Visualizing Confusion Matrices
for Multidimensional Signal Detection Correlational Methods

Yue Zhou[a], Thomas Wischgoll*[a], Leslie M. Blaha[b], Ross Smith[c], Rhonda J. Vickery[c]

[a]Wright State University, 3640 Colonel Glenn Hwy Dayton, OH, USA 45435; [b]711th Human Performance Wing, Human Effectiveness Directorate, Air Force Research Laboratory, WPAFB, OH, USA 45433; [c]Dynamics Research Corporation, 2900 Presidential Dr, Fairborn, OH 45324, *Contact information: email: thomas.wischgoll@wright.edu, phone: 937-775-5057

ABSTRACT

Advances in modeling and simulation for General Recognition Theory have produced more data than can be easily visualized using traditional techniques. In this area of psychological modeling, domain experts are struggling to find effective ways to compare large-scale simulation results. This paper describes methods that adapt the web-based D3 visualization framework combined with pre-processing tools to enable domain specialists to more easily interpret their data. The D3 framework utilizes Javascript and scalable vector graphics (SVG) to generate visualizations that can run readily within the web browser for domain specialists. Parallel coordinate plots and heat maps were developed for identification-confusion matrix data, and the results were shown to a GRT expert for an informal evaluation of their utility. There is a clear benefit to model interpretation from these visualizations when researchers need to interpret larger amounts of simulated data.

Keywords: Information Visualization, Analytics, Infovis, Web-browser-based Visualization, D3, Javascipt, Scalable Vector Graphics (SVG), General Recognition Theory

1. INTRODUCTION

General Recognition Theory (GRT) is a multidimensional signal detection theory framework for capturing sources of perceptual and decisional dependence in human perceptual decision making tasks[1]. The primary type of data for GRT models is an identification-confusion matrix (IDCM), typically derived from a complete factorial identification task. In a complete factorial identification task, observers are given all combinations of a finite set of features (often limited to two features at two different values each for a total of four stimuli, for experimental tractability) and asked to provide the unique identification assigned to each stimulus. In an experiment with four stimuli, there are four unique responses. For example, Figure 1(a) depicts IDCM data from face stimuli where the distance between the eyes and the width of the mouth were manipulated, each at two levels, and the factorial combinations produced four unique face stimuli. A single IDCM tabulates the number of responses given to each stimulus over the course of an experiment, with one IDCM created for each individual observer in the experiment. If the observer is able to perfectly recognize all the stimuli, then the IDCM would have non-zero values on the diagonal, reflecting correct responses, and zero values everywhere else. Non-zero off-diagonal entries in the IDCM, on the other hand, indicate identification mistakes, and GRT researchers use the patterns of confusion in these mistakes to draw inferences about the perceptual and cognitive mechanisms supporting the task under investigation.

GRT attempts to infer three types of independence from IDCM data: perceptual independence, perceptual separability, and decisional separability[1]. While we leave the detailed technical definitions to the GRT literature[1,2,3] to focus here on the analysis and visualizations processes, we note that GRT modelers seek to understand, for a given task, whether or not each of the independence constructs exhibited a violation. Lack of violations means that independence holds on all constructs, while violations of any type of independence result in response patterns that reflect some form of correlation in the GRT space, which we review below. The goal of the analysis and visualization, then, is to identify the violations present in any dataset and then to understand how the patterns of confusion and (in)dependence in the data reflect the underlying psychological representation of the task under investigation.

One theoretical challenge for GRT researchers is model mimicry, where the patterns of confusion in a given IDCM can be explained by multiple combinations of violations in the GRT model[2,4]. The basic tools of GRT analysis[1,3,5] may not be enough to tease apart mimicking models because of limited theory-to-data implications that are necessary but not sufficient. This has prompted researchers to explore additional statistical tools that may help provide converging evidence to distinguish candidate models in mimicking situations; together with advances in computational power, this work has begun to entail large-scale simulations to test the new and existing measures, and to explore nuanced GRT model situations for their potential to result in model mimicry GRT predictions[2,6,7]. In Meneer and colleagues[6,7], for example, through simulations of 0-3 model violations, each violation at 2-4 parameter levels, with 1000 IDCMs simulated per model, researchers generated over 2 million datasets. As explained in more detail below, the current analysis and visualization techniques that focus on the modeling of a single IDCM do not effectively scale up to the volume of data in these simulation studies. Thus, new visualization tools are needed to support GRT research with larger volumes of data. Additionally, as one of the goals of the simulation studies is to compare the model spaces for different combinations of independence and violations of independence together with comparing the inferences made with the various statistical analyses, visualization tools are needed that encode as much detail as possible within the visualization.

This paper describes the application of different information visualization techniques to a large number of IDCMs to help domain specialists more effectively analyze their data. These techniques were specifically adapted and expanded to improve the visualization of IDCM data[2,6,7]. The paper is structured as follows. Section 2 discusses standard techniques for visualizing GRT models, as well as work related to the visualization approaches presented in this paper. Section 3 describes the methodology developed for visualizing the IDCM data, to augment traditional GRT analyses and to visualize large numbers of simulated datasets. Section 4 presents the results together with discussion and commentary by a GRT domain expert. Finally conclusions and recommendations are presented.
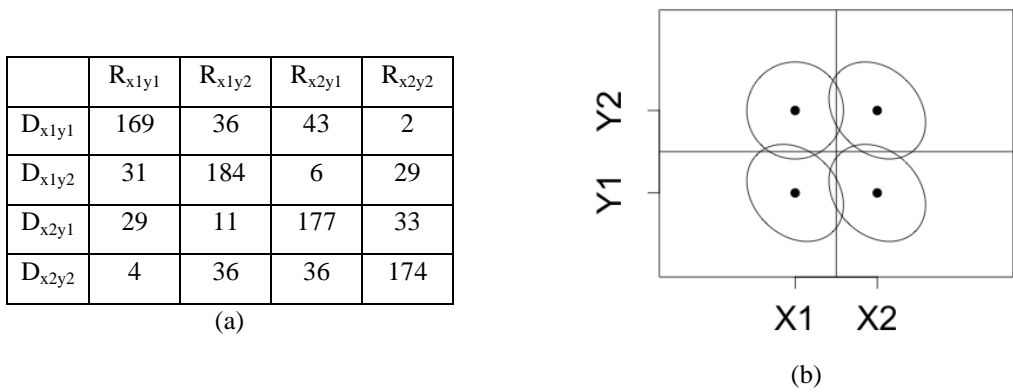
	R_{x1y1}	R_{x1y2}	R_{x2y1}	R_{x2y2}
D_{x1y1}	169	36	43	2
D_{x1y2}	31	184	6	29
D_{x2y1}	29	11	177	33
D_{x2y2}	4	36	36	174

(a)

(b)

Figure 1.   (a) Sample identification-confusion matrix (IDCM) from one observer in a GRT study. The model stimuli are denoted D_{xiyj} for the four possible combinations, and the responses are denoted R_{xiyj}. Thus, the diagonal tabulates the correct responses, and the off-diagonal entries show the pattern of mistakes, or confusions, made by the observer. (b) Typical visualization of the GRT model resulting from the analysis of the IDCM in (a). Drawn are the equal probability contours from the four bivariate Gaussian distributions comprising a GRT model space. The raw IDCM data is not visualized. Data taken from the simulation study by Menneer and colleagues[6,7].

2. RELATED WORK

This section discusses previous work that is related to the methods used in this paper. It first provides an overview of current techniques used for visualizing GRT data and subsequently presents visualization methodologies that could be used as an alternative with some modifications as outlined in the methodology section.

2.1 GRT Visualization

The typical analysis pipeline for a GRT study is to (1) collect the IDCM data, (2) perform non-parametric tests for independence and separability and/or parametric multidimensional signal detection theory analyses, and (3) develop a GRT model space visualization. The goal of the visualization is to capture the psychological representation model that reflects the independence or violations of independence found in step 2. An example of a typical GRT model space is shown in Figure 1(b), derived from the data in Figure 1(a). The two dimensions in the space (X, Y) represent two

stimulus features from the experiment; each of these feature dimensions has two values (e.g. X1, X2), which reflect a manipulation of two stimulus levels used in the human observer experiment. This creates a space with four unique quadrants defined by the factorial combination of two features each having two levels. The boundaries between the quadrants represent the observer's decision bounds that divide the space into four mutually exclusive response regions. The ellipses in the model space represent the equal probability contours of four bivariate Gaussian distributions, with each distribution reflecting the probabilistic psychological response to one of the four stimuli. The shape of these contours, the amount of overlap of contours, and the relationships of the contours to the decision bounds reflect the degree of violation of perceptual independence, perceptual separability, and decisional separability, respectively. Thus, inferences about the GRT model for a single observer can be made from a single GRT model space visualization. A typical GRT domain expert would look at Figure 1(b) to draw inferences about the psychological relationships between the dimensions of the stimulus and the task. For example, the contours shown here exhibit correlations between X and Y dimensions for most of the stimuli, shown by the non-circular tilted ellipses. The centers of the contours form a rectangle, which is interpreted as separable dimensions.

More recent work by Silbert and colleagues in visualizing GRT modeling has involved efforts to incorporate modeling results and parameters into the visualizations of the model spaces[8,9]. The goal of these visualizations is to convey in a concise manner how robust the model fits are to the data, as well as the variability across the observers in the parameter estimates, which serves as a proxy for the variability of the underlying psychological representations across observers (i.e. this attempts to tell researchers how similar, or not, everyone's internal psychological models are). While these are helping add information about fitted model results, they do not yet incorporate any information about the traditional non-parametric GRT analyses[1] or signal detection analyses[5].

Note that the GRT analysis pipeline process is completed for each individual observer, which results in *N* model spaces for *N* observers. Comparisons across those observers are usually done by comparing the images from step 3, with *N* spaces presented as small multiples, or by tabulating the statistical test results and parameter estimates of step 2. When a study has a relatively small number of observers, say 5-10, small multiples presentation of GRT spaces like Figure 1(b) can be completed in a relatively easy manner. However, as the number of plots increases, comparisons of all GRT space images, especially if all plots must be presented like small multiples, is no longer an efficient undertaking for the experimenter, and communication of the results becomes an unwieldy task from the visualization standpoint. In a study like Menneer, Wenger and Blaha[2], where over 2 million IDCMs were simulated, small multiples comparisons is not an option to reasonably consider. Thus, in order to effectively capture patterns in large-scale simulations or experiments, new visualization techniques are needed for the GRT framework.

Additionally, the reader may observe that while the GRT inferences are all based on the IDCM data, this "raw" data is not visualized in any way for direct comparisons; currently, only the results of GRT analyses or model fitting of the IDCM data are visualized. However, given that violations of the three independence constructs of GRT produce correlations in the underlying model, they may also produce distinct patterns in the IDCMs that visualization could illuminate for the researchers. Visualization of the IDCMs may also lend itself to predicting patterns in the GRT model space visualizations that could further aid GRT experts in predicting if they have data that will be vulnerable to model mimicry or is otherwise problematic for GRT analyses or model fitting.

2.2 Multidimensional Data Visualization

The fact that IDCMs are multidimensional in nature requires a visualization technique that is capable of conveying higher-dimensional data. Parallel coordinates are one of the first two-dimensional visualization techniques used for high-dimensional data[10]. The basic idea is to present *N* dimensional data of the attribute space by mapping *N* equidistant parallel axes to the two-dimensional space. The axis of each dimension represents a property. The axis of the corresponding attribute values range from minimum to maximum uniform distribution. Thus, each data item can be used in accordance with its property values in a line segment on *N* parallel axes. The parallel coordinates technique has been used as a powerful visualization tool. However, when large datasets are visualized using parallel coordinates, it will cause confusion due to large numbers of lines overlapping. Many improvements have been suggested in the literature. For example, Peng defined visual clutter as any aspect of the visualization that interferes with the viewer's understanding of the data, and presents the concept of clutter-based dimension reordering[11]. Dimension order is a technique that can significantly affect the visualization's expressiveness. By varying the dimension order in a display, it is possible to reduce clutter without reducing information content or modifying the data in any way. Hauser et al. presented angular brushing for parallel coordinates as a new approach to highlighting rational data-properties, i.e., features which - in a non-separable way - depend on two data dimensions[12]. They also demonstrated smooth brushing as an intuitive tool for

specifying non-binary degree-of-interest functions (for *focus+context* visualization). Siirtola introduced two web browser based novel techniques to manipulate parallel coordinates[13]. The first technique provides the capability of dynamically summarizing a set of polylines with polyline averaging, which can replace computationally much more demanding methods. The second technique interactively visualizes correlation coefficients between polyline subsets, helping the user to discover new information in the dataset. Zhao et al. proposed a technique based on parallel coordinates that utilizes an edit-distance-based technique to rearrange variables so that patterns of interest can be easily detected[14]. This system, V-Miner, includes both automated methods for visualizing common patterns and a query tool that enables the user to describe specific target patterns to be mined/displayed by the system. Johansson et al. introduced a method to allow the user to simultaneously examine the relationships of a single dimension with many others in the data[15]. The single dimension can then be interactively changed to allow the user to quickly examine all possible combinations. This is accomplished by extending the standard parallel coordinate approach to a 3D clustered multi-relational parallel coordinate representation.

Another simple, but very effective information visualization tool is a heat map[16]. The basic idea of a heat map is essentially a two-dimensional visualization that utilizes a predefined gradient color palette to encode the individual data values by color mapping. For example, the closer the data value gets to a range associated with "hot", the brighter the color used to represent that data value. A single data point is rendered with a single color value from the color palette. Numerous variations of the basic heat map[17,18,19] have evolved for use with large multidimensional datasets. We describe our variation of this technique for use with IDCM data in Section 4.

Data visualization is a means of conveying information but can also add art and design through visual elements to more clearly communicate information to the viewer. Using web-based techniques can further help reach a broader audience. For this, Bostock et al. proposed a web-based JavaScript library called D3 (Data-Driven Documents)[20]. When people use this kind of web-based visualization as a means of communication, it not only changes the design semantics but also provides numerous options for 2D display of multivariate data. For these reasons, we used the D3 library as a starting point for our visualizations.

3. MOTIVATION

The goal of the presented visualization process is for relevant GRT data to be presented in a form (e.g. an adjustable map) that is optimized for the human visual perception system and provides insight into the particular issues being evaluated. This process resembles a standard information visualization pipeline that describes the transition from the original data to the final representation. In a standard information visualization pipeline, the data may need to undergo a series of data conversions first. The process of converting the original data to the final visualization is likely to be expressed in a series of transformations that may include user operations to fine-tune the results. For example, the displayed data is a conversion of the original data to a table with correlation descriptions. Visual mapping of that data table into a visual structure is a combination of space-based marking and graphical attributes. View transformations defining the position, scale, cropping, and other graphical parameters create a visual view of the data representation. The user can interactively control the operational parameters of the transformations within the visualization pipeline, thereby binding to a specific view of the range of data or changing the attributes of the transformations.

The information in the data table to be communicated is based on mathematical relationships, whereas the visualization structure is based on being able to deal effectively with the visual attributes. Scientific visualization often focuses on location-based datasets that are typically the results of scientific experiments or simulations, whereas information visualization research generally includes more abstract information. In many cases, the information itself cannot automatically be mapped to geometry or physical space. This means that different types of information with no natural or obvious physical representation are common. Therefore, a key objective in information visualization is to discover appropriate visual metaphors and structure to represent the pertinent information, combined with an understanding of the information supported by a related analysis task.

In the visualization process, the data table is mapped to a visual representation. The visual representation is encoded by graphical properties and annotations. In order to get a good visual representation, the mapping of the data onto visual elements has to preserve the original information and only the data in the table is displayed within the visual representation. To find a good mapping is not an easy task. The visual structure is prone to contain less important data, and the data encoded in the visual representation needs to be easily comprehensible by the user. Using a mapping that cannot be easily understood can cause misinterpretation of the results.

View transformations interactively change the view onto the visual representation. One method used to navigate a larger visual representation is to create different viewpoints to investigate specific details encoded within the visualization of the data. Another viewpoint control technique commonly used encodes both an overview and a detailed representation. The overview provides a view of the context, and acts as a control unit to change the detailed view. The detailed view is used to magnify the selected area or focus. Such a *focus+context* visualization can use separate areas to depict the overview and focused representation, or both can be combined within a single view[21]. Typically, the screen real-estate that can be used for the visualization is very limited. To compensate for the imbalance between screen real-estate and space required by the visualization, specifically designed deformations, such as hyperbolic projections, are used. If such a deformation can allow users to perceive a larger visual representation, then such modifications are effective.

Interaction and control of information visualization are dependent on the form of encoding of the data within the visualization. Based on the previous information visualization reference model, the user needs to be able to control the visual mapping by tweaking specific parameters. These controls can be in the form of simple user interface elements, such as various control buttons, scroll bars, etc. Many interactive information visualization techniques, in essence, are a form of selection method by selecting subsets of the data table objects to retrieve the desired visual representation. These interactive techniques can be used to position data elements, reveal patterns within the data, and select the transform parameters, as well as techniques useful for multivariate data, such as *Details-on-demand*[22] and *Brushing*[23]. These interactive techniques not only improve the speed at which the user interacts with the information, but can also help the user avoid taking a wrong path or getting lost in solving a specific problem. Furnas introduced the concept of a fish-eye view, implemented as a zoom in of a small local area combined with a fish-eye lens-type effect[24]. The surrounding information around the enlarged area is pushed toward the outer perimeter, but is still visible. This allows the user to observe a small central focus area while maintaining a greater visibility of the surrounding area, which is the meaning of *focus+context* view. The fish-eye view can be considered as a *collection of information* detail view of a specific portion (focus), where the overall structural view of a set of information (context) is mixed together. The layout can also be considered as many large information spaces at the same time (context), one of which is viewed with more details (focus).

4. METHODOLOGY

4.1 Data structures

The data to be visualized in this paper are the simulated IDCMs from the work of Menneer and colleagues who generated over 2 million individual IDCMs with large-scale GRT simulations in order to have a standard set of data for testing novel statistical tools for GRT analysis against the performance of traditional analysis tools[2,6,7]. Again, an IDCM summarizes the frequency with which the observer (here, a model) gives each of the R responses to each of the D stimulus options. Correct responses are tabulated along the diagonal cells of the IDCM, while confusions (incorrect responses) are tracked in the off-diagonal cells. This simulation of such a large number of IDCMs makes data and analyses comparison a hard problem and renders the traditional individual IDCM analyses unwieldy for the experimenters. Appropriate visualization tools are needed to examine the data in two ways. First, 1000 IDCMs were simulated for each GRT model under study; that is 1000 replications of the same model were run. Visualizations are needed to capture patterns *within* each model's set of data. Second, visualizations are needed to compare *across* the models to allow the changes of patterns to be identified.

To develop the techniques described in this paper, we had to interface with the set of 1000 data files for each model. Each simulated IDCM is contained in its own text file, for over 2 million text files to analyze. Figure 2 shows sample data files from two models with different types of simulated GRT theoretical constructs. A single file contains header information about the model parameters, and then a single labeled IDCM, where D_{xiyj} labels the simulated stimulus and R_{xiyj} labels the model responses. As they illustrate, a typical IDCM exhibits higher values along the main diagonal that indicate predominantly correct responses. Values off the main diagonal indicate confusion between the actual stimulus shown and the response made. It is hard to see in numerical format, but these two IDCMs show different patterns of confusion in the off-diagonal cells. Figure 2(a) has fairly evenly spread confusion, while Figure 2(b) shows most of the confusion responses in the middle two columns and few confusion responses in columns 1 and 4.

In the present paper, the goal was to preprocess the data and then provide the application domain specialists with a web-based user interface and visualization of the data. Current options for web-based visualizations include WebGL- and SVG-based methods. In order to make it easier for the domain specialists to analyze the data using our proposed information visualization techniques, a 2D-based visualization was deemed more appropriate. As a result, we avoided the

extra overhead of integrating WebGL with various web browsers, and we chose an SVG-based approach using the D3 JavaScript library.

Mean(dPrime): 1   AnyZeros: 0   vPI_Rho: 0.00
vPS_Shift: 0.00
vDSc_piecewiseShift_or_continousSlope: 0.00
SampleSize: 100

	R_{x1y1}	R_{x1y2}	R_{x2y1}	R_{x2y2}
D_{x1y1}	48	21	28	3
D_{x1y2}	20	50	12	18
D_{x2y1}	23	16	42	19
D_{x2y2}	11	20	24	45

(a)

Mean(dPrime): 1   AnyZeros: 0   vPI_Rho: -0.25
vPS_Shift: 0.25
vDSX_piecewiseShift_or_continousSlope: 0.25
SampleSize: 100

	R_{x1y1}	R_{x1y2}	R_{x2y1}	R_{x2y2}
D_{x1y1}	29	32	34	5
D_{x1y2}	17	56	21	6
D_{x2y1}	16	10	62	12
D_{x2y2}	2	27	35	36

(b)

Figure 2.   Sample data files (formatted for better readability) for model with violations on all sources of independence (a), indicated by the parameter values equal to zero in the header, and (b), indicated by the non-zero parameter values in the header and no violations.

There is support for numerous data formats available in the D3 library. Due to the size of the overall dataset it was impractical to convert the data by hand. Instead, Visual Basic for Applications (VBA) was used to convert the raw text-based data into an appropriate format for D3, specifically the comma separated variable (CSV) format. The CSV format was chosen as it is readily human readable and can be imported into common spreadsheet applications, such as Micosoft Excel. Once the data was imported, multiple approaches were tested to find appropriate ways to visually analyze the data, namely parallel coordinates and heat maps following various different layout paradigms.

In some of the visualizations, numerous duplicated data values caused a significant loss in rendering speed, particularly when the user employed interactive brushing and other selection techniques in the parallel coordinates layout, even though there was no perceived visual difference whether the duplicated values were included or not. This performance issue was resolved in the preprocessing step by only including unique rows in the CSV file to be imported into D3. However, this data culling was not performed during heat map rendering since a representation for each value is displayed.

4.2 Data visualizations

In order to develop the techniques describing different elements in this paper, several heat maps and parallel coordinate plots were created using D3. In the parallel coordinates layout, the series of lines represent the four individual sets of data and the average data across IDCMs, and styles were set for the browser web page using Cascading Style Sheets (CSS). The values corresponding to D_{x1y1} in the confusion matrices datasets are colored red. Values for D_{x1y2} are green, while D_{x2y1} and D_{x2y2} are colored blue and orange respectively, as shown in Figure 3. A line representing the average value for the IDCM dataset is over-drawn in black. Because of the large number of lines overlapping each other, the opacity has been set to 50% to improve readability. The data is read using the D3 library and stored such that it is accessible via a *layout* function, and can be used anywhere by a simple function call. D3 reads the CSV data and preprocesses them as different *traits arrays*. In this confusion matrices dataset, D_{x1y1}, D_{x1y2}, D_{x2y1}, and D_{x2y2} are different *traits arrays*.

The values of those arrays need to be transformed into different lines for the parallel coordinates plot and plotted at the correct height positions. In order to show different values in the arrays, a larger value is set to a higher position on the *y axis* in this layout. The height of each position shows the difference of each value. Now it is simple to show only this dataset because the minimum is 0 and the maximum is 100. If the maximum height in this layout is 50 pixels, a value of 20 can be set to 20% of 50 pixels and a value of 80 can be set to 80% of 50 pixels. However, some datasets have a maximal value of up to 1000. Hence, a scale function within D3 is deployed to rescale the values appropriately. In order to use this function, *domain* and *range* need to be provided as input parameters. The *domain* describes the range of input values whereas the *range* is the range of output values. In this case it should output a pixel unit to set the position. For example, if the minimum value is 0 then the position is 50 pixels; if the maximum value is 100, then the position is 200

pixels, and the range of the output is 50 to 200. In this case, the center of the *range* is 125, so if the input value is 50, then the output will be 125 pixels. D3 provides the function *y[d] = d3.scale.linear()* to apply the linear scaling. To set the *domain* input, *.domain(d3.extent(layout, function(p) { return +p[d]; }))* determines the range of input values in the array. The function called *layout* has been generated before, so it provides access to all the values in this dataset. The *d3.extent* function is used to determine the minimum and maximum value in that array. To set the range of output values, *.range([h, 0])* sets the range of output in pixels. At this point, the position of each value has been set. Finally, the line for each path is generated via an SVG object. An SVG function in D3 is used to create a new path, the function call *svg.append("svg:path")* is used to create a new HTML element, and *.attr("d", path) .attr("class", function(d) { return d.rows; })* is called to set the attributes. Because the style of the CSS for the path is already set, the lines generated as new HTML elements use different colors for each array.

5. RESULTS AND DISCUSSION
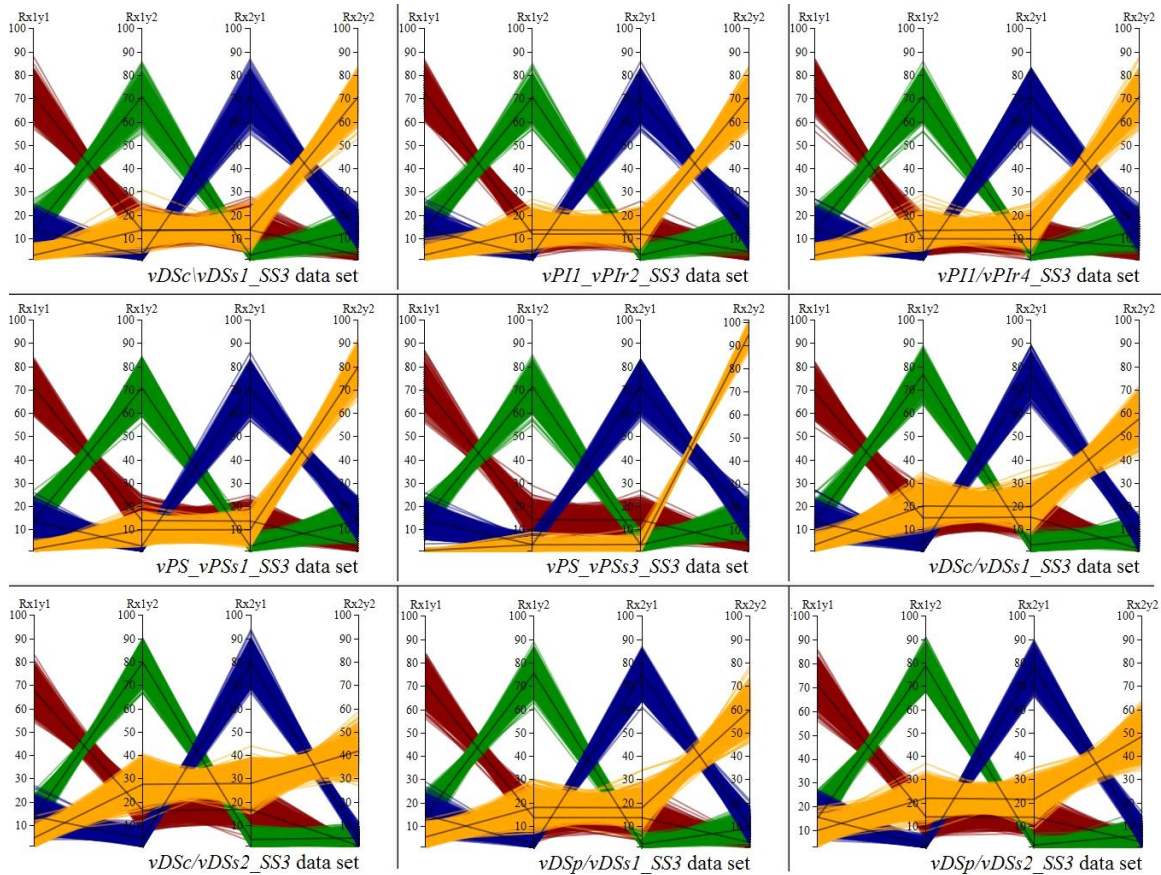
5.1 Parallel coordinates

Figure 3.　　　Parallel coordinate plot with 9 different GRT model simulation datasets. The colors encode different stimulus conditions, with red for D_{x1y1}, green for D_{x1y2}, blue for D_{x2y1}, and yellow for D_{x2y2}. The black line within each set of colored lines represents the mean response rates for that stimulus.

Parallel coordinates are a very useful layout to compare each row and each column in every matrix. When properly preprocessed, parallel coordinates can easily handle the 1000 IDCMs for a single simulated model in one visualization. In the IDCM datasets, the rows D_{x1y1}, D_{x1y2}, D_{x2y1}, and D_{x2y2} are drawn with different colors in order to make it much easier to compare (red, green, blue, and yellow, respectively). The columns, R_{x1y1}, R_{x1y2}, R_{x2y1}, and R_{x2y2}, are then plotted as the axes of the parallel coordinate system to show the values. This allows the user to compare a value with other rows using a different color, or compare with different columns using the same color. If the spread of a group of same colored lines becomes wider, that means the overall values represented from the matrix tend to be noisier; and if the spread of a group of same colored lines becomes narrower, that means the represented values are much closer to each other and

therefore less noisy. The black lines in the centers of each different colored group show the average response rate values, which enables the user to more easily understand the difference of the spread of values encoded within the matrix. Also a small multiples layout is capable of showing several different datasets together within a single visualization, so comparing different datasets can be done easily (see Figure 3).

The nine parallel coordinate plots in Figure 3 illustrate nine different GRT theoretical model situations with specific conditions annotated at the bottom of each plot. Researchers are trying to understand how different combinations of GRT model construct violations influence the pattern of responses within the IDCMs. Figure 3 shows a possible way to quickly compare the rates of responding to the different stimulus conditions (as shown by the different colors) across multiple models. The parallel coordinate plots in Figure 3 offer GRT experts a means of examining trends in the IDCMs both within an across model simulations, which was one of the main goals for our IDCM visualizations.

When presented to two GRT experts, they quickly recognized that the yellow lines, representing the D_{x2y2} stimulus, show a high degree of variability across the plots; this was the stimulus modified by changes in the simulation parameters, confirming that the simulations were behaving as expected. The center plot shows very little confusion for the D_{x2y2} stimulus, because the response rates are low for R_{x1y1}, R_{x1y2}, and R_{x2y1}, and high R_{x2y2}; the narrow spread of the yellow lines in this plot shows very consistent responses across the 1000 IDCMs for this model. This is very different from the models in the middle right and bottom three plots, which show a large amount of confusion of D_{x2y2} with the other stimuli (wider spread of lines, lower total height on R_{x2y2} and higher rates of R_{x1y2} and R_{x2y1}). Thus, those three models produced more confusion and might be harder to analyze in a human perception experiment. The models showing larger variation for the spread of yellow lines all exhibited a violation of decisional separability, where the middle model did not; rather it had violations of perceptual separability. So visually experts can quickly see gross differences in the types of violations exhibited based on the shapes in the parallel coordinate plots. Domain experts see much potential for using these parallel coordinate plots for model comparison based on the IDCM patterns alone (without additional statistical analyses), especially if the parallel coordinate plots can quickly group datasets based on the types of violations exhibited by the data, which can then be quantified with the statistical analyses.

5.2 Heat map visualization

A heat map is another way of visualizing the matrix data, and has the potential for preserving the matrix-like form of the IDCM data. Generally, heat maps, while not depicting the exact absolute value encoded in the confusion matrix, are useful for providing an overview of the values of the confusion matrices relative to each other. The goal here was to utilize the heat maps to visualize the data from multiple individual IDCM datasets, including layouts that might enable the visualization of all 1000 individual IDCMs for each simulated model. Different layouts for heat maps were tested to identify the more suitable ones for looking within and across datasets.

Figure 4.　　　Layout example of heat map - a combination of each matrix in the dataset. The image on the right illustrates the details on demand offered by a mouse hovering over a single IDCM.

Figure 4 shows an example of a heat map layout for all multiple IDCMs from the same model; it is a combination of each matrix in the dataset. It highlights each individual value within the entire dataset using color coding ranging from white (low value) to red (large value). Also for each exact value, the user can hover the mouse over the heat map, which will automatically show the partitioning of the values in that confusion matrix to which that cell belongs (Figure 4, right hand side). This basic heat map allows for several IDCMs to be displayed together, which can help identify some

patterns of confusion. However, with 16 cells for each IDCM, this particular layout can get unwieldy rather quickly. Figure 4 only captures 30 IDCMs from a model, so it will take a large amount of real estate to capture all 1000, let alone multiple model datasets. GRT experts quickly recognize that this layout will not scale well to larger numbers of IDCMs, and also reported difficulty segmenting the individual IDCMs from the larger set without the partitioning in the right-hand side of Figure 4. They were not too interested in using this for large-scale pattern identification.

An alternative heat map layout that better assembles all values in the dataset to one matrix layout, to better combine the large number of data files, is shown in Figure 5. In this layout, the specific cells within the IDCMs are lumped together in the same heat map block. For example, the values of the upper left cell (row 1, column 1) of each IDCM gets plotted in a single block where the value of the first confusion matrix is used for the first cell within the block, the second value is used for the second cell, and so on. This better preserves the overall layout of the confusion matrix (4-by-4 grid), which makes it appear more familiar with the domain specialists. It also tends to better show the overall confusion for the entire study, because all 1000 IDCMs are included.

Comparing the two heat maps in the top row of Figure 5, domain experts quickly identified that both heat maps exhibit the darkest blocks along the main diagonal, which means the responses overall were correct (this is important for ensuring that the model or human participant are doing the task as expected). However, the heat map in the upper left has darker blocks surrounding the main diagonal compared to the upper right, which indicates greater confusion in response to the stimulus for the upper left model. The bottom left heat map shows an example with even less confusion during the study, as many of the off-diagonal blocks are close to white in color (very low numbers of responses). Similar to the previous layout from Figure 4, in the lower right heat map in Figure 5, the user can hover the mouse over the heat map to automatically select each value within a single block, thereby showing the exact values of that block. Similar to the parallel coordinate plots in Figure 3, the domain experts will be able to use these blocked heat maps to quickly compare multiple simulated model conditions for their influence on the overall pattern of confusions.
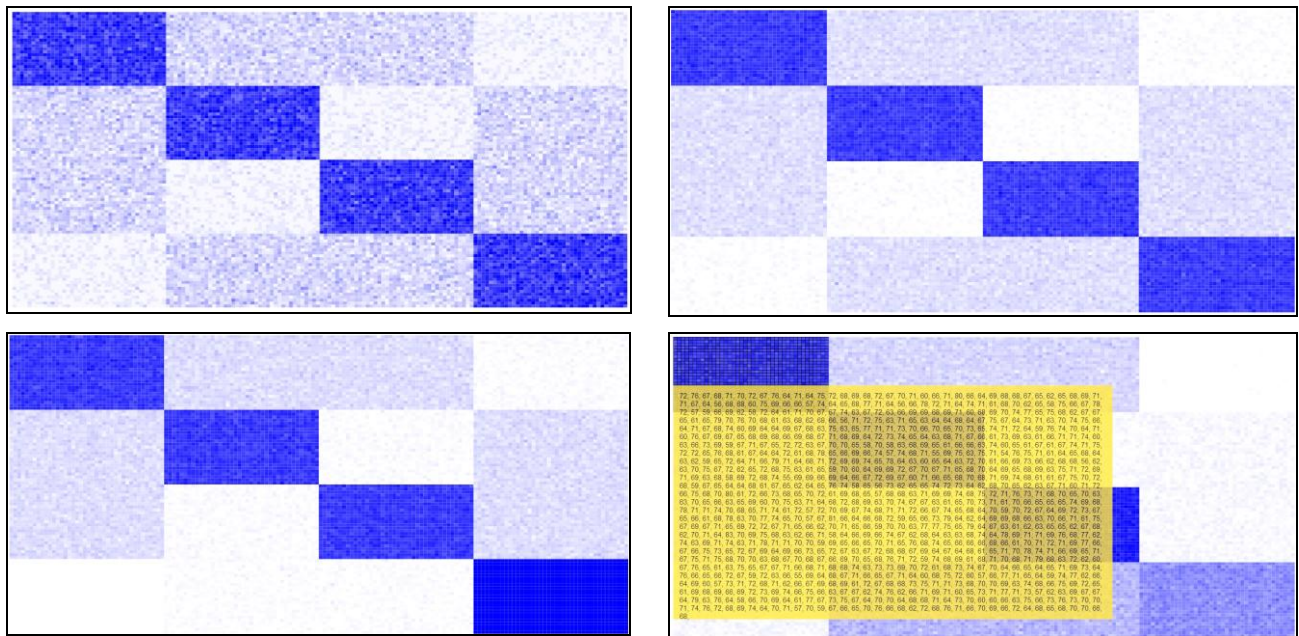

Figure 5.   Heat map layout in which values are lumped by their position within the confusion matrix. Each of the four heat maps illustrates a different simulated set of data (i.e. different underlying model) but all capture the 1000 IDCMs simulated per model. This is to show how this layout also highlights differences in the overall confusion patterns (color changes in the off-diagonal blocks). The mouse over action illustrated in the bottom right corner now shows all the values from the 1000 IDCMS for the selected cell (here, the row 1, column 1 cells of all 1000 IDCMs).

When viewing these enhanced visualizations, the two GRT experts liked the ability to quickly relate the heat map layout to the traditional numerical IDCMs with which they were familiar. They additionally liked seeing how the patterns of confusion shifted the coloring of the off-diagonal blocks as the types of model violations were changed. It was obvious to them that the pattern shown in the lower left of Figure 5 is the same as the middle plot of Figure 3. However, they

think that the parallel coordinate plots of Figure 3 may provide a more compact, though less familiar, means of comparing a large number of models or conditions. The layout of Figure 5 may provide a unique advantage for identifying sets of data that will not be modeled well, due to missing data or not enough confusion for the GRT framework (some confusion is needed for the models to work at all). That is, if a heat map reveals only responses on the diagonals and none in the off-diagonals, then the analysis will not succeed. The quick visual comparisons offered by the heat map in Figure 5 will enable researchers to avoid wasting time on larger numbers of data files that cannot properly be analyzed.
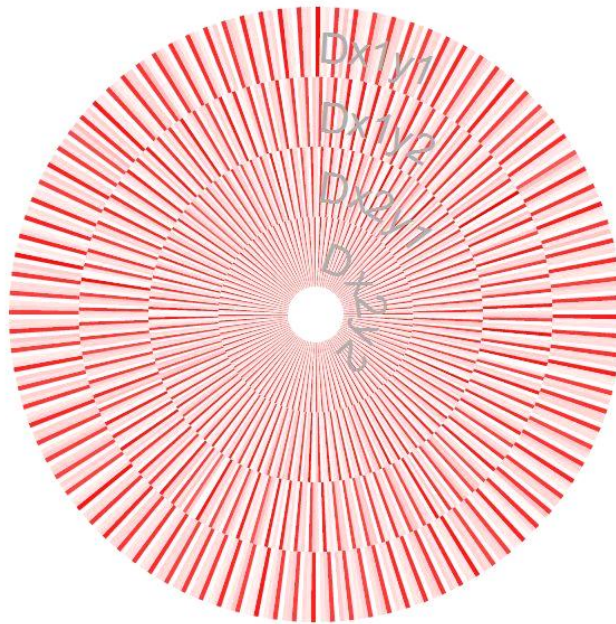

Figure 6.   The same dataset as shown as in Figure 5, lower right, using a circular layout for a heat map that takes up less screen space to show the same information.

A circular heat map is another way of arranging the values in this type of visualization. It provides the same functionality as the previous heat map layouts, but uses a different basic shape to create a more compact layout. Instead of arranging the individual values in a rectangular layout, a circular metaphor is used. Figure 6 shows an example using the circular layout with the same data as depicted in Figure 5. Note that the circular view uses less space to show the same amount of information. Figure 7 and Figure 8 also show the same dataset as Figure 5 (lower right), but using variations of the circular layout. Figure 7 follows a similar paradigm as Figure 5 in grouping the individual IDCMs by their responses; in the circular layout, each column (i.e. each response R_{xiyj}) is grouped into a quadrant of the polar array. Figure 8, on the other hand, breaks each block into a ring in the circle, which uses much less space to show each value in each block. In this case, the rows belonging to each response option are separated by a dark gray circle, grouping every four rings together.

Across these circular layouts, GRT experts like Figure 7 the most for a quick grasp of the IDCM confusion patterns. While Figure 6 does track the individual rows as rings of the circle, it is harder to separate the columns of each individual IDCM; the data in the innermost ring (D_{x2y2}) is also quite small and the colors blur together to become indistinguishable. This is not useable on a small screen or in print because of the perceptual difficulty distinguishing the data. It is likely this would happen for any color map used in Figure 6. The rings in Figure 8 capture patterns within each ring well, but experts find it hard to compare the same response ring across the different stimuli (D_{x1y1}-R_{x1y1}, and D_{x2y2}-R_{x1y1}, for example) because of the extra rings in between them. With the quadrant layout in Figure 7, domain experts are reminded of the traditional means of illustrating GRT model space (Figure 1(b)) which plot a bivariate normal curve in each of four quadrants representing the four stimuli. Thus, Figure 7 may lend itself to easy integration with other GRT plotting tools, which may provide a means of visualizing both the "raw" IDCM data together with GRT model analyses. Domain experts are eager to see further development of visualizations based on Figure 7.
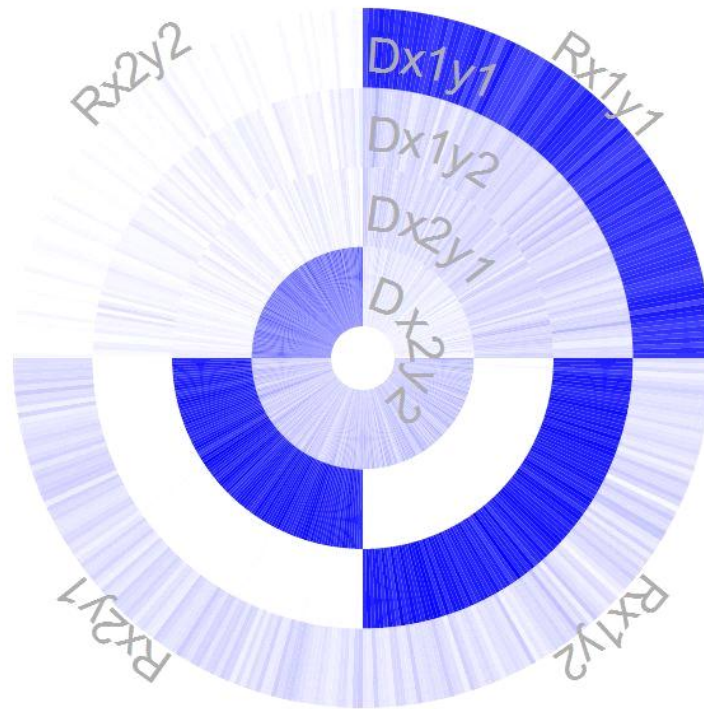
Figure 7.   Circular layout showing the same dataset as Figure 5, lower right.

6. CONCLUSIONS

Information visualization as a tool has matured into a valuable analysis technology. It is based on computer graphics, as well as general graphics principles and methods. It converts massive data into graphics, images, and other visual representations. It involves computer graphics, image processing, computer vision, computer-aided design, graphical user interfaces, and many other research areas. Humans obtain information retrieved from external data mostly through the visual channel, which has high-speed, high-capacity, and parallel features. Hence the sayings "seeing is believing" and "a picture is worth a thousand words" are particularly relevant. With the rapid development of human society, more people in scientific research and production environments are producing ever-increasing amounts of data. To cope with the information overload, appropriate visualization techniques are needed. Within this study, several different visualization approaches were developed to determine those methods that are most suitable to visualize the data from the IDCMs simulated for GRT models, namely parallel coordinates and modified heat maps using different layout approaches. As is expected, some are better suited than others depending on the task the domain specialists are trying to achieve as pointed out during the discussion of the results. Overall, the visualization methods were very successful in visualizing aggregated IDCMs in a way that allowed the domain specialists to easily interpret the changes in the patterns of confusion across different models. While the domain specialists were very pleased with the current results, a more formal user study will have to be performed as part of the future work.

REFERENCES

[1] Ashby, F. G. & Townsend, J. T. "Varieties of perceptual independence." Psychological Review, 93, 154-179, (1986).
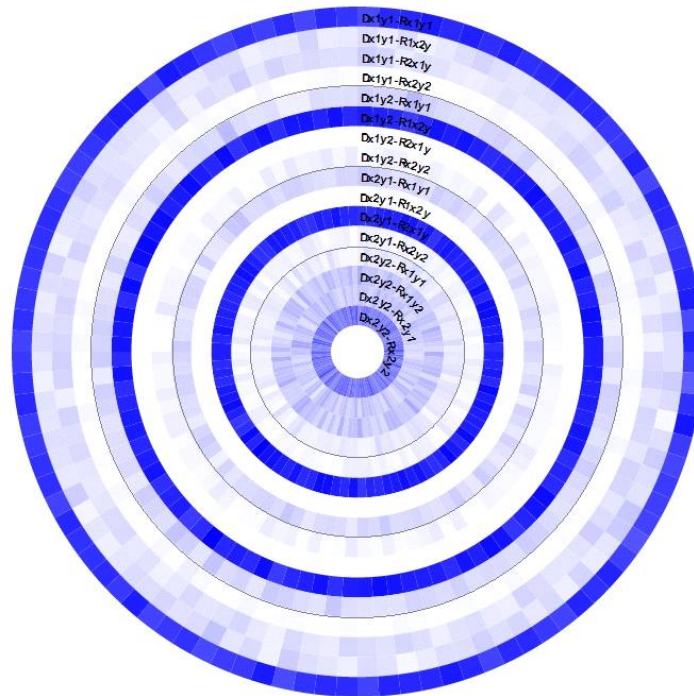
Figure 8.   Heat maps of the same data-set with Figure 5 (lower right), but it breaks each block into each ring in the circle, and will use much less space to show each value in each block.

[2] Menneer, T., Wenger, M. J. & Blaha, L. M. "Multiple methods of of modeling and detecting perceptual and cognitive configurality. 9th International Conference on Cognitive Modeling, Manchester-ICCM, Manchester, GB, (2009).

[3] Thomas, R. D. "Characterizing perceptual interactions in face identification using multidimensional signal detection theory." In M.J. Wenger & J.T. Townsend (eds.) Computational, Geometric, and Process Issues in Facial Cognition: Progress and Challenges. Erlbaum Press, 193-227 (2001).

[4] Silbert, N. J. & Thomas, R. D. "Decisional separability, model identification, and statistical inference in the general recognition theory framework." Psychonomic Bulletin and Review, 20(1), 1-20, (2013).

[5] Kadlec, H. & Townsend, J. T. "Implications of marginal and conditional detection parameters for the separabilities and independence of perceptual dimensions." Journal of Mathematical Psychology, 36(3), 325-374, (1992).

[6] Blaha, L. M., Menneer, T., & Wenger, M. J. "Correlational Methods for Multidimensional Signal Detection: An Evaluation and Comparison with Standard Analyses." 2013 Vision Sciences Society Annual Meeting, Naples, Florida. (2013).

[7] Menneer, T., Wenger, M. J., & Blaha, L. M. "Probit analysis for multidimensional signal detection: An evaluation and comparison with standard analyses." Poster accepted for the 2013 Vision Sciences Society Annual Meeting, Naples, Florida. (2013).

[8] Silbert, N. H. "Syllable structure and integration of voicing and manner of articulation information in labial consonant identification." Journal of the Acoustical Society of America, 131 (5), 4076-4086, (2012).

[9] Silbert, N. H., Townsend, J. T., & Lentz, J. J. "Independence and separability in the perception of complex non-speech sounds." Attention, Perception & Psychophysics, 71 (8), 1900-1915, (2009).

[10] Inselberg, A., & Dimsdale, B. "Parallel coordinates for visualizing multi-dimensional geometry" Springer Japan, 25-44 (1987).

[11] Peng, W., Ward, M. O., & Rundensteiner, E. A. "Clutter reduction in multi-dimensional data visualization using dimension reordering". In Information Visualization, 2004. INFOVIS 2004. IEEE, 89-96 (2004).

[12] Hauser, H., Ledermann, F., & Doleisch, H. "Angular brushing of extended parallel coordinates". In Information Visualization, 2002. INFOVIS 2002. IEEE, 127-130 (2002).

[13] Siirtola, H. "Direct manipulation of parallel coordinates". In Information Visualization, 2000. Proceedings. IEEE International Conference, 373-378 (2000).

[14] Zhao, K., Liu, B., Tirpak, T. M., & Schaller, A. "Detecting patterns of change using enhanced parallel coordinates visualization". In Data Mining, 2003. ICDM 2003. Third IEEE International Conference, 747-750 (2003).

[15] Johansson, J., Cooper, M., & Jern, M. "3-dimensional display for clustered multi-relational parallel coordinates". In Information Visualisation, 2005. Proceedings. Ninth International Conference, 188-193 (2005).

[16] Wilkinson, L., & Friendly, M. "The history of the cluster heat map". The American Statistician, 63(2) (2009).

[17] Vester, D "Creating Heat Maps with Virtual Earth and SQL Server 2008" http://jkebeck.wordpress.com/2008/12/01/creating-heat-maps-with-virtual-earth-and-sql-server-2008 (December, 2008)

[18] Nierman, W. C., Pain, A., Anderson, M. J., Wortman, J. R., Kim, H. S., Arroyo, J., ... & Lu, C. "Genomic sequence of the pathogenic and allergenic filamentous fungus Aspergillus fumigatus". Nature, 438(7071), 1151-1156 (2005).

[19] Špakov, O., & Miniotas, D. "Visualization of eye gaze data using heat maps". Electronics and electrical engineering, 2, 55-58 (2007).

[20] Bostock, M.; Ogievetsky, V.; Heer, J., "D³ Data-Driven Documents", IEEE Transactions on Visualization and Computer Graphics, vol.17, no.12, pp.2301,2309 (2011)

[21] Chi, E. H. H., & Riedl, J. T. "An operator interaction framework for visualization systems". In Information Visualization, 1998. Proceedings. IEEE 63-70 (1998).

[22] Shneiderman, B. The eyes have it: "A task by data type taxonomy for information visualizations". In Visual Languages, IEEE, 336-343 (1996, September).

[23] Becker, R. A., & Cleveland, W. S. "Brushing scatterplots". Technometrics, 29(2), 127-142 (1987).

[24] Furnas, G. W. "Generalized fisheye views". In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '86), Marilyn Mantei and Peter Orbeton (Eds.). ACM, New York, NY, USA, 16-23 (1986).