

Radial Visualization of Multidimensional/Multivariate Data: A Survey

Gabriel Zhou

Abstract— Visualizing multidimensional/multivariate data is essential but can also be challenging. Among all the techniques that can be applied to visualize these high dimensional data, radial visualization turns out to be an interesting approach. By mapping data in a circular pattern, radial visualization takes advantage of the unique characteristics of a circle. We classify multidimensional/multivariate radial visualization techniques into two major categories: axis-based and segment-based. For axis-based visualization, data attributes are assigned to the axes of the circle. For segment-based visualization, data attributes are assigned to the segments of the circle. We put techniques that are hard to fit in these two categories into the “others” category. By surveying papers related to this topic, we prove the validity and effectiveness of radial visualization as a distinct set of techniques to present multidimensional/multivariate data.

Index Terms—Multivariate visualization, visualization techniques and methodologies, radial visualization

1 INTRODUCTION

With the rapid development of computing and sensing technologies, the available number of data continuously grows. Observation of data can have more than one outcome variable, and such data is referred to as multidimensional/multivariate data. These data appear in various scientific fields including chemistry, biology, engineering, etc. Studying and analyzing these multidimensional/multivariate data are essential to answer domain questions and make research progresses. Visualization helps with data analysis by presenting data in a graphical form to improve understanding. However, high dimensional visualization has always been a big challenge. The current solutions contain some dimensionality reduction algorithms (e.g. PCA or t-SNE) and direct multidimensional/multivariate data visualization techniques (e.g. Parallel Coordinates or Scatterplot Matrix).

Radial visualization is a prevalent methodology that displays data in a circular pattern. Radial visualization can be essentially considered a circle, and it shares the unique and interesting characteristics of a circle. Giving a point at any position in a radial layout, that location can be represented with an angle and a radius, which can be further considered as a length. Length along with position are ranked at the top among all visualization channels, and the angle channel is placed right after the length channel. Therefore, when using a radial layout, we are naturally applying two most reliable visualization channels in one system. Another important characteristic of a circle is the inherent periodicity. There is no defined starting or end point in a circle, any point on the circle circumference can go around in either direction and back at its initial position. The inherent periodicity of a circle is a huge advantage when presenting periodic data (e.g. time referenced data). Moreover, circle is a compact shape. For the same perimeter, circle is proved to have the largest area. Its compactness also reflects on how data are distributed in the layout. Many other visualization layouts restrict mapping data in certain directions. However, in a radial layout, data are usually separated out in every direction and fill the whole canvas. These characteristics make the radial visualization one of the most flexible visualization methods out there.

We are interested in the feasibility and effectiveness of using radial visualization as a possible solution to present multidimensional/

multivariate data. In this survey paper, we go through several visualization papers that focus on applications of radial visualization techniques on multidimensional/multivariate data. The contents include origins and designs of these radial visualization techniques, limitations and improvements on these original techniques, comparison between different techniques, and case studies that apply them in real-world problems. In Section 2, we explain some important terminologies. Section 3 discusses related works with previous survey papers on radial visualization or multidimensional/multivariate data visualization. We divide all radial visualization techniques into two major types: axis-based and segment-based. Each is named after the way data attributes are assigned in a circle, either to the axes or to the segments. Section 3 includes three axis-based radial visualization techniques. They are Radar Chart, RadViz, and Star Coordinates. Section 4 includes two segment-based radial visualization techniques, which are Circle Segments and Circle View. Some techniques cannot be simply classified into these two categories and we present them in Section 5 as “others”.

2 TERMINOLOGY

2.1 Radial Visualization

Defined by Draper et al. in their survey paper [1]:

“radial visualization describes any interactive system that arranges data in a circular fashion.”

In this given definition, radial visualization represents an interactive system by default. This holds to be accurate as all radial visualization techniques discussed in this paper provide certain interactions. On the other hand, a traditional non-interactive radial layout is referred to as a *radial display*. The simplest radial visualization can be a pie chart. When a visualization idiom is arranged into a radial layout, it can also be seen as an application of radial visualization. For example, using a radial node-link diagram to present a tree. Radial visualization is a widely used methodology in information visualization.

2.2 Multidimensional/Multivariate Data

The terms multidimensional and multivariate are often used vaguely in visualization papers. Despite being the common terms, the definitions and differences between these two often remain clarifications. A formal definition proposed by Santos and Brodlie [2]:

“Think of data as a sample from k -variate function $F(x)$ defined over an n -dimensional domain D . Thus $F = (f_1, f_2, \dots, f_k)$ has k components, and $X = (x_1, x_2, \dots, x_n)$ is a point in D . We shall allow k

• Gabriel Zhou is with University of British Columbia, Department of Mechanical Engineering, 6250 Applied Science Ln #2054, Vancouver, BC V6T 1Z4. E-mail: gabriel_zxs@outlook.com.

to be zero, in which case we just have a point in D , and we allow n to be zero, in which case we just have a value of F . We shall talk in terms of dependent variables F and independent variables X .”

In this case, data with multiple independent variables X is called multidimensional data and data with multiple dependent variables F is called multivariate data. Although being very exact and rigorous, this definition may seem to be less intuitive and comprehensible. Munzner describes the two terms in her book [3] with respect to key attributes and value attributes:

For a table dataset, a key attribute acts as an index to lookup value attributes. “Multidimensional” refers to multiple key attributes, and “multivariate” refers to multiple value attributes.

This definition is easier for readers to understand. It is also worth to mention that a dataset can be multidimensional and multivariate at the same time. Visualization papers use *attributes*, *dimensions*, and *variables* these three terms interchangeably. When the term “*multidimensional data*” is mentioned, it is usually referred to as *multivariate data* rather than the “actual” *multidimensional data* we described above. In this survey paper, we focus on using radial visualization to present *multivariate data*. *Multidimensional data* is used as a synonym of *multivariate data* in the following sections.

3 RELATIVE WORK

Despite the term *radial visualization* had been mentioned as early as in 1997 by Hoffman et al. [4] and became increasingly popular around the early 21st century, it was not considered as a distinct methodology of its own. Until in 2009, Draper et al. [1] presented their survey work on radial visualization, and this idea had finally brought up in front of people. Their contributions are majorly from three perspectives. First, they identified different problems that radial visualization had applied. Second, they proposed seven design patterns for the radial visualization system at that time. Third, they came up with five design choices for designers to consider when creating new radial visualization systems. Draper’s work was a significant milestone on the road of radial visualization. It drew people’s attention to this rising visualization technique, elevated it from an interesting visual gimmick to a unique methodology. Moreover, the taxonomy presented in this paper built a solid foundation for further research. As we talk about the various radial visualization techniques in later sections, we will use these well-defined patterns as references.

Table 1: Design Patterns for Radial Visualization [1]

	Pattern
Polar Plot	<i>Tree</i>
	<i>Star</i>
Space Filling	<i>Concentric</i>
	<i>Spiral</i>
	<i>Euler</i>
Ring	<i>Connected Ring</i>
	<i>Disconnected Ring</i>

Although Draper et al. did a great job of summarizing the status of radial visualization at that time, one decade has passed and radial visualization has become a sophisticated visualization metaphor. The scope of its applications on various domain problems has undoubtedly grown as well. Moreover, Draper’s work presented a very general overview and guideline about radial visualization, while we would narrow down the coverage and focus on the performance of radial visualization in multivariate data, a topic few have studied.

On the other hand, visualizing multivariate data has always been a hard challenge for data scientists and visualization designers. Enormous researches and papers were presented to overcome this difficulty. A few comprehensive survey papers have concluded the current techniques for multidimensional/multivariate data

visualization. He et al. proposed a survey of visualization of multivariate spatial data [5], focused on a certain type of multivariate data rather than the type of visualization. Maalej et al. introduced various visualization techniques as valid solutions to multidimensional data [6]. Radial visualization is only briefly mentioned in this overview without an in-depth exploration. Chan and dos Santos conducted highly comprehensive surveys on multidimensional/multivariate data visualization [7] [8], respectively. Again, with full coverage on this topic, they did not present applications of radial visualization in detail. Also, they did not consider radial visualization as a separate class of techniques for multidimensional/multivariate data visualization.

4 AXIS-BASED

4.1.1 Radar Chart

Radar Chart (RC) is also known as web chart, spider chart, star chart, star plot, cobweb chart, irregular polygon, polar chart, or Kiviati diagram. The first appearance of RC can go way back to 1877 in Mayr’s book “Die Gesetzmäßigkeit im Gesellschaftsleben” written in German. However, the formal concept and conventions about this technique were proposed by Kolence in 1973 [9]. With some minor differences in how the graph is encoded, the general idea of this technique is the same. RC consists of a series of spokes projecting from the center point, with each spoke representing one of the variables. The value of each variable is encoded with the length of each spoke. The plotted values are connected to form an enclosed figure. In this case, RC is equivalent to the parallel coordinates, with axes spaced radially and originated at a single point. In a RC, usually the size and shape of the resulting polygons are interested.

Being the sibling of the parallel coordinates, RC shares some bad genes. One significant limitation of RC is its scalability, with multivariate datasets of more than a few hundred items, the plot can become overwhelming. That is only the case RC is plotted the same way as the parallel chart. In other implementations, the enclosed polygons are color coded to emphasize the resulting size and shape (also called Filled Radar Chart). The color may or may not be transparent, but in either case, its scalability issue is more severe and significantly reduces its effectiveness.

Comparing the values of data from one single axis in a RC is pretty straightforward, but comparing values across different axes is less intuitive. This is because it is harder to judge the length in a radial layout than in a rectilinear layout, in which case using a line graph or bar chart would be more preferred. A solution to this problem is to use guidelines or gridlines to help with cross-axes comparison. However, in many cases, comparison across axes in a RC is meaningless in the first place, as variables on axes are nominally independent with different scales. In these cases, although the problem with cross-axes comparison is automatically resolved since there is no point for comparison at all, another problem arises at the same time. RC may mislead people to compare these variables naturally, such comparison may cause confusion and false conclusions [10].

Sorting of axes has a significant impact on the shape of the enclosed figure. Since size and area is usually the primary factor in a RC, sorting of axes can further affect people’s judgement. When data is presented as a thin and narrow shape, it becomes difficult and misleading to compare with another relative full shape. Although Leary et al. remarked “a well-performing unit is always likely to have a larger area than a less well-performed unit” [11], Feldman [12] argued against it with his example as shown in Fig. 1. Another issue related to the project area is its nonlinearity. The area of the enclosed shape in a RC is proportional to the square of the values, meaning a change in values results in a greater change in the area. One could very possibly exaggerate the significance of a change in values by observing the resulted change in the area.

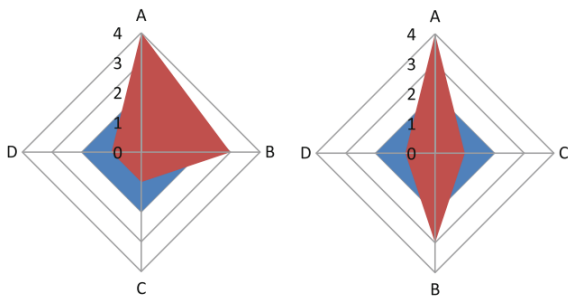


Fig. 1. Two filled radar charts with measures on different axis [12]. The area in red is reduced on the right.

Porter and Pooya apply RC to compare the multidimensional functional performance of natural biological materials based on their mechanical properties [13]. Since the performance of a functional task is not determined by a single mechanical property but a combination of properties, the correlations between properties and functions are interested. Each axis represents a mechanical property (strength, modulus, etc.) and is normalized by maxima from 0 to 1. The permutation of axes is searched and determined by the maximal total area of resulting profiles. Shape descriptors like Jaccard indices, shape moments are also calculated as metrics to further compare these data. The Jaccard index, defined as the size of the intersection divided by the size of the union, determines the similarity of two profiles. The relative multidimensional performance is indicated by its normalized first moment of area relative to the boundary of the property space. As shown in Fig. 2, a “gap” on the chart would represent the relative intensity of a trade-off, larger gap suggesting the opposition of properties. With permuted radar charts, the research group validates that natural materials can hold strength and toughness at the same time, while they are considered to be mutually exclusive in mechanical materials. They further apply their radar charts to comparison beyond materials, on the performance of feeding vs singing of Darwin’s finches. They are able to successfully draw conclusions from the presented radar charts, which verifies the applicability of their method.

Researchers also point out some constrains with their design, including “relative” performance comparisons and the cautious choices of various descriptors. Overall, researchers are satisfied with the performance of permuted radar charts and suggest potential applications on other filed of biological science or engineering.

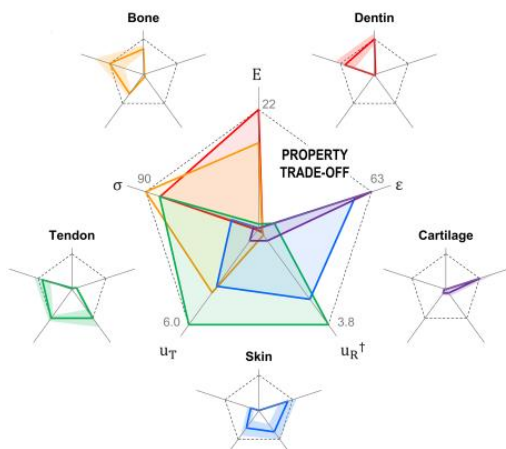


Fig. 2. Normalized, permuted radar charts comparing five collagenous tissues, where the profile averages (lines) and standard deviation/ranges (shaded regions) are shown on the outer edges; the center plot illustrates a trade-off between stiffness and extensibility [13].

4.1.2 RadViz

As mentioned above, Hoffman was the first person who brought up the term “radial visualization”. Not only that, but he also introduced his visualization approach in the same paper: RadViz [4], which later became one of the most popular techniques to present multivariate data. The name of this visualization system is somewhat ambiguous as people may inadvertently interpret it as the abbreviation of “radial visualization”. We clearly state here that RadViz is a standalone name for this specific visualization technique.

The design theory behind RadViz is innovative but simple at the same time. Dimensions are equally spread out on the circumference of the circle and act as pins or anchors. Imaging an invisible spring is connected to a data point and the anchor of the i -th dimension. The spring constant K_i would equal to the value of the data point on the i -th dimension. Now for n -dimensional data, there will be n springs connecting the data point to the anchors, and the position of the data point is where the sum of the spring forces equals 0. Usually, the data points values need to be normalized between 0 to 1 to have a reasonable position, and the dimensions must be nonnegative. In RadViz, all data points lie inside the convex hull of anchors.

In this case, a data point that has a higher value in one dimension would have a shorter distance toward the anchor of the corresponding dimension anchor. With some basic knowledge in mechanics, we can further infer that a data point with equal values in all dimensions locates right at the center of the circle. An n -dimensional line will be map to a line, and a sphere will map to an ellipse. An intuitive graph interpretation of RadViz is illustrated in Fig. 3. Beware that in practical applications, these “spring lines” are usually hidden.

$$\mathbf{x} = (0.5, 0.25, 0, 0.25, 0.5, 1)$$

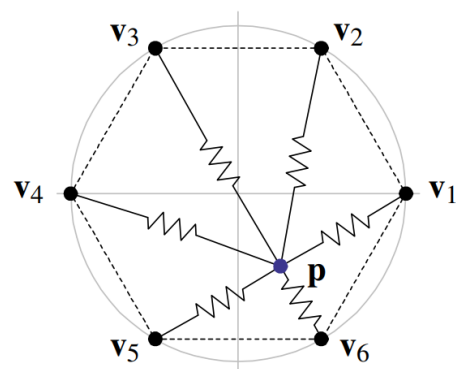


Fig. 3. RadViz encoding illustration [14].

Hoffman’s motivation of creating RadViz was to find a new way to distinguish coding DNA sequences (exons) from non-coding DNA sequences (introns). In order to do this, symbolic DNA sequences are represented by numbers or vectors. In his application, the dimensions are nucleotides (A, C, T, G) with their appears in position (1, 2, 3, 4, etc.). Exons and introns are encoded by shape and color channels with exons as red crosses and introns as green circles. As shown in Fig. 4, since most data points lie close to the center, zooming is provided for a clear observation of the data.

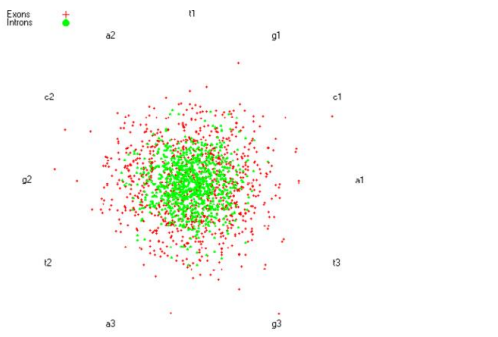


Fig. 4. displays 2000 points zoomed up by a factor of 5. In this picture we can see that the exons (red +) are more spread out, and the introns (green) are closer to the center of the circle [4].

Despite the popularity of RadViz, it has some significant drawbacks. Data points with different values but the same scale numbers in each dimension will end in the exact same position, which means data points that are different in original space cannot be distinguished from the visualization layout [15]. This particular drawback is referred to as ambiguity. Another issue with RadViz is its high dependency on the order of dimension anchors. When applying visualization on multivariate data, people are often seeking to find the correlations, patterns and outliers in these data. However, in RadViz, the order of anchors can change the visualization view significantly, further influence the detection of statistical relationships in these data. It is a tough task to adjust the anchor order and hope to find the optimal view in the end. For even higher dimensional data, finding the optimal order of dimensions in RadViz is proved to be NP-complete [16]. A very detailed research on the properties of RadViz by Daniels et al. provides a more in-depth insight on this technique [17].

We mentioned the ambiguity problem of RadViz, which occurs in special cases where values of data on each attribute are different by the same factor. However, even in normal cases, this ambiguity still exists to a certain extent. Data points on the RadViz layout cannot reveal their actual relationship in original space accurately. Data points that are not related may be presented very close on the graph, mislead observers to false conclusions. RadViz Deluxe [18] solves this issue by adding a three-stage refinement algorithm. Each stage reduces the error of distance between data-to-data, data-to-variable, and variable-to-variable respectively. The first stage reduces the variable-to-variable error by arranging anchors based on the approximate solution for Hamilton Cycle. The second stage reduces the data-to-variable error by using an iterative algorithm to reduce the error polygon area of each data point. The third stage reduces the data-to-data error by adjusting data points positions via force-directed layout, based on the magnitudes of errors between data points. RadViz Deluxe results in a distance-spaced layout, which the anchors are not evenly spaced on the circumference, a comparison between RadViz and RadViz deluxe is shown in Fig. 5. RadViz Deluxe effectively mitigates the ambiguity issue.

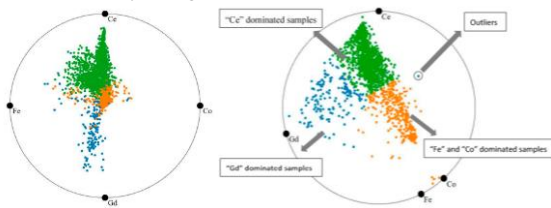


Fig. 5. The battery data set visualized with RadViz (left) and RadViz Deluxe (right). Compared to the original RadViz, the samples in the visualization generated by RadViz Deluxe are more scattered based on their components [18].

RadViz++ [19] presents another advanced RadViz layout with a set of interactive techniques. It sorts the anchor ordering based on the similarity of variables, using the Pearson correlation coefficient and the average-linkage Agglomerative Hierarchical Clustering (AHC). The anchor ordering follows the ordering of leaves in the computed cluster dendrogram.

Two metaphors are proposed for variable-to-variable analysis. First is the variable hierarchy, which maps the prementioned cluster dendrogram on the graph. Parent nodes surround the basic circle layout as stacked fragment rings, and colormap from blue (similar) to green to red (dissimilar) indicates the similarity. Second, to further clarify the similarity between variables, RadViz++ uses bundled edges and hierarchical edge bundles (HEB) to connect similar variables, with transparency representing the level of similarity. To provide more detailed information about variable values, a histogram is plotted in each ring showing the value distribution of each variable. Clicking on a certain bin in the histogram highlights the data points in that value range. Based on the variable hierarchy, users are allowed to interactively aggregate or filter variables to compute a new layout with reduced dimensionality. To enhance data-to-data analysis, RadViz++ provides animation to switch from mapping generated by RadViz and mapping generated by other dimensionality reduction methods such as LAMP or t-SNE. When the layout is switched, RadViz anchors are greyed out to prevent misinterpretation. These DR methods usually present better cluster separation. In addition, analyzing high dimensional data with different visualizations is proved to be very effective. Sometimes users may be interested in analyzing a group of data points on the plot. In this case, a brushing-and-linking tool can be used to link selected data points to their variable histogram bins. Comparing to the original RadViz, RadViz++ has better scalability in variables and reduces the ambiguities in data analysis.

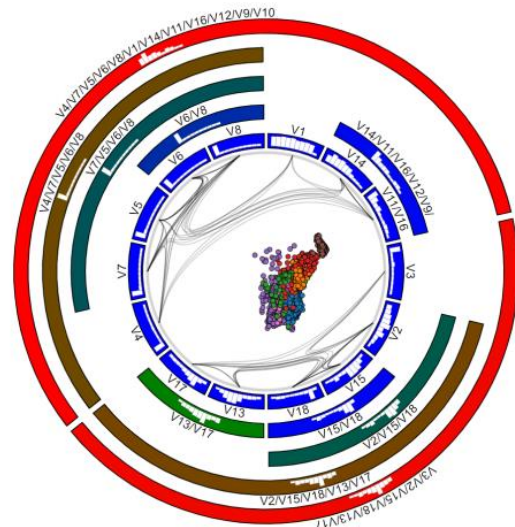


Fig. 6. RadViz++, with HEB and histograms [19].

4.1.3 Star Coordinates

Kandogan created star coordinates (SC) for the purpose of help people to understand their data easily at the early stage of their analysis tasks [20]. This technique is similar to scatterplot in spirit as it extends a typical scatter plot to higher dimensions.

Axes representing attributes are evenly radiated out from the center point of the circle with the same length (initially). The final position of a data point would be the sum of all unit vectors on each coordinate multiplied by the value of the data for that coordinate. Fig. 8 can be used to help with understanding this approach. Theoretically, we can start from any axis and follow any order we want. In this case, as the figure illustrated, we start with axis C_1 and going clockwise, locate the data point based on the first axis as its coordinate. Then change the coordinate to the next axis and move the data point according to its value on that axis. The choice of axes sequence does not matter, eventually we end up in the position P.

SC also has some defects. It has the ambiguity issue just like RadViz, which a single data point on SC can represent various actual data. To overcome or at least reduce the effect of this problem, Kandogan provides several interaction approaches. These interactions are further adapted in latter SC applications and became essential components of SC.

The two major interactions are scaling and rotation. Scaling allows the user to change the length of the axis by dragging the end point of an axis away or towards the origin. By doing so, the contribution of a particular attribute can be increased or decreased on the visualization plot. Rotation allows the user to change the direction or angle of the axis by dragging any point on the axis and move to the desired position. Rotation changes the correlations of a specific axis to other axes.

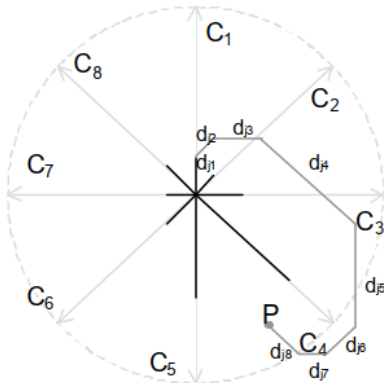


Fig. 8. Calculation of data point location of an 8-dimensional dataset [20].

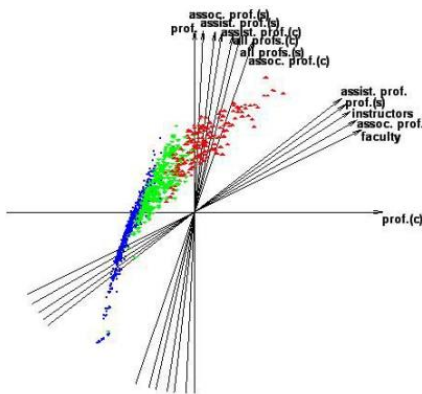


Fig. 9. The visualization result of ASC [21].

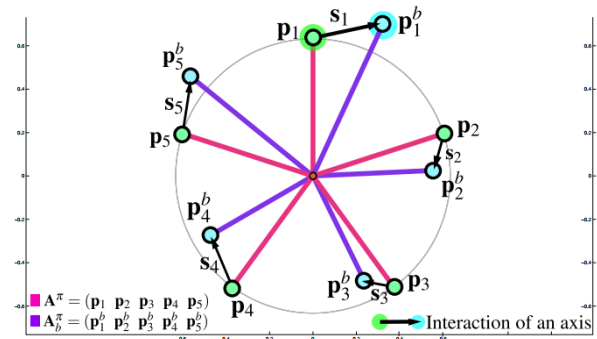


Fig. 7. Orthography-preserving interaction with the axes of Orthographic Star Coordinates [22].

Advanced Star Coordinates (ASC) [21] uses the diameter line across the circle instead of the spoke radiated out from the origin as the attribute axis. Combining the new coordinates with an optimized algorithm, the ambiguity problem of SC is resolved. The data structure and spatial relationship in the original data space is reserved in the ASC layout. ASC also applies a dimension sorting strategy based on Ankerst's dimension arrangement algorithm [16]. Both the axis ordering and the angle between axes are determined based on correlation, with higher correlation axes next to each other and a small angle between them.

Orthographic Star Coordinates (OSC) [22] restricts SC to orthographic projections to prevent data shape distortion. Interactive adjustments of axes are responded by on the fly repeated optimization, keeping the orthographic projection after interaction. OSC provides a distortion-free global patterns visualization in the original data space.

4.1.4 Comparison between RadViz and Star Coordinates

RadViz and SC are the two most popular techniques to present multivariate data in radial layouts. They are two visualizations techniques that support high dimensional data-to-data, data-to-variable, and variable-to-variable analysis at the time [18]. Rubio-Sanchez et al. point out that the biggest difference between RadViz and SC is the linearity [14]. SC generates linear mappings, while the normalization step in RadViz introduces nonlinearities. In some special propositions, RadViz and SC are equivalent. However, in most applications, RadViz and SC generate different graphs due to the differences in algorithm and linearity. Rubio-Sanchez et al. claim that nonlinearities affect the performance of RadViz in various aspects.

Interpretation of the clusters can be more difficult in RadViz than in SC. When mapping a multivariate dataset with normal distribution, a plot generated by RadViz may not follow a normal distribution. Data at ends may be significantly far from others, misleading people to consider them as outliers. Naturally, the same problem would also influence outlier detection. RadViz has a lower detection probability than SC for extreme value outliers. However, for low likelihood outliers, both approaches have a poor performance. RadViz also tends to clump data points close to its origin. The average distance of data points to origin decreases as the number of data points increases, while it is the other way around for SC. However, the shape of distribution becomes very close for the two methods as the number of variables increases. The problem can also be solved by using the zooming feature in RadViz. The location of the cluster in its original data space can affect its appearance on RadViz. A cluster near origin will have a larger size when it is presented on RadViz. In a special case where the data only contains one non zero value, users can easily read out that value in SC. Whereas RadViz cannot provide any information about the actual value because the data is mapped at the anchor. The most significant advantage of SC over RadViz is the high degree of freedom of the

axes, which mainly contribute to data classification and separation. Users are able to manipulate the axes configuration of SC to reproduce other linear transformations like PCA, LDA, biplots, etc. Although some advanced algorithms support RadViz to perform such tasks as well, the accuracy is not as good as SC. Moreover, although recovering the actual data values is different for both techniques, users can get a relatively accurate estimation with proper axes configuration in SC. On the other hand, Rubio-Sanchez et al. admit the advantage of RadViz in analyzing sparse data. As shown in Fig. 10, RadViz presents sparse data close to anchors and non-sparse data close to the origin, while on SC they may be mapped in the same area and hard to distinguish.

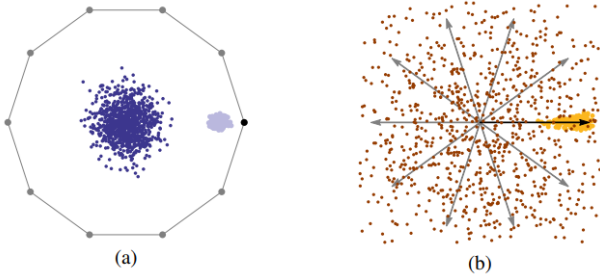


Fig. 10. Usefulness of RadViz for visualizing sparse data [14].

5 SEGMENT-BASED

5.1.1 Circle Segments

Circle Segments (CS) uses the pixel-oriented technique [23] to visualize a large amount of multivariate data in a circle. First introduced by Keim, the pixel-oriented technique uses each pixel in the display to encode one data value in order to visualize the largest amount of data possible. Since this technique uses pixels to map data, the arrangement becomes extremely important. Keim proposed several pixel-orientated arrangements for two different purposes. The first one is the query-independent visualization, which the data has a natural ordering in one or more data variables (e.g. time). The second one is the query-dependent visualization, which the data set has no natural ordering of data.

Based on Keim's concept, Ankerst et al. [24] propose CS as a new pixel-oriented visualization method. CS assigns dimensions to the segments of a circle. A k dimensional data would cut the circle into k equivalent segments with each segment representing one dimension. As illustrated in Fig. 11, pixel representing the data point from one dimension goes back and forth along the "draw line", filling the whole plot and end up being a regular polygon. In this proposed layout the trace of pixel represents the time dimension, with the initial data at the center and the more recent data on the outside. Coloring is used to indicate the magnitude of the data value, with high values to light colors and low values to dark colors. CS requires data set with at least three dimensions, and users can change the ordering of dimensions. Comparing CS to the traditional line graph (Fig. 12 and Fig. 13), CS does not have the drawback of data overlap, which makes the line graph unreadable in high dimensions. Moreover, it is easier to find analogic tendencies in different dimensions based on CS rather than the traditional line graph. CS is further compared with other pixel-oriented visualizations proposed by Keim, including the spiral technique and the recursive pattern technique. CS is proved to be more perceivable in terms of finding patterns, correlations and dependencies between dimensions.

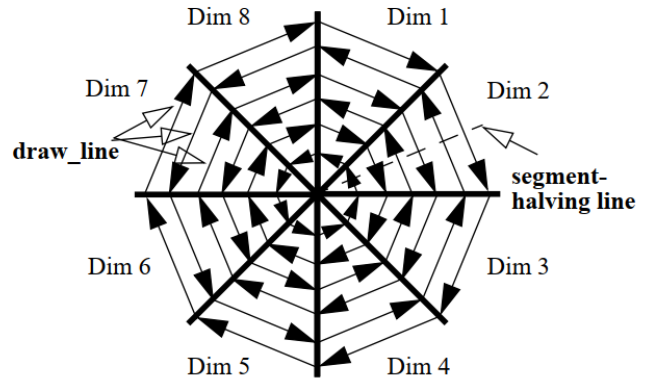


Fig. 11. "Circle Segments" technique for 8-dimensional data [24].

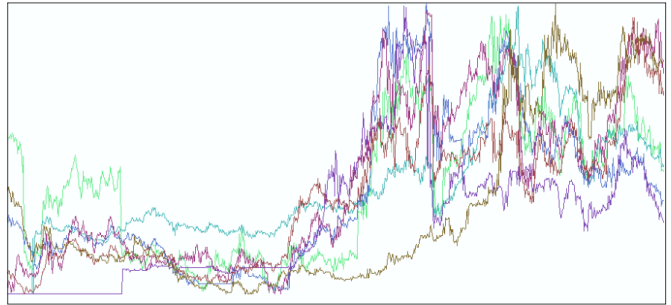


Fig. 12. Visualizing 7-dimensional data using the "Line Graph" visualization technique [24].

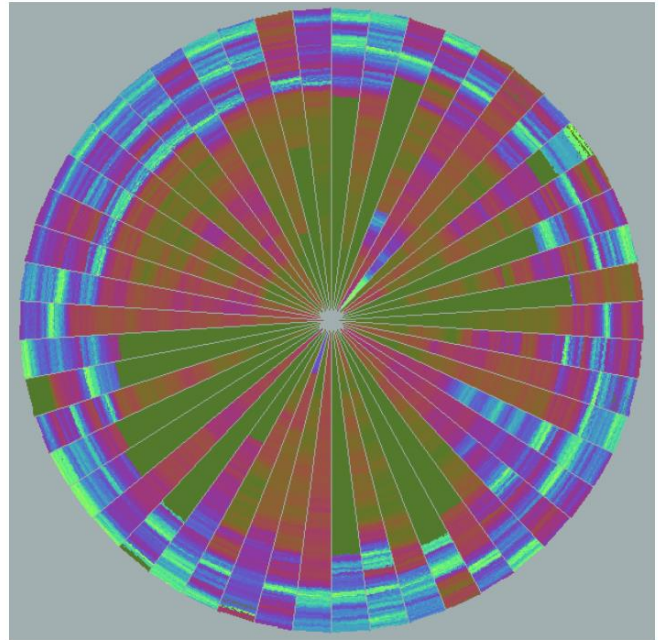


Figure 5: Representing about 265,000 50-dimensional Data Items with the 'Circle Segments' Technique

Fig. 13. Representing about 265,000 50-dimensional data items with the "Circle Segments" technique [24].

CS is not restricted to present time-series data, following case studies of using CS on Neural Networks (NNs) building are the examples. NNs are widely used sets of algorithms that generate simple models of the animal brain to recognize patterns. The most important property of NNs is the ability to learn from previous data, and apply the learned knowledge on the new data sample to give predictions. However, NN acts like a black box without giving information about how the prediction is generated. To further

improve this approach and make it more convincing, it is important to include domain users in the data exploration and analysis process. The selection of input features of a NN can affect its performance in various aspects, and the number of input features is an essential factor. The performance of a NN model is directly related to the number of input features used. On the other hand, redundancy of irrelevant inputs can deteriorate the performance of the NN model due to the “curse of dimensionality” problem, which states that the performance of a NN model reduces after the number of inputs exceeds a certain point. Lim et al. try to analyze the effect of the input features to the target outputs and select the significant input features to build NN models by introducing visualization [25]. They propose a coupling of CS and the most widely used NN architecture, the multiplayer perceptron (MLP). This proposed approach is applied to three case studies for evaluations.

In this approach, the CS layout design is divided into three stages: dividing, ordering and coloring. In the dividing stage, the circle is divided into the number of total inputs and outputs. In the ordering stage, it is important to make sure the ordering shows the correlations between the input features and the output classes. The inputs are ordered around the circle based on their magnitudes of correlation to the outputs. The ordering stage also includes the ordering of different data samples, which is the major difference between the CS layout in this application and the original CS layout mentioned above. The original CS is filled with time-series data from the center to the outside, which limits the CS to depict only one data sample at different times. In this CS layout, data are not time referred, and the space is filled by lines across each segment. Each line represents the value of the corresponding input or output, and each data sample is encoded as a regular polygon. The thickness of lines is depended on the total number of data samples. With a large number of data samples, the lines are compressed to width of one pixel. In other words, this CS layout is not completely pixel oriented but still remains a CS displaying structure. The ordering of data samples is first determined by the output value. Data with the highest output value is encoded as the perimeter and data with the lowest output value is encoded at the center. If the data have the same output value, the ordering is then determined by the value of the input with the highest correlation to the outputs. If there is still a tie, then the values of the second-high correlation input are compared. This rule is followed until all data samples are sorted properly. In the coloring stage, each data value is normalized between 0 and 1, and a colormap is used to encode the magnitude of each value.

The three case studies are the wire electrical discharge machining (WEDM) process, the Wine data set, and the stroke patient records. The objective of the WEDM case study is to test the ability of using CS to find out the effects of input features to target outputs. By comparing users’ observations from the CS and the main effects plot, it proves that users can conclude correct effectiveness information of inputs towards outputs using the CS diagram. For the Wine data set and the stroke patient records case studies, the objective is to see the improvements of using CS to select significant input features on the performance of the MLP network. Comparisons are made between three different methods: MLP without feature selection (referred as MLP), MLP with CS feature selection (referred as CS-MLP), and MLP with principal component analysis for feature selection (referred as PCA-MLP). In both cases, CS-MLP has the highest accuracy and lowest standard deviation. All three case studies support the usefulness of CS in selecting and analyzing the effects of the input features to the target outputs when building a NN model.

5.1.2 Circle View

Circle View (CV) is a radial visualization technique specified in analyzing fast changing time dependent data streams [26]. To avoid confusion caused by the overwhelming data information, CV presents only aggregated information to help analysts at the initial

overview data analysis stage. It combines the radial layout and the hierarchical display technique to present time referenced multidimensional data.

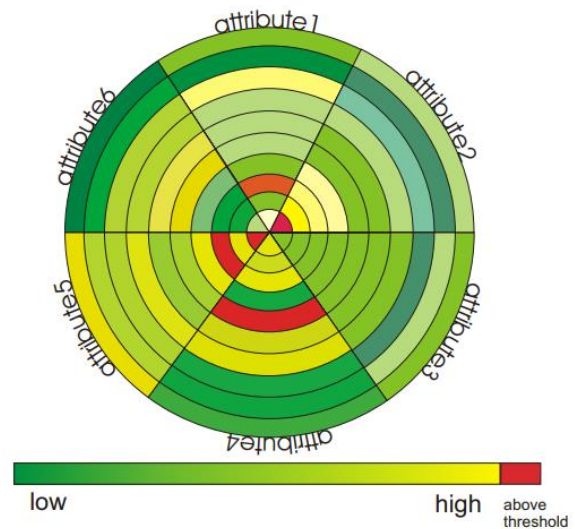


Fig. 14. Circle View showing the evolution of multiple attributes over time [26].

As shown in Fig. 14, the basic idea of CV is to use circular segments to represent data dimensions and radius of the circle to represent the time dimension. The circle is divided into a number of segments based on the number of dataset dimensions. These segments are further divided by concentric circles with different radiuses representing the propagation of time. The direction of such propagation can either start at the center point toward the edge or vice versa. The change in radius of the concentric circles is the time interval, also referred as to the time slot, which its unit can be determined by the user based on the application scenario (second, minute, hour, etc.). The level of the time interval can be selected interactively when the corresponding data is available. Each cell in the CV would represent the aggregated value of an attribute over a certain period of time. This value is encoded in colors to identify its magnitude. The ordering of attributes is computed based on similarity, but changing the ordering manually is possible which helps to compare attributes that are far apart. By using CV, it is easy to compare the data values in neighbor attributes at a certain time. It is also convenient to compare data values in one attribute from neighbor time slots, and observe the changes of the value over time. Although CV presents historical data and the current time data at the same time, people usually find the current time data more important. CV emphasizes the current time data by adjusting the space of each time slot. The change of radius of the concentric circles is no more constant but proportional to the time propagation. Depends on the user defined time direction, CV now presents as concentric circles with either increasing radius or decreasing radius. When using CV to visualize continuous data streams, the time slots are shifted to the center or to the edge depending on the preset time propagation direction. As shown in Fig. 15, information with the new time slot is added to the plot with the oldest time slot disappeared. One way to interpret this transformation is to see it as a sink or source of fluid flow. The user can choose a start and end time of the dataset, and the update of CV is presented by animation. The user can also choose the update time to control the speed of animation. An extension of multiple views is also available for CV, which is called Multi-Circle Views. Multi-Circle Views can be applied to view different data with the same attributes, or different combinations of attributes from one data. Such visualization is useful in finding the local patterns in different groups of attributes. When using Multi-Circle Views to

visualize different groups of attributes, each CV has a different diameter representing the weight of the group of attributes.

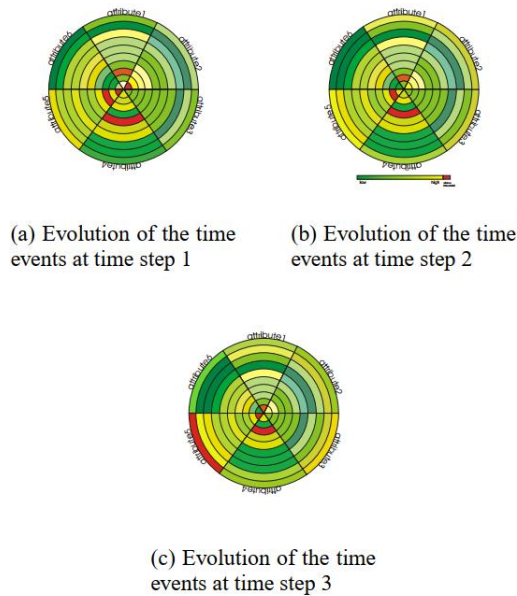


Fig. 15. Circle View visualize a continuous data stream over time [26].

Mariano et al. introduce CV to the field of Phenology [27]. Phenology is a science that studies the periodic life cycles of plants and animals and their relations to the environment. Mariano et al. create a novel visualization approach by combining CV with visual rhythms to present phenological data. Visual rhythm is a methodology that samples the images from a video using different criteria and combines them into one image that shows the changes of the video over time. In this paper, the researchers apply the visual rhythm approach on the time series tabular data derived from on-the-ground phenology observations. Instead of sampling images from a video, researchers are sampling small tuples of table from the large phenology observations data table. Each tuple of table contains all observations data in a certain time, analogic to one frame in a video. The radial visualization idiom is a modified CV. The modifications are mainly two parts. This idiom still uses concentric circles to represent time attributes, such as years, months, or days. However, this idiom utilizes the inherent cyclic property of the radial visualization. Instead of using segments to assign other dependent attributes, they are used to represent the periodic property of the time attributes. For example, when days are assigned to the concentric circles, the circle can be divided into twenty-four segments with each one represents one hour. Each segment is further divided into sub-segments as the nominal attributes, including individuals, species, phenophases, or weather sensors. Comparing to the original CV, where each attribute appears only on one segment, the modified CV has attributes that appear various times around the circle. Second, they leave an empty hole in the center, so the innermost information is not too small to observe. The hole can also be potentially used to display some extra information. Other than these minor changes, this idiom follows the same design as the CV. Propagation of time, in the unit of years, is encoded as concentric circles from inside out. Colors are used to represent the magnitude of data values in cells.

In the testing and evaluation stage, the design team proposes seven prototypes for detailed visualization, which visualize individual-related data; and five prototypes for summarized visualization, which visualize species-related data. The general encoding design is the same for every prototype. Concentric circles are assigned to years, and the whole circle is divided into twelve major segments representing twelve months of the year. The major

segments are further divided into sub-segments representing phenophases. Each cell represents the intensity of the corresponding phenophases, denoted by a score of 0, 1, or 2. For detailed visualization, the score is a direct measurement of phenophases' activity. 0 is absence, 1 represents a low activity between 0% to 50%, and 2 represents a high activity between 50% to 100%. For summarized visualization, the value in each cell represents the percentage of individuals from a particular species that have a phenophase intensity score above 0. In one of the prototypes, climate data of precipitation is presented at the center of the CV. In other prototypes, the precipitation is assigned to an extra sub-segment. The differences between prototypes are in their labels, legends, and colormaps. According to testers' feedbacks, Prototype 3 (Fig. 16) for individuals and Prototype 2 and 3 for species received the highest rating on analyzing complex phenology data. Based on the test results and feedbacks, the research team concludes two potential improvements to the current design. The two types of climate data encodings should both be reserved as options for users depending on different research objectives. In addition, the visualization idiom should include an interaction mechanism for the user to choose the colors used in the colormap.

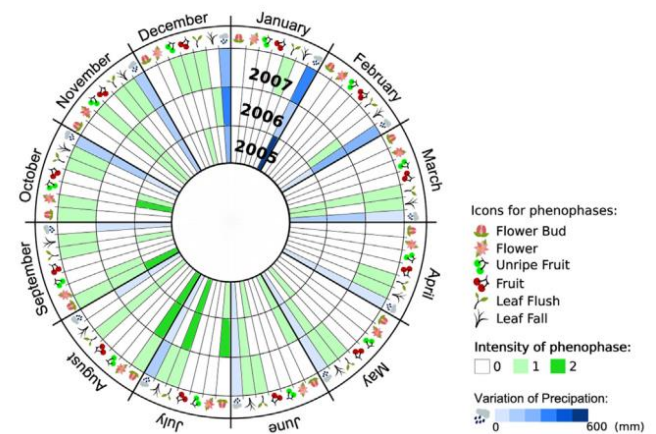


Fig. 16. Prototype 3, phenophases represented by icons [27].

Based on the previous studies and researches, Mariano et al. create a visualization tool named as RadialPheno [28]. While the last paper focuses on visualizing on-the-ground observations data, this paper pays more attention to camera-derived near-surface remote data. Digital cameras called phonecams are installed at near ground level or at the top of towers. Sequential images are taken by these cameras. When analyzing these images, researchers define the region of interest (ROI) from the images. For each ROI, vegetation indices are extracted from the digital images. One possible index is the green chromatic coordinate (Gcc) index, which describes the normalized green amount in a digital image. These data are generated in a CSV file and can be directly inputted into RadialPheno. Users can manually enter the on-the-ground observations data to a CSV file and input it into RadialPheno as well. Therefore, RadialPheno can be used to either analyze and compare the on-the-ground data and the camera-derived together, or present and analyze these data separately. RadialPheno provides interactive filtering base on the attributes in the CSV file. The filtering options include species, years, and periodicity. There are many differences between RadialPheno to the previously built prototypes. The biggest one is that sub-segments are no longer used for nominal attributes. Therefore, RadialPheno is restricted to represent data of one specie over years, or data from different species within one year, as shown in Fig. 17. Users can switch plots of different species or different years by using the previous (<<) and next (>>) buttons. The sub-segments are now assigned to periodicity. A more detailed time data will have finer sub-segments. When the periodicity is select to be monthly, only the major segments appear on the graph. RadialPheno

has some limitations. It is hard to identify data presented in the inner circle as the size of segments is constrained by the radius. Also, additional interaction mechanisms are needed for data selection based on database-oriented technologies and selection of visualization settings.

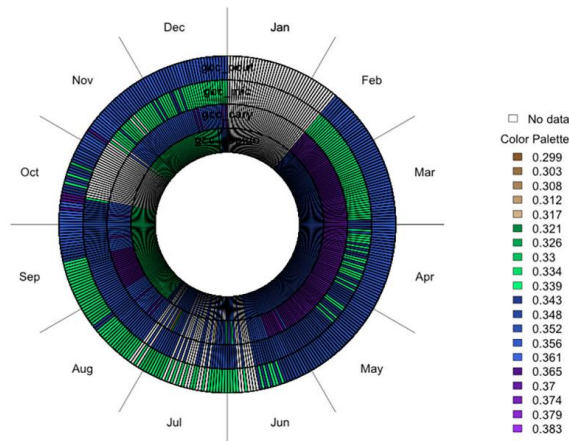


Fig. 17. RadialPheno visualizing daily Gcc values during the year 2012, observed for individuals of four species [28].

6 OTHERS

6.1.1 VisAxes

VisAxes is a visualization framework that consists of two novel radial visualization techniques – TimeWheel and MultiComb [29]. Both techniques use the same design of axes. Each axis is corresponded to one variable and scaled from the variable’s minimum to maximum. What makes the axes design distinctive is its interactivity. VisAxes creates three types of axis depends on the characteristic of the variable: scroll axis, hierarchical axis and focus within context axis. The scroll axis is mainly used for variables with a large number of values. It combines a slider to the axis and users can set the start and the end of the slider to choose the size of interest. Users can further drag the slider along the variable domain to switch the range of interest. The hierarchical axis is used for hierarchically structured variables. The hierarchical axis is first divided into segments with the number of nodes in the first level (children of the root). Selecting one segment and that segment is further divided into sub-segments of corresponding children nodes. Animation is used in this axis variation process to avoid visual discontinuity. The focus within context axis is also used for variables with a large number of values. It is different from the scroll axis because it always maps the entire variable range, and creates a focus within a range of interest while remaining the context values on the axis. Users can interactively change the focus on the axis. The fundamental concept of both TimeWheel and MultiComb is they both set an axis as the reference axis representing an independent variable. Time would be the most common independent variable, but any other appropriate variables can be applied. In order to distinguish between data from different variables, depended variables are assigned to different colors.

For TimeWheel, the time axis is placed at the center of the plot and circularly surrounded by other depended axes to form a regular polygon shape. The line is connected from the variable value to its time value. From Fig. 18, it is clear that the relation between the dependent variable and the time is easier to observe when the variable axis is parallel to the time axis. Therefore, TimeWheel provides interactive rotation for move the axis to the desired position. On the other hand, when axes are not parallel to the time axis, the connected lines become crowded. To reduce clutter, the TimeWheel uses coloring fading based on the angle of lines. It also automatically

adjusts the length of value axes, so the information on axes that are interested (parallel) is amplified, and the less important information on other axes is weakened. The TimeWheel can be considered as a transformation of Parallel Coordinates to present multidimensional time referenced data.

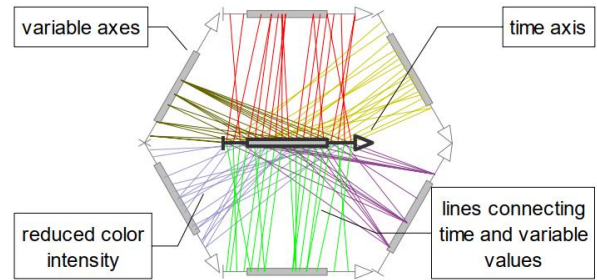


Fig. 18. A TimeWheel. Six variable axes are arranged circularly around an exposed centered time axis [29].

MultiComb is a combination of multiple plots for every dimension. As shown in Fig. 19, it has two possible arrangements. In the first arrangement, the time axes are arranged radially with value axes pointing outward. In the second arrangement, the value axes are arranged radially with time axes spreading out. Like TimeWheel, MultiComb also provides interactive rotation for users. The arrangement of the MultiComb naturally creates a blank space in its central area, which is available to present additional information. The paper suggests using a spike glyph to fill this blank hole. The spike glyph is useful in direct data value comparison and to provide a history of past data value. For value comparison, the selected data value is represented by the length of the corresponding spike, which has the same color as the data dimension and points in the direction of that dimension. The arc at the end of each spike acts like a gridline as mentioned in RC and helps with comparison. For the aggregated history view, spikes are extended fully, and a certain number of arcs are plotted on each spike representing aggregated values. For each arc a specified number of values are aggregated, and the aggregated value is encoded as the angle of the arc. The aggregated history view gives the user an idea about the pattern of values in the last time period.

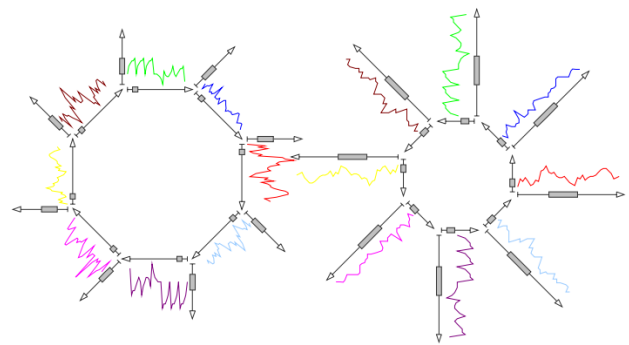


Fig. 19. The MultiComb. Left: The axes of depending variables extend outwards from the center. Right: The axes of reference extend outwards [29].

Based on testing, VisAxes has a scalability of up to 15 axes. However, the analysis of data becomes more difficult once it reaches 10 axes. TimeWheel and MultiComb are most suitable to compare different time depended variables and try to find their correlations. Like other radial visualization techniques presented in previous sections, the arrangement of the variable axes affects the performance of revealing data patterns, tendencies and correlations. For the future work, the research team intend to arrange the axes based on similarity or entropy, and compare the effectiveness of different arrangement using computer human interaction study.

6.1.2 Wind Rose

The history of wind rose goes back a long time, even before the creation of the compass. In the 13th century, Italian and Spanish sailors already included the wind rose in their map to show the directions of the eight major winds [30]. Later on, with the appearance of the compass, the wind rose is combined with it to be the compass rose or the compass card. However, the early historical wind rose has no distinction between the cardinal directions and winds coming from these directions. The fairly simple configuration of the historical wind rose provides very little information and does not need a detailed analysis. Therefore, we only focus on the design and application of the modern wind rose in this section.

The modern wind rose (hereinafter referred to as wind rose) presents the summarized information about the wind at a particular location over an observation period. The summarized information often includes the directions, speeds and frequencies of winds. Here the frequency is interpreted as the percentage of time of the wind coming from a certain direction. The wind rose uses the angel channel to display cardinal directions, just like the compass. A typical wind rose has 16 cardinal directions, but the resolution can go higher to 32 or even more based on users' demand. Sharp triangular-shaped ray radiates out from the center and points to the direction which the wind is coming from. The length of each ray indicates the frequency of the wind, usually with concentric gridlines to help identifying the actual value. In some applications, the length can directly represent the total time (hours, minutes, etc.) of the wind blowing from the specified direction. A segmented colormap is applied in these rays to show the speed of winds, with each color represents a range of speed. Similar to the length of the rays, the length of each color segment in a ray is proportional to the speed frequency of the wind in the corresponding direction. A common trick in the wind rose is to display the frequency of calm (zero speed) or nearly calm air at the center. It does not only save a color channel for the speed frequency, but also provide an indentation so the information at the center of the wind rose is less crowded. Fig. 20 shows an example of a wind rose. The wind rose is useful in various fields, including sailing, architecture, wind farms and airport runways constructing [31].

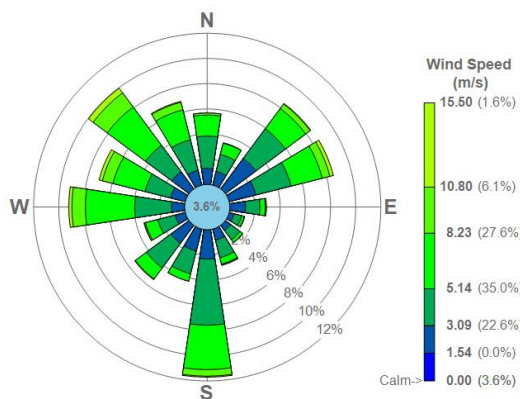


Fig. 20. Wind rose plot for LaGuardia Airport (LGA), New York, New York, 2008 [32].

6.1.3 Sun Path Polar Plot

The sun path polar plot presents the path of the sun across the sky over a day as seen from a given location. Due to the nature of earth orbit, any location at the same latitude has the same sun path [33]. Assuming the earth to be a perfect sphere and projecting the 3D sphere on a 2D canvas with the observer at the center point. Angel is used for cardinal directions, in this case the azimuth, starting at north as 0/360 degree and going around to the east of north direction. Radius represents the altitude, with outer perimeter as 0 degree

(horizon) and the center point as 90 degree (zenith). The horizontal arc across the circle is the path of the sun on a specific day. A sun path polar plot integrates the sun paths over a year using multiple horizontal arcs, with lines vertically cross these paths indicating the particular solar time of the day. Date and time are labeled right next to the corresponding arc respectively. Noticing only half a year's sun paths are plotted, that is because the remaining half has the exact same paths due to the orbital revolution of the earth. It is easy to locate the position of the sun at the specific date and time from the sun path polar plot. Moreover, it is readily to conclude the pattern and tendency of the sun path. For example, Fig. 21 presents the sun paths in Rotterdam. It experiences its longest daytime in June of around 16 hours, and the shortest daytime in December of around 8 hours. Interactive sun path polar plot is also possible using computer programming languages. A very good example is [34] based on D3, which provides interactive selections of latitude, longitude, time zone, date of the year and time of the day. It is also different from the traditional sun path polar plot by presenting the local standard time on the plot with more accurate hour lines.

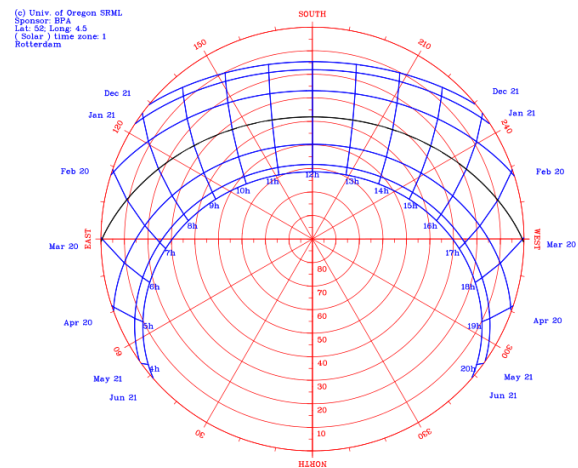


Fig. 21. Sun path polar chart for any location at the latitude of Rotterdam [35].

7 CONCLUSION

This paper proposes a survey on visualizing multivariate data in a radial layout. We categorize these techniques into two major types: axis-based and segment-based. In the axis-based category, data attributes are assigned on the axes of the circle. In the segment-based category, the circle is divided into a number of segments, and each segment will represent one data attribute. We found some radial visualization techniques being more complex that cannot be simply classified into these two categories, and we put them into the others category. When applying axis-based radial visualization, researchers are usually trying to compare these multivariate data, look for potential clusters, patterns, outliers, etc. We include Radar Chart, RadViz and Star Coordinates for axis-based radial visualization. Segment-based radial visualization is mostly applied, but not limited, to visualized time referenced multivariate data. We include Circle Segments and Circle View for segment-based radial visualization. In the others category, we introduce VisAxes, Wind Rose, and Sun Path Polar Chart. Going through these papers and articles, all techniques included are proved to be effective in multivariate data analysis. Therefore, this survey confirms the validity of radial visualization being a noteworthy class of techniques to perform multivariate data analysis. However, no technique outperforms others in every single aspect, and radial visualization is not the optimum solution to visualize multivariate data. Users should carefully select their method to analyze multivariate data based on different tasks and objectives.

REFERENCES

- [1] G. M. Draper, Y. Livnat and R. F. Riesenfeld, "A Survey of Radial Methods for Information Visualization," *IEEE Trans. Visualization and Computer Graphics*, vol. 15, no. 5, pp. 759-776, 2009.
- [2] S. dos Santos and K. Brodlić, "Gaining Understanding of Multivariate and Multidimensional Data through Visualization," *Computers & Graphics*, vol. 28, no. 3, pp. 311-325, 2004.
- [3] T. Munzner, *Visualization Analysis and Design*, Florida: A K Peters/CRC Press, 2014.
- [4] P. Hoffman, G. Grinstein, K. Marx, I. Grosse and E. Stanley, "DNA Visual and Analytic Data Mining," in *Proc. Eighth Conf. Visualization*, 1997.
- [5] X. He, Y. Tao, Q. Wang and H. Lin, "Multivariate Spatial Data Visualization: a Survey," *Journal of Visualization*, vol. 22, no. 5, pp. 897-912, 2019.
- [6] A. Maalej, N. Rodriguez and O. Strauss, "Survey of Multidimensional Visualization Techniques," in *Computer Graphics, Visualization, Computer Vision and Image Processing (CGVCVIP)*, 2012.
- [7] W. Y. Chan, "A Survey on Multivariate Data Visualization," *Dep. Comput. Sci. Eng. Hong Kong Univ. Sci. Tech.*, vol. 8, no. 6, pp. 1-29, 2006.
- [8] S. R. dos Santos, "A Framework of the Visualization of Multidimensional and Multivariate Data," Ph.D. dissertation, School Comput., Univ. Leeds, Leeds, 2004.
- [9] K. W. Kolence and P. J. Kiviat, "Software Unit Profiles & Kiviat Figures," *ACM SIGMETRICS Performance Evaluation Review*, vol. 2, no. 3, pp. 1-12, 1973.
- [10] G. Odds, "A Critique of Radar Charts," *Scott Logic*, 23 9 2011. [Online]. Available: <https://blog.scottlogic.com/2011/09/23/a-critique-of-radar-charts.html>. [Accessed 17 11 2019].
- [11] T. Leary, S. Ridley, K. Burchet, K. Kong, P. Crispin and M. Wright, "Assessing Critical Care Unit Performance: a global measure using graphical analysis," *Anaesthesia*, vol. 57, no. 8, pp. 751-755, 2002.
- [12] R. Feldman, "Filled Radar Charts Should not be Used to Compare Social Indicators," *Social Indicators Research*, vol. 111, no. 3, pp. 709-712, 2012.
- [13] M. M. Porter and P. Niksiar, "Multidimensional Mechanics: Performance Mapping of Natural Biological Systems Using Permutated Radar Charts," *PLoS One*, vol. 13, no. 9, pp. 1-18, 2018.
- [14] M. Rubio-Sanchez, L. Raya, F. Diaz and A. Sanchez, "A Comparative Study between RadViz and Star Coordinates," *IEEE Trans. Visualization and Computer Graphics*, vol. 22, no. 1, pp. 619-628, 2016.
- [15] T. V. Long and V. T. Ngan, "An Optimal Radial Layout for High Dimensional Data Class Visualization," in *Int. Conf. Advanced Technologies for Communications (ATC)*, 2015.
- [16] M. Ankerst, S. Berchtold and D. A. Keim, "Similarity Clustering of Dimensions for an Enhanced Visualization of Multidimensional Data," in *Proc. IEEE Symp. Information Visualization (InfoVis)*, 1998.
- [17] K. Daniels, G. Grinstein, A. Russel and M. Glidden, "Properties of Normalized Radial Visualizations," *Information Visualization*, vol. 11, no. 4, pp. 273-300, 2012.
- [18] S. Cheng, W. Xu and K. Mueller, "RadViz Deluxe: An Attribute-Aware Display for Multivariate Data," *Processes*, vol. 5, no. 4, pp. 1-18, 2017.
- [19] L. d. C. Pagliosa and A. C. Telea, "RadViz++: Improvements on Radial-Based Visualizations," *Informatics*, vol. 6, no. 2, pp. 1-23, 2019.
- [20] E. Kandogan, "Star Coordinates: A Multi-Dimensional Visualization Technique with Uniform Treatment of Dimensions," in *Proc. IEEE Information Visualization (InfoVis)*, 2000.
- [21] Y. Sun, J. Tang, D. Tang and W. Xiao, "Advanced Star Coordinates," in *The Ninth Intl. Conf. Web-Age Information Management*, 2008.
- [22] D. J. Lehmann and H. Theisel, "Orthographic Star Coordinates," *IEEE Trans. Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2615-2624, 2013.
- [23] D. A. Keim, "Pixel-Oriented Visualization Techniques for Exploring Very Large Data Bases," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 58-77, 1996.
- [24] M. Ankerst, D. A. Keim and H.-P. Kriegel, "'Circle Segments': A Technique for Visually Exploring Large Multidimensional Data Sets," in *Proc. Visualization, Hot Topic Session*, 1996.
- [25] C. P. Lim, S. L. Wang, K. S. Tan, J. Navarro and L. C. Jain, "Use of the Circle Segments Visualization Technique for Neural Network Feature," *Neurocomputing*, vol. 73, no. 4, pp. 613-621, 2010.
- [26] D. A. Keim, J. Schneidewind and M. Sips, "CircleView: A New Approach for Visualizing Time-related Multidimensional Data Sets," in *Proc. Advanced Visual Interfaces (AVI)*, 2003.
- [27] G. C. Mariano, V. G. Staggemeier, L. P. C. Morellato and R. d. S. Torres, "Multivariate Cyclical Data Visualization Using Radial Visual Rhythms: A Case Study in Phenology Analysis," *Ecological Informatics*, vol. 46, pp. 19-35, 2018.
- [28] G. C. Mariano, B. Alberton, L. P. C. Morellato and R. d. S. Torres, "RadialPheno: A Tool for Near-surface Phenology Analysis through Radial Layouts," *Applications in Plant Sciences*, vol. 7, no. 6, pp. 1-5, 2019.
- [29] C. Tominski, J. Abello and H. Schumann, "Axes-Based Visualizations with Radial Layouts," in *Proc. ACM Symp. on Applied Computing (SAC)*, 2004.
- [30] B. Thoen, "Origins of the Compass Rose," GISNet, 2001. [Online]. Available: <http://www.gisnet.com/notebook/comprose.php>. [Accessed 25 11 2019].
- [31] K. C. Heidorn, "By Any Other Name: The Wind Rose," *The Weather Doctor's Weather Almanac*, 2005. [Online]. Available: <https://www.islandnet.com/~see/weather/almanac/arc2005/alm05mar.htm>. [Accessed 25 11 2019].
- [32] Anonym, "Wind Roses and Plotting," BREEZE Software, 2019. [Online]. Available: <https://www.breeze-software.com/Software/MetView/Product-Tour/Wind-Roses-and-Plotting/>. [Accessed 2 12 2019].
- [33] V. Khavrus and I. Shelevytsky, "Introduction to Solar Motion Geometry on the Basis of a Simple Model," *Physics Education*, vol. 45, no. 6, pp. 641-653, 2010.
- [34] A. Marsh, "2D Sun-Path," ANDREWMARSH.COM, 2014. [Online]. Available: <http://andrewmarsh.com/apps/releases/sunpath2d.html>. [Accessed 26 11 2019].
- [35] Anonym, "Sun Path Chart Program," University of Oregon, Solar Radiation Monitoring Laboratory, 2010. [Online]. Available: <http://solarlat.uoregon.edu/SunChartProgram.html>. [Accessed 2 12 2012].