

Multivariate Hybrid Visualisation of Ornithological Sensor Data

Richard C. Roberts, Robert S. Laramée & Mark W. Jones

Department of Computer Science, Swansea University, Swansea, UK

Abstract

To overcome the challenges of displaying multivariate sensor data, we propose a novel work-in-progress, hybrid, polar method of visualisation. Sensor data is collected by marine biologists in high volumes and using multiple sensors. Challenges arise when trying to unlock the marine wildlife behaviour from the vast amount of time series data collected. The proposed method filters uninteresting behaviour and isolates the features of interest within the set. Multi-layer polar plots are used to visualise local pressure, temperature, temporal behaviour and energy expenditure, all of which are given upper and lower plotting ranges to ensure no overlap. This results in a feature centred visualisation that focuses on the most important behaviour. The value in this method is that the visualisation can show many instances of the chosen activity. Each animal can be examined individually, or multiple animals and behaviours can be compared side-by-side for the first time. An interactive software system enables the user to navigate such that individual instances of the marine wildlife behaviour can be studied at high resolution or the user may choose an overview of every animal. This paper uses ornithological sensor data to demonstrate the proposed visualisation. Although it can be applied to other multivariate data sets that are linked with a temporal dimension.

1. Introduction Motivation

A fundamental challenge when attempting to visualise multivariate time-series data is one of resolution. Creating a visualisation that encompasses an too many attributes in one image may result in an illegible visualisation. Overviews can be presented but high resolution analysis may be beyond the limitations of traditional data visualisation. Euclidean space visualisations can be created but are still limited, at most, to three spatial dimensions. A fourth dimension of time can also be utilised.

Traditionally, parallel coordinates are used to visualise multi dimensional data but may be vulnerable to overplotting. There comes a point in a parallel coordinate visualisation that adding new data reduces the clarity (overplotting [GTC01]). Moving beyond this requires creative use of design and a firm knowledge of the dataset to develop a custom visualisation.

The linear nature of the parallel coordinate can be complemented by a method based on a polar coordinate system that plots around a central axis as opposed to a line [DLR09]. This paper presents an extension of these methods by adding two-level system to a radial polar plot. The proposed layering system reduces clutter. A feature centred approach and

the utilisation of the temporal dimension is a key factor in simplifying a large dataset into a more manageable size.

In this work-in-progress paper, we present a novel method for handling and visualising a multivariate data set using a custom made hybrid visualisation and an interactive piece of software. To create the visualisations we used an ornithological data set that spans eight hours using 14 different micro sensors. The species of wildlife used to collect the data is the sea cormorant, a diving sea bird that can spend several minutes under water hunting for food. Our visualisation looks at the data from a biological viewpoint and we have designed the visualisation accordingly. This paper provides the following contributions:

- a hybrid, multi layer polar plot visualisation of multivariate, time-series data.
 - a novel method to compare multiple animals side-by-side.
 - The application of our high-dimensional visualisation to real-world data for the study of marine wildlife behaviour.
- The rest of the paper is organised as follows:
- Related work
 - Visualisations and software design
 - Conclusions and further work

2. Related Work

This paper presents a visualisation for ornithological, time-series, sensor data in order to study marine wildlife behaviour.

Multi Dimensional Data Ward et al. categorise multi dimensional visualisations into four main groups. Point based techniques, line based techniques, region based techniques and a combination of techniques [WGK10].

A popular line based method of high dimensional visualisation is parallel coordinates which plots each dimension on its own axis. This enables large amounts of data to be in view at any one time [Weg90] [BBP08] [GPL*11]. The benefits of this technique is the ability to handle any number of data dimensions, however when the data set becomes too large, detail may be lost due to overplotting.

A solution to the issue of overplotting was suggested Artero et al. where clusters of data could be made more prominent [AdOL04] [GPL*11]. This validates the use of parallel coordinates for high dimensional data in the context of data overview, but it still does not enable the viewer to look at the individual data points or run a direct comparison of two 'non parallel' dimensions.

To enable a more direct comparison of the data dimensions, many solutions have been presented. Hyper-slice and Hyperbox [vWvL93] [AC91] present a side-by-side comparison of n dimensional data sets using a matrix of individual but related graphs. These scatterplot matrices overcome overplotting by comparing each dimension against each other but in compartmentalised boxes [CLNL87] [VMCJ10]. Many solutions focus on the adaptation of the traditional axis to accommodate more dimensions or allow more data to be mapped before overplotting occurs [CVW11].

Time-series Data Multivariate time series data presents a special challenge to visualisation. Although it is beneficial for the time variable to calibrate the remaining dimensions with one another, the specific visual requirements of that dimension may place limits on the methods of visualisation. Time series data may require a continual, constant transition axis, presenting difficulties for scatterplots and other point plotting methods. See Aigner et al. [AMST11] for a complete overview of the time-series visualisation techniques. The time series axis can be linear [VWVS99], where the data is plotted on a single spacial dimension axis. This method enables an easy comparison of multiple variables but can use a large amount of display space.

An alternative to the linear method of visualising is the radial method. These involve a circular plot that can either map data points to multiple axis' radiating from a central common point, or have a time series data set plotted around that point [DLR09]. Tominski & Schumann developed a spiral display that visualises cyclic nature of a data set [TS08].

Particularly useful in seasonal data, this type of visualisation allows comparison of time-series data across events. Alexa et al. developed upon this spiral data visualisation to allow a day-by-day comparison of time series data [WAM01]. Zhao et al. used this concept to add layers to a time series plot that can be displayed at a spiral [ZCB11].

Whitfield & Cannon implemented a polar line plot using seasonal time series data where the months are labelled around the edge of the polar plot and the seasons are separated by equal distances [WC00]. The benefit of this method is that lines can be overlaid on the plot showing different years of data, allowing for a direct comparison across time frames.

Multivariate time-series data The same data has previously been visualised using parallel & spherical coordinates [WGJL] [GPL*11], also using 3D scatterplots and spherical histograms [GJL*09]. Walker et al. took a more holistic approach to the data and visualised the full set but did not take into account the time series nature of the data. Walker et al. created a parallel coordinate hybrid visualisation that integrated spherical coordinates [WGJL]. The combination of the two methods enables the viewer to see the data from multiple coordinate systems. Grundy et al. isolated individual instances of the features found within a single animal manually and visualised those. The holistic approach does successfully visualise the data so that the complete set can be seen. Grundy et al. produced a visualisation for the biologists that the data is valuable to, but loses some of the context due to the data filtering. One key ability that distinguishes this work from its predecessors is the ability to compare animal behaviour amongst an entire group, instead of just focusing on an individual.

3. Multivariate Hybrid Ornithological Visualisation

Here we present a custom-made hybrid visualisation designed to accommodate the multivariate, large data set consisting of temporal sensor readings for the study of ornithological data. The design goal is the creation of a simple dynamic that can be used as a template for comparison of multiple animals and multiple behaviour instances.

Visualisation Design In order to focus on the data at a high resolution we have isolated the features in the set using the pressure sensor values and discarded any activity recorded between low pressure events. This was done by noting the change in pressure value to indicate when the bird has entered the water. As the pressure value rises, the birds depth increases. This approach of data filtering has the resulting effect of high visibility for the most useful aspects of the data, and thus addresses the challenge of data resolution. Given the dataset is a collection of sensor data involving a diving sea bird, the features of interest are the bird dives themselves.

Figure 1 shows a schematic diagram of the proposed visualisation which is centred around two layers of polar plotting. An inner layer, r , and an outer layer, R . The inner layer

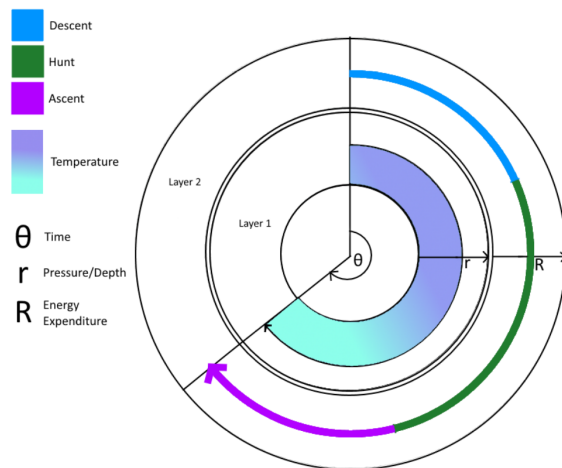


Figure 1: The inner layer, r , of the visualisation shows an area polar plot that rotates around the central axis representing the pressure sensor values. The outer layer's line polar plot represents the 'VeDBA' (Energy Expenditure), R , of the animal. The colour of the first layer changes according to the temperature sensor readings and the colour of the second layer changes depending on what the current stage of dive is (Descent, Hunt, Ascent).

r , graphs an area polar plot mapped to the pressure sensor values. The outer layer, R , graphs a polar line plot that represents the energy expended by the bird and its current behaviour.

The θ value represents the passage of time which enables two data attributes to be directly compared over time. The maximum θ value is scaled according to the largest dive in the set which maps to 360° .

The benefit of this proposed method is the reduction of overplotting as the two layers of information are separated but can be directly compared. Other attributes of data are brought into the visualisation using colour. The area plot is coloured according to the temperature sensor and the energy expenditure line plot changes colour based on the birds current stage of dive. The dive is broken down into three stages; descent, hunt, and ascent.

Pressure and Depth Local pressure is mapped to the radius, r , of the inner polar area plot. The pressure value is inherently linked with the depth the bird dives. A comparison of two separate dives through this aspect of the visualisation shows, not only the length of the dive mapped to θ , but how deep the bird descends to hunt for prey. As well as pressure, r shows the environmental temperature using the colour of the area plot. Thus the temperature gradient can be observed at different depths and times and any relationships or trends can be seen in the visualisation.

The area polar plot rotates around the centre of the visualisation. The base of the plot is not directly at the centre of

the circle as to create a lower bound to ensure all dives are visible at a reasonable resolution. This first layer will never overlap the second as the dimensions are scaled such that the deepest dive is mapped to the maximum radius r just short of base line for the next layer of visual space.

The uniformity of r in a visualisation indicates that the bird descends to a certain depth to hunt for fish and does not stray from that depth until it ascends back out again. All recorded dive instances show only slight variations in r , but all depict a uniform curve at the hunt stage of the dive. This suggests the bird descends to an instinctive depth where it is known food is typically found and remains at that depth until food is found or they need to resurface for air.

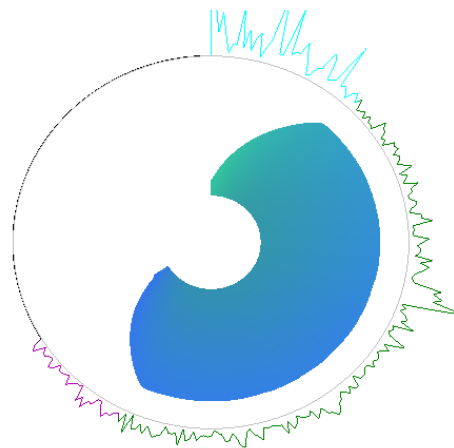


Figure 2: Single instance of a dive: Starting at the 'midnight' position, the visualisation plots the data clockwise around the centre. Here you can see the energy expenditure spike half way through the dive; potentially at the point the bird catches a fish.

Energy Expenditure To display the energy expenditure of the bird a polar line plot is used, starting at the midnight position and plotting data clockwise to show temporal progression. The line is split into three colours denoting the three stages of a dive; descent, hunt, and ascent. This division enables us to go beyond the raw data by dynamically pinpointing events and features within it.

This polar line plot, again, revolves around the centre point of the graphic but occupies the outer ring layer. The radius R of the line plot from the second layer base line measures the amount of energy expended by the animal (Calculated using [GWS11]). The two layers can be directly compared as they are synchronised by the θ value. This way you can directly see the relationship between energy expenditure and their dive activity, alongside other dive attributes such as depth and temperature.

Energy expenditure is calculated by rooting the summed square values for each accelerometer axis [GWS11],

[GHWF09]. This ‘VeDBA’ represents the Vectorial Dynamic Body Acceleration and directly corresponds to the energy expended by the bird. As well as the VeDBA, the three components of the dive are extracted using the pressure sensor values. From this we can see the value of multi dimensional data collection where multiple sensors are being used to create one hybrid visualisation.

Software Design The software allows the user to navigate through the complete data set using a graphical slider that visualises the features within the data. In figure 4 (bottom) you can see the horizontal scroll bar that moves the visual bar through the data. A blue bar shows descent activity, green shows the birds hunt, and purple shows the ascent back out of the water. A red line displays the current index of the software. The user can click anywhere on the line to move the index to that location. This method enables easy navigation through eight hours of data and allows the user to quickly find the dive occurrences in the data.

The software plots the data in real-time, concurrently across the different visualisations. On the left, the software plots raw sensor data continuously on three graphs. Each graph can plot up to three different sensors at a time. Listed below are the available sensors to visualise;

- Accelerometer (X, Y & Z)
- Hall sensor (Beak activity)
- Pressure sensor
- IR sensor (Measures the underwater speed of the bird)
- Compass (X, Y & Z)
- Temperature sensor
- Lux Sensor (light sensor)
- Pitch & Roll sensor

In addition to the raw data plotting, figure 3 shows a 3D visualisation of the accelerometer data. Again plotted in real-time, this interactive window allows user to rotate the sphere in multiple axes to see the real-time plotting of accelerometer data. New points are plotted in red and fade to yellow over time. This allows the user to see the birds physical activity in three dimensions, adding to the context based plot of the main visualisation.

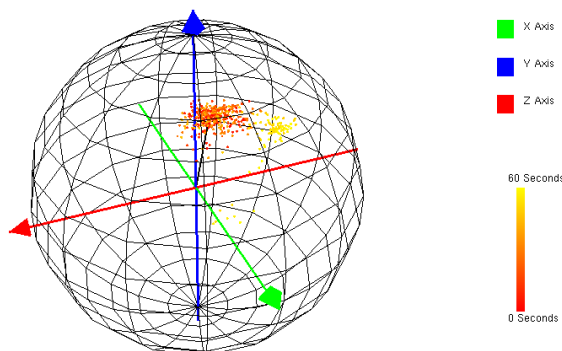


Figure 3: This figure shows the 3D accelerometer visualisation. The user can click and drag the sphere to rotate along the X & X axis and around the Y axis.

The main visualisation will only appear when the current real-time index falls within a dive activity. If the current dive is being visualised, a temperature gauge will appear next to it, plotting the temperature in real time on a sliding colour legend. This main visualisation is accompanied by a calculated average dive. This is as a reference to show how far deviated the current dive is from an average dive.

Additionally, a user option changes the view to a side-by-side comparison of all 78 of the behaviour instances in the data. This overview allows the user to see patterns and trends in the data or pick out outlying behavioural instances (Figure 5).

4. Conclusions and future work

In this paper we present a novel, work-in-progress approach to visualising sensor data. Combining two levels together in a polar coordinates system enables direct comparison of dimensions in the data set. The bird diving activity that was visualised can be directly compared against other dive instances in the same set, or compared against dive sensor data from other species. This cross-species comparison is a currently desired area of research in marine biology.

The rule based system for capturing individual dives provides a fast and accurate means to identify behavioural sequences. By dividing each dive into its various phases and associating each phase with a visual indicator in the overview, it presents an easy to interpret summary of the sequence. We have not previously been able to see all sequences presented together so clearly, so the visualisation provides a novel overview of the behaviour of a particular bird or bird group. We can view as a whole, look for a dive sequence of interest to zoom into, and then compare and relate that sequence against others.

In future work, we hope to add more layers to the visualisation and investigate the challenges that may occur in doing so. We will also refine the current visualisation by extending the user options and upgrading the user interface. Additional features to be implemented are dive sorting in the overview visualisation. The dives can be sorted by each of the data dimensions. In addition to the sorting feature, the dives could be clustered into groups by a selection of their data components. Visualising the dives in the sorted order or cluster could reveal a new layer of information previously unseen.

The development of a more in-depth visualisation could enable a better understanding of the animal behaviour. We could look at the variation in dive depths and compare the values against other dimensions of the data set to find potential correlations. Another aspect of the dive we can perform further analysis on is the speed of dive and ascent; from a look at figure 5 it is difficult to see subtle variations in descent and ascent speed across multiple dives. Here we could create a visualisation tool to display Δ to add a new layer to the information we are creating.

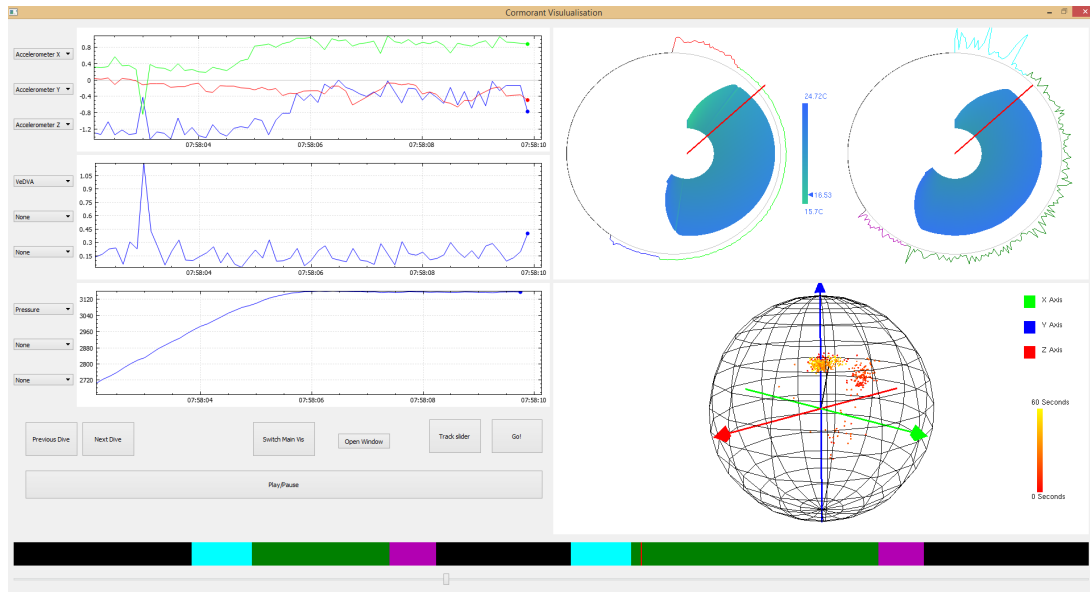


Figure 4: Top Left: Raw amplitude vs time data for five sensor attributes plotting in real time. User selection of sensors to plot from drop down box. Top Right: Main Visualisation (Left: Average Dive, Right: Current dive). Bottom right: 3D Accelerometer Visualisation. Bottom Left: User interface (Play/Pause, Next/Previous dive, slider options and switch view). Bottom Row: Navigation slider.

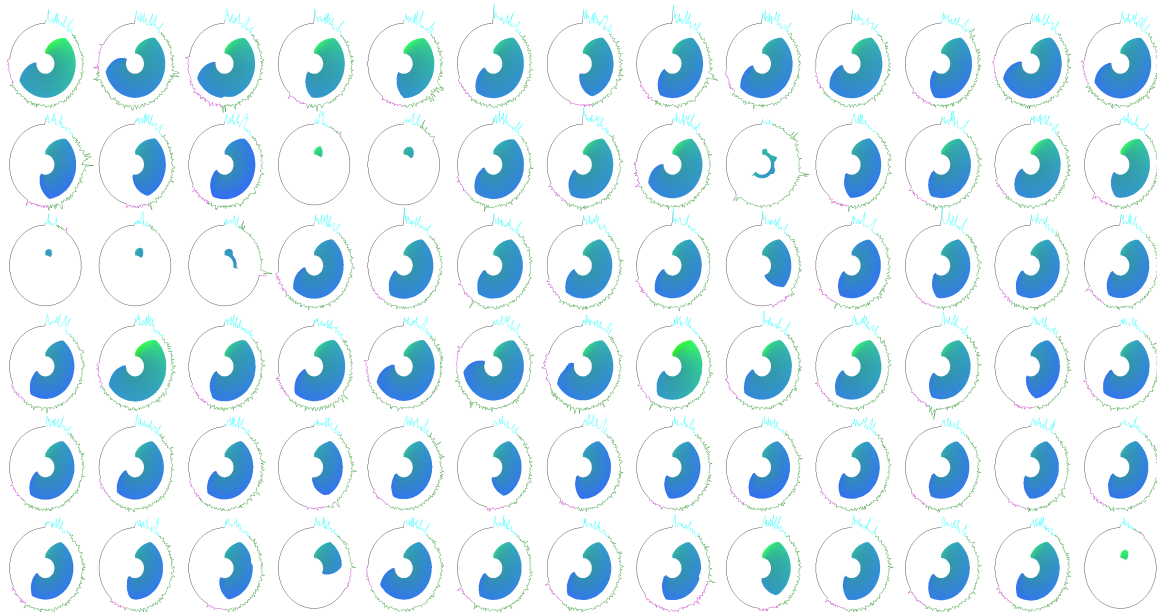


Figure 5: Here is a high resolution image of the complete dive set in chronological order. This is the alternative view of the main visualisation window in the software; pressing 'Switch Main Vis' toggles between the single dive view and complete dive view. Here you can observe the range of dive sizes. There are seven outlying dives in the visualisation where the bird quickly retreated from the dive, indicating that a part of the environment was not optimal for the activity.

References

- [AC91] ALPERN B., CARTER L.: The hyperbox. In *Visualization, 1991. Visualization'91, Proceedings., IEEE Conference on (1991)*, IEEE, pp. 133–139. [2](#)
- [AdOL04] ARTERO A. O., DE OLIVEIRA M. C. F., LEVKOWITZ H.: Uncovering clusters in crowded parallel coordinates visualizations. In *Information Visualization, 2004. INFOVIS 2004. IEEE Symposium on (2004)*, IEEE, pp. 81–88. [2](#)
- [AMST11] AIGNER W., MIKSCH S., SCHUMANN H., TOMINSKI C.: *Visualization of time-oriented data*. Springer Science & Business Media, 2011. [2](#)
- [BBP08] BLAAS J., BOTHA C. P., POST F. H.: Extensions of parallel coordinates for interactive exploration of large multi-timepoint data sets. *Visualization and Computer Graphics, IEEE Transactions on* 14, 6 (2008), 1436–1451. [2](#)
- [CLNL87] CARR D. B., LITTLEFIELD R. J., NICHOLSON W., LITTLEFIELD J.: Scatterplot matrix techniques for large n. *Journal of the American Statistical Association* 82, 398 (1987), 424–436. [2](#)
- [CVW11] CLAESSEN J. H., VAN WIJK J. J.: Flexible linked axes for multivariate data visualization. *Visualization and Computer Graphics, IEEE Transactions on* 17, 12 (2011), 2310–2316. [2](#)
- [DLR09] DRAPER G., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. *Visualization and Computer Graphics, IEEE Transactions on* 15, 5 (2009), 759–776. [1](#), [2](#)
- [GHWF09] GREEN J., HALSEY L. G., WILSON R., FRAPPELL P.: Estimating energy expenditure of animals using the accelerometry technique: activity, inactivity and comparison with the heart-rate technique. *Journal of Experimental Biology* 212, 4 (2009), 471–482. [4](#)
- [GJL*09] GRUNDY E., JONES M. W., LARAMÉE R. S., WILSON R. P., SHEPARD E. L.: Visualisation of sensor data from animal movement. In *Computer Graphics Forum (2009)*, vol. 28, Wiley Online Library, pp. 815–822. [2](#)
- [GPL*11] GENG Z., PENG Z., LARAMÉE R. S., ROBERTS J. C., WALKER R.: Angular histograms: Frequency-based visualizations for large, high dimensional data. *Visualization and Computer Graphics, IEEE Transactions on* 17, 12 (2011), 2572–2580. [2](#)
- [GTC01] GRINSTEIN G., TRUTSCHL M., CVEK U.: High-dimensional visualizations. In *Proceedings of Workshop on Visual Data Mining, ACM Conference on Knowledge Discovery and Data Mining (2001)*, pp. 1–14. [1](#)
- [GWS11] GLEISS A. C., WILSON R. P., SHEPARD E. L.: Making overall dynamic body acceleration work: on the theory of acceleration as a proxy for energy expenditure. *Methods in Ecology and Evolution* 2, 1 (2011), 23–33. [3](#), [4](#)
- [TS08] TOMINSKI C., SCHUMANN H.: Enhanced interactive spiral display. In *Proceedings of the annual SIGRAD conference, special theme: interactivity (2008)*, pp. 53–56. [2](#)
- [VMCJ10] VIAU C., MCGUFFIN M. J., CHIRICOTA Y., JURISICA I.: The flowvizmenu and parallel scatterplot matrix: Hybrid multidimensional visualizations for network exploration. *Visualization and Computer Graphics, IEEE Transactions on* 16, 6 (2010), 1100–1108. [2](#)
- [vWvL93] VAN WIJK J. J., VAN LIERE R.: Hyperslice: visualization of scalar functions of many variables. In *Proceedings of the 4th conference on Visualization'93 (1993)*, IEEE Computer Society, pp. 119–125. [2](#)
- [VWVS99] VAN WIJK J. J., VAN SELOW E. R.: Cluster and calendar based visualization of time series data. In *Information Visualization, 1999.(Info Vis' 99) Proceedings. 1999 IEEE Symposium on (1999)*, IEEE, pp. 4–9. [2](#)
- [WAM01] WEBER M., ALEXA M., MÜLLER W.: Visualizing time-series on spirals. In *Information Visualization, IEEE Symposium on (2001)*, IEEE Computer Society, pp. 7–7. [2](#)
- [WC00] WHITFIELD P. H., CANNON A. J.: Polar plotting of seasonal hydrologic and climatic data. [2](#)
- [Weg90] WEGMAN E. J.: Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association* 85, 411 (1990), 664–675. [2](#)
- [WGJL] WALKER J., GENG Z., JONES M., LARAMÉE R. S.: Visualization of large, time-dependent, abstract data with integrated spherical and parallel coordinates. [2](#)
- [WGK10] WARD M. O., GRINSTEIN G., KEIM D.: *Interactive data visualization: foundations, techniques, and applications*. CRC Press, 2010. [2](#)
- [ZCB11] ZHAO J., CHEVALIER F., BALAKRISHNAN R.: Kronominer: using multi-foci navigation for the visual exploration of time-series data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011)*, ACM, pp. 1737–1746. [2](#)