, Jun. 2000. doi> 10.1016/S0097-8493(00)00033-9
- [5] S. Bachthaler and D. Weiskopf, "Continuous Scatterplots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1428–1435, Nov./Dec. 2008. doi> 10.1109/TVCG.2008.119
- [6] R. Maciejewski, S. Rudolph, R. Hafen, A. M. Abusalah, M. Yakout, M. Ouzzani, W. S. Cleveland, S. J. Grannis, and D. S. Ebert, "A Visual Analytics Approach to Understanding Spatiotemporal Hotspots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 2, pp. 205–220, Mar./Apr. 2010. doi> 10.1109/TVCG.2009.100
- [7] J. J. van Wijk and A. Telea, "Enridged Contour Maps," in *Proc. VIS*. Los Alamitos: IEEE Computer Society, 2001, pp. 69–74. doi> 10.1109/VISUAL.2001.964495
- [8] R. J. Phillips and L. Noyes, "An Investigation of Visual Clutter in the Topographic Base of a Geological Map," *Cartographic Journal*, vol. 19, no. 2, pp. 122–132, Dec. 1982. doi> 10.1179/000870482787073225
- [9] S. Openshaw, "The Modifiable Areal Unit Problem," in *Concepts and Techniques in Modern Geography*. Norwich, UK: Geo Books, 1984, vol. 38.
- [10] M. Swink and C. Speier, "Presenting Geographic Information: Effects of Data Aggregation, Dispersion, and Users' Spatial Orientation," *Decision Sciences*, vol. 30, no. 1, pp. 169–195, Jan. 1999. doi> 10.1111/j.1540-5915.1999.tb01605.x
- [11] C. L. Eicher and C. A. Brewer, "Dasymetric Mapping and Areal Interpolation: Implementation and Evaluation," *Cartography and Geographic Information Science*, vol. 28, no. 2, pp. 125–138, Apr. 2001. doi> 10.1559/152304001782173727
- [12] R. Spence, *Information Visualization*. Reading, MA, USA: Addison-Wesley, 2001.
- [13] L. Wilkinson, *The Grammar of Graphics*, 2nd ed. Heidelberg/Berlin: Springer-Verlag, 2005.

- [14] A. M. MacEachren, *How Maps Work: Representation, Visualization, and Design*. Guilford Press, 1995.
- [15] S. Chainey, L. Tompson, and S. Uhlig, "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime," *Security Journal*, vol. 21, no. 1–2, pp. 4–28, Feb.–Apr. 2008. doi> 10.1057/palgrave.sj.8350066
- [16] B. W. Silverman, "Density Estimation for Statistics and Data Analysis," in *Monographs on Statistics and Applied Probability*. New York: Chapman & Hall, 1986, no. 26.
- [17] C. Ahlberg and B. Shneiderman, "Visual Information Seeking using the FilmFinder," in *Proc. CHI*. New York: ACM, 1994, pp. 433–434. doi> 10.1145/259963.260431
- [18] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner, "Structure-Based Brushes: A Mechanism for Navigating Hierarchically Organized Data and Information Spaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 6, no. 2, pp. 150–159, Apr.–Jun. 2000. doi> 10.1109/2945.856996
- [19] A. Dix and G. Ellis, "By Chance: Enhancing Interaction with Large Data Sets Through Statistical Sampling," in *Proc. AVI*. New York: ACM, 2002, pp. 167–176. doi> 10.1145/1556262.1556289
- [20] A. MacEachren, "Visualizing Uncertain Information," *Cartographic Perspectives*, no. 13, pp. 10–19, Fall 1992.
- [21] R. Dunn, "A Dynamic Approach to Two-Variable Color Mapping," *The American Statistician*, vol. 43, no. 4, pp. 245–252, Nov. 1989. doi> 10.1080/00031305.1989.10475669
- [22] J. Olson, "Spectrally Encoded Two-Variable Maps," *Annals of the Association of American Geographers*, vol. 71, no. 2, pp. 259–276, Jun. 1981. doi> 10.1111/j.1467-8306.1981.tb01352.x
- [23] A. MacEachren and D. DiBiase, "Animated Maps of Aggregate Data: Conceptual and Practical Problems," *Cartography and Geographic Information Systems*, vol. 18, no. 4, pp. 221–229, Oct. 1991. doi> 10.1559/152304091783786790
- [24] H. Hagh-Shenas, S. Kim, V. Interrante, and C. Healey, "Weaving Versus Blending: A Quantitative Assessment of the Information Carrying Capacities of two Alternative Methods for Conveying Multivariate Data with Color," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1270–1277, Nov./Dec. 2007. doi> 10.1109/TVCG.2007.70623
- [25] T. Saito, H. N. Miyamura, M. Yamamoto, H. Saito, Y. Hoshiya, and T. Kaseda, "Two-Tone Pseudo Coloring: Compact Visualization for One-Dimensional Data," in *Proc. InfoVis*. Los Alamitos: IEEE Computer Society, 2005, pp. 173–180. doi> 10.1109/INFOVIS.2005.35
- [26] M. Sips, J. Schneidewind, D. A. Keim, and H. Schumann, "Scalable Pixel-Based Visual Interfaces: Challenges and Solutions," in *Proc. IV*. Los Alamitos: IEEE Computer Society, 2006, pp. 32–38. doi> 10.1109/IV.2006.95
- [27] C. Panse, M. Sips, D. Keim, and S. North, "Visualization of Geo-spatial Point Sets via Global Shape Transformation and Local Pixel Placement," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 749–756, Sep./Oct. 2006. doi> 10.1109/TVCG.2006.198
- [28] D. Dorling, A. Barford, and M. Newman, "Worldmapper: The World as You've Never Seen it Before," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 757–764, Sep./Oct. 2006. doi> 10.1109/TVCG.2006.202
- [29] P. C. Wong, K. Schneider, P. Mackey, H. Foote, G. Chin, R. Guttmerson, and J. Thomas, "A Novel Visualization Technique for Electric Power Grid Analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 3, pp. 410–423, May/Jun. 2009. doi> 10.1109/TVCG.2008.197
- [30] D. Fisher, "Hotmap: Looking at Geographic Attention," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1184–1191, Nov./Dec. 2007. doi> 10.1109/TVCG.2007.70561
- [31] C. Tominski, P. Schulze-Wollgast, and H. Schumann, "3D Information Visualization for Time Dependent Data on Maps," in *Proc. InfoVis*. Los Alamitos: IEEE Computer Society, 2005, pp. 175–181. doi> 10.1109/IV.2005.3
- [32] J. Bertin, *Semiology of Graphics*. Redlands, California: ESRI Press, 2011.
- [33] J. Tarbell, "Substrate," Web site & simulation: <http://www.complexification.net/gallery/machines/substrate/>, 2003, accessed February 2012.
- [34] T. Isenberg, "Visual Abstraction and Stylisation of Maps," *The Cartographic Journal*, vol. 50, no. 1, pp. 8–18, Feb. 2013. doi> 10.1179/1743277412Y.0000000007
- [35] P. Rheingans, "Task-Based Color Scale Design," in *Proc. SPIE*, vol. 3905. SPIE, 2000, pp. 35–43. doi> 10.1117/12.384882
- [36] C. Ware, "Color Sequences for Univariate Maps: Theory, Experiments and Principles," *IEEE Computer Graphics and Applications*, vol. 19, no. 5, pp. 41–49, Sep./Oct. 1988. doi> 10.1109/38.7760
- [37] C. A. Brewer, *Designing Better Maps: A Guide for GIS Users*. Redlands, CA, USA: ESRI Press, 2005.
- [38] M. Prasad, "Intersection of Line Segments," in *Graphics Gems II*, J. Arvo, Ed. Boston: Academic Press, 1991, pp. 7–9.
- [39] A. Michotte, G. Thinès, and G. Crabbé, *Les Compléments Amodeux des Structures Perceptives (Amodal Completions of Perceptual Structures)*. Louvain: Institut de Psychologie del'Université de Louvain, France: Studia Psychologica, 1964.
- [40] M. S. Farrar, *Magic Squares*. Charleston, SC, USA: BookSurge Publishing, 1996.
- [41] R. A. Likert, "A Technique for the Measurement of Attitudes," *Archives of Psychology*, vol. 22, no. 140, pp. 5–55, 1932.
- [42] D. Acevedo and D. Laidlaw, "Subjective Quantifi-

cation of Perceptual Interactions among some 2D Scientific Visualization Methods,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 1133–1140, Sep. 2006. doi> 10.1109/TVCG.2006.180

- [43] M. Stone, “In Color Perception, Size Matters,” *IEEE Computer Graphics & Applications*, vol. 32, no. 2, pp. 8–13, Mar./Apr. 2012. doi> 10.1109/MCG.2012.37



**SungYe Kim** received her Ph.D. in Electrical and Computer Engineering from Purdue University in May, 2012. She also received her masters degree in Computer Science and Engineering from Chung-Ang University, South Korea in 2000. She is currently a graphics software engineer at Intel Corporation. Prior to this, she was employed as a research engineer at the Electronics and Telecommunications Research Institute from 2000 to 2006. Her research interests are

computer graphics, illustrative visualization, visual analytics and information visualization.



**Ross Maciejewski** received his Ph.D. in Electrical and Computer Engineering from Purdue University in December, 2009. He is currently an Assistant Professor at Arizona State University in the School of Computing, Informatics & Decision Systems Engineering. Prior to this, he served as a Visiting Assistant Professor at Purdue University and worked at the Department of Homeland Security Center of Excellence for Command Control and Interoperability in the Visual Analytics for

Command, Control, and Interoperability Environments (VACCINE) group. His research interests are geovisualization, visual analytics and non-photorealistic rendering.



**Abish Malik** is a Ph.D. student in the School of Electrical and Computer Engineering at Purdue University and a research assistant at the Purdue University Rendering and Perception Lab. He received his B.S. degree in Electrical Engineering from Purdue University in 2009. His research interests include visual analytics, correlation and predictive data analytics.



**Yun Jang** is an assistant professor of computer engineering at Sejong University, Seoul, South Korea. He received the masters and doctoral degree in electrical and computer engineering from Purdue University in 2002 and 2007, respectively, and received the bachelors degree in electrical engineering from Seoul National University, South Korea in 2000. He was a postdoctoral researcher at CSCS and ETH Zürich, Switzerland from 2007-2011. His research interests

include interactive visualization, volume rendering, visual analytics, and data representations with functions. He is a member of the IEEE.



**David S. Ebert** is a professor in the School of Electrical and Computer Engineering at Purdue University, a University Faculty Scholar, director of the Purdue University Rendering and Perceptualization Lab, and director of the Purdue University Regional Visualization and Analytics Center. His research interests include novel visualization techniques, visual analytics, volume rendering, information visualization, perceptually based visualization, illustrative visualization, and procedural ab-

straction of complex, massive data. Ebert has a PhD in computer science from Ohio State University and is a fellow of the IEEE and member of the IEEE Computer Society's Publications Board.



**Tobias Isenberg** is a senior research scientist with INRIA in France. He received his doctoral degree from the University of Magdeburg, Germany. Previously, he held positions as assistant professor for computer graphics and interactive systems at the University of Groningen, the Netherlands, and as post-doctoral fellow at the University of Calgary, Canada. He works on topics in interactive non-photorealistic and illustrative rendering as well as computational aesthetics and ex-

plores applications in scientific visualization. He is a member of the IEEE.