International Journal of Trend in Research and Development, Volume 4(2), ISSN: 2394-9333
www.ijtrd.com

The Data Visualization Pattern and Methods in 1D, 2D, 3D and 4D

[1]Wasim Khan, [2]Narendra Kumar Sahu and [3]Navnit Choudhury
[1,2]Lecturer, [1,2]Government Women's Polytechnic College, Indore, Madhya Pradesh, India
[3]System analysts, Shri Vaishnav Polytechnic College, Indore, Madhya Pradesh, India

Abstract -- Data visualization is a new and promising field in computer science. It uses computer realistic effects to show the patterns, relationships out of datasets and trends. In this paper, we first get familiar with techniques of data visualization and some related concepts, then we will look some general algorithms to do the data visualization. To get deeper about it, we also have some discussion about multidimensional data visualization. With the combination of some methods, we present a new algorithm to do 4 dimensional data visualization.

Keywords -- *Data Visualization, Maps, Multidimensional, Graphs, 1D, 2D, 3D, 4D*

I. INTRODUCTION

Human has a very long history with basic data visualization, and data visualization is still a very hot topic today. The history of visualization was cast to some extent by available technology and the pressing needs of the time, they include: maps on walls, primitive paintings on clays, photographs, table of numbers (with rows and columns concepts), these all are the way of data visualization – although we may not call them under this name at that time. Now a day's visualization very hot and is begin using in every field like Big data, healthcare analysis, social trends, finance, banking etc.

Visualization is the art of presenting the information in graphical way, with the goal of providing the viewer an easy way to understand the information contents. It is also the process of transforming objects, numbers and concepts into a form which is visible to the human eyes. When we say "information", we may refer to processes, data, relations, or concepts.

Data visualization is all about understanding relationships and ratios among numbers. Not about understanding individual numbers, but also about understanding the trends, patterns and relationships that exist in the set of numbers [8].From the point of user understanding, it may involve measurement, detection, and comparison, and is enhanced by different types of interactive techniques and also by providing the information from multiple views and with multiple techniques.

A. Why we do with data visualization?

To see and understand images is one of the natural instincts of human, and to understand statistical data is a year's training skill from schools, and even so, many people are still not good with numerical data [8]. A well-drawn picture, make much easier to find the trends and relations. Because visual presentation helps human to understand the data very fast. Data visualization shifts the load from numerical to visual reasoning. Gathering information from pictures is more time-saving than looking through numbers and text – that's why many decision makers would prefer to present information in graphical form rather than textual form .Another thing we should mention is that: data visualization is NOT as scientific visualization. Scientific visualization uses simulation, animation and sophisticated computer graphics to create visual models of structures and processed that can seen in sufficient detail. While data visualization is a way that display and present information in a way that encourages appropriate interpretation, association, and selection. It utilizes human skills for trend analysis, pattern recognition and exploits the ability of people to extract a great deal of information in a short period of time from visuals presented in a standard format.

The Importance of Visualizations in Business In A graph can communicate more information than a table in a smaller space. This trait of visualization makes them more effective than tables for presenting data. For example, see the table below, and try to spot the month with the highest sales.

Month -Jan Feb Mar Apr May Jun

Sales -45 56 36 58 75 62

This data when visualized gives you the same information in a second or two.

Figure 1: Plot of month with the highest sales

We have some commonly used data representation ways:

Line Graphs -A line chart or line graph shows the relationship of one variable to another variable. They are mostly used to track changes over time. These types of charts are very useful when comparing multiple items over the same time period (see Fig 2). The stacking lines are used to compare the individual values for several variables.

You may also want to use line graphs when change in a variable or variables clearly needs to be displayed or/and when trending or rate of change information is of value. It is very important to note that you shouldn't pick a line chart merely because you have data points. Rather, the number of data points that you are working with may dictate the best visual to use. For example, if I only have 10 data points to show, the easiest way to understand those 10 points might be to simply make a list of them in a particular order using a table. When deciding to use line chart, you should need to consider whether the relationship between data points needs to convey. If it does, and the values on X axis are continuous, a simple line chart may be what you need

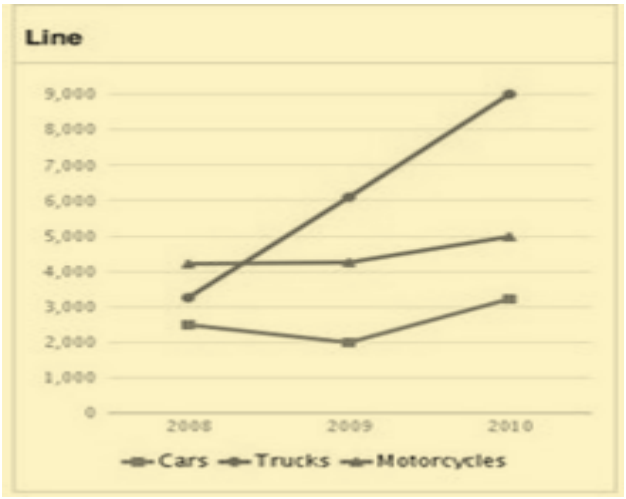


Figure 2: Line Graph

Pie Chart -Pie Charts are suitable for illustrate the results of a single variable that has less outcomes and the outcomes are shown as percentages rather than absolute numbers. It provides a visual aid to grasp the contrast in those outcomes.

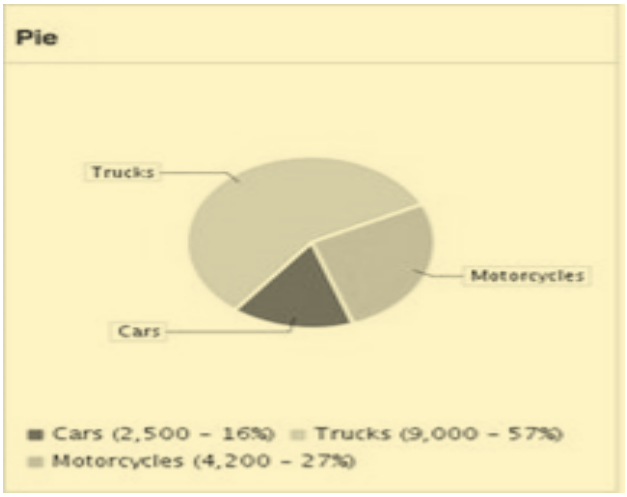


Figure 3: Pie chart

Stacked bar chart- A stacked bar chart is also known as stacked bar graph. It is a graph that is used to break down and compare the parts of a whole. Every single bar in the chart represents a whole, and segments in the bar represent different class or parts of that whole. Different types of colors are used to illustrate the different categories in the bar.

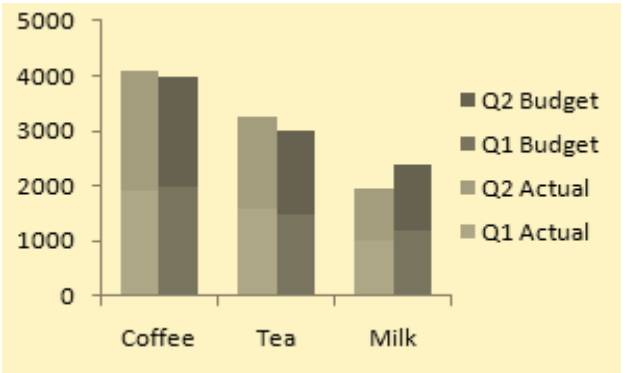


Figure 4: Stacked bar chart

Scatter Plots - A scatter plot (or we can say X-Y plot) is a two-dimensional plot that shows the joint variation of two data items. In a scatter plot, each marker (symbol such as stars, plus signs, squares and dots) represents an observation. The marker position indicates the value for each of the observation. Scatter plots also support grouping. If you assign more than two measures, a scatter plot matrix is generated. A scatter plot matrix is a series of scatter plots that displays each possible pairing of the measures that are assigned to visualization.

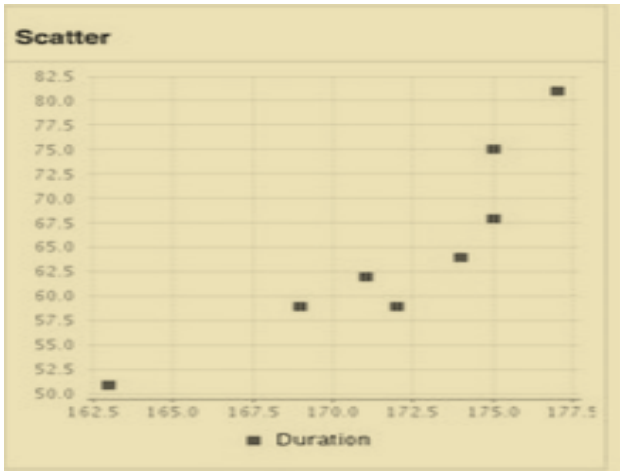


Figure 5: Scatter Plot

B. Maps

Maps are the graphs where the x and y axes correspond to distance: feet, meters, and kilometer whatever you want. Thus, an x-y coordinate represents a location of some place. It is one of the most effectual graphing techniques for map.

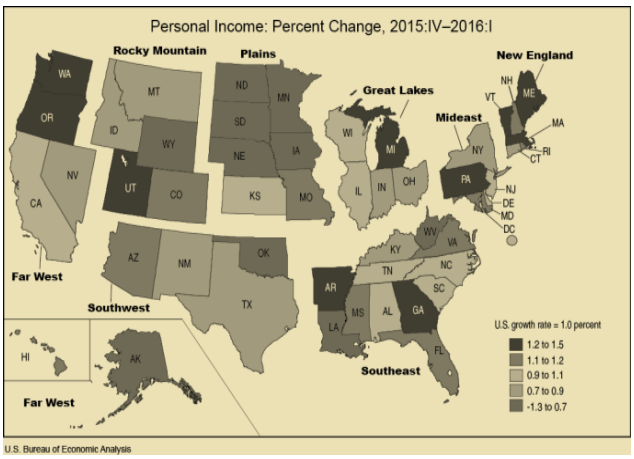


Figure 6: Map Graph

3-D Graph - A 3-dimensional *graph* is the *graph* of a function f(x, y) of 2 variables, or the *graph* of a relationship g(x, y, z) among 3 variables. Provided that x, y, and z or f(x, y) are real numbers, the *graph* can be represented as a planar or curved surface in a 3-dimensional Cartesian coordinate system. Example- Figure 3d graph for the expression x*x-y*y where x range (-50, 50) and y range (-40, 40)

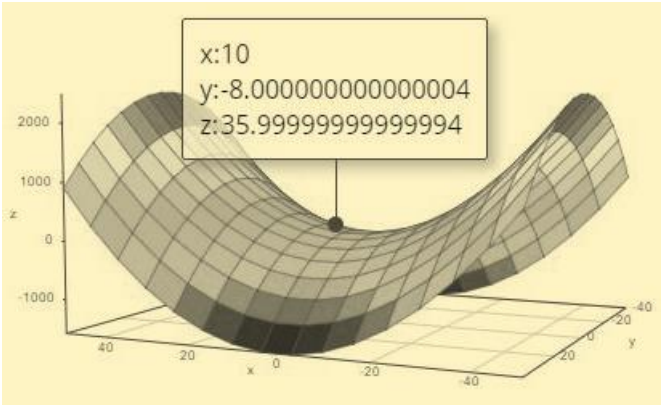


Figure 6: 3D-Graph

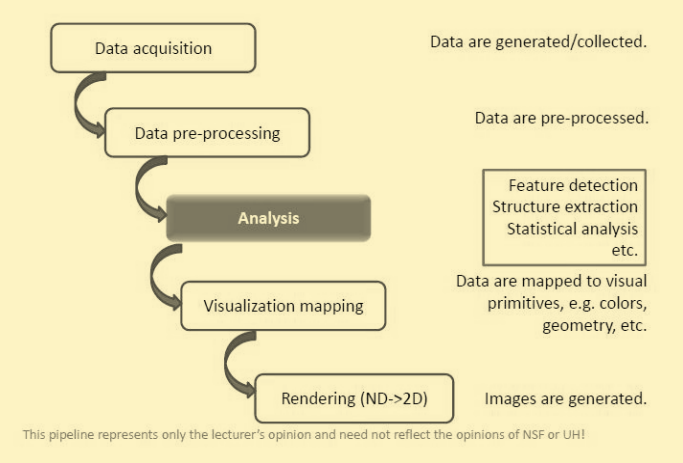


Figure 7: Data visualization analytic pipeline

C. Multi-dimensional Data Visualization (N>=3)

a. Visualizing Big Data

Big data brings challenges to visualization because of the size, speed and diversity of the data that must be taken into account. The most common definitions of big data is data that is of such velocity, volume and variety[7] that an organization must move beyond its comfort zone technologically to derive intelligence for effective decisions.

Data has become an integral part of politics, history, economics, business structures and science, and now even social lives. This trend is clearly visible in social networks such as Twitter, Face-book and Instagram where users produce an enormous stream of different types of information daily (music, text, pictures etc.) [1]. Now, scientific, government, and technical laboratory data as well as space research information is available not only for review, but also for public use. For instance, there is the 1000 Genomes Project [2, 3], which provide over 260 terabytes of human genome data. More than 22 terabytes of data are publicly available at ClueWeb09 [6], Internet Archive [4, 5] among others.

Efficient visualization tools should consider cognitive and perceptual properties of the human brain. The data is so huge that the human cannot imagine and it is also not possible to real the data one by one. So we use visualization tools to show trends and prediction. Various visualization tools are in the market like table au, raw, digraphs, and d3.js to enable real-time human interaction (scaling, touring etc.).

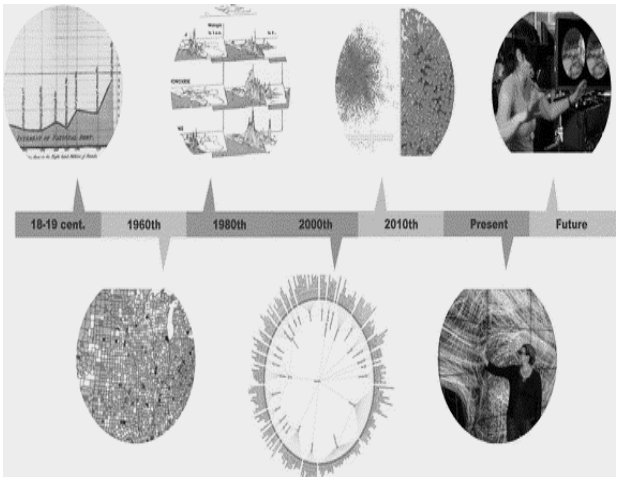


Figure 8: The rise of visualization methodology. Development of visualization methods start from 18th century and it is speedily improving today due to technical sophistication.

D. Data Visual Analytic Pipeline

Many of the data visualization methods are taken from the past days. When paper publish industry control the world. As they derive from paper publish, they take use of the paper as a media, which is a 2-D media. They handle 2-Ddata quite well. They also can present basic 3-D data with the help of projection. But when it comes to higher dimensional data, these data visualization technique gets failed.

When talking about N-D (N >=3) data visualization, we have many ways to do it. They are:

1. Translate to N-1 dimensional data for visualization.
2. Use special viewing instrument
3. Use special viewing methods.
4. Use ordinary 2/3-D algorithms plus various attributes to represent detain other dimensions.
5. Use animation.

We exam these ways one by one. We will mainly focus our attention on 4D data visualization if applicable.

E. Translate to n-1 dimensional data

This is the most used way to handle N >=3 data visualization. The primary method for it is projection. Given an N-D data, we may first project it into (N-1)-dimension, if required, we also can continue our projection into (N-2)-Dimension, until we can properly handle the visualization. For example, if we visualize data set of (x, y, z), then we use (x, y, z) as three axis to get a 3-D object, given the assumption that we have some way to picture this3-D object for our data examination (parallel viewing projection).

For example: Another method of translate N-D data to N-1 is also invested [9].In the case where 1D has a very finite range. It is possible to effectively combine2 dimensions, A and B into a single dimension which is called as C. We can call this multiplexed dimension.

F. Use special viewing instrument

One type of special glasses are called "stereoscope". With these glasses, when looking 2 intentionally prepared color pictures, one can see a 3dimension image out of these two2 dimension pictures. With the help of this, we can visualize our 3dimensiondata into 3dimension objects, and present them on two 2dimension pictures, while viewers can get the 3dimension image easily.

G. Use special viewing methods

As I shall claim that it is human that is the viewer of the result of data visualization. So, scientists are trying to find whether human eyes, without any other subordinate, can directly get some N-D information out of some (N-1)-D data representation. The research is successful in 3 dimensions, but (I shall say) "very limited". By purposely adjust our eyes' focusing point; we can get a 3 dimension image out from a 2 dimension paper image without any other aid. This is sometimes called "stereograph". But the big disadvantages are: it is not a convenient way. And while some people can do this, some people just can't.

I. Use ordinary 2/3-D algorithms plus various attributes to represent

Data in other dimension directly put, 4=3+1 (N= m + i, m<N). The primary idea is to present a 3-Drepresentation showing some of the points of the object, and to add the 4D information to it. There are two broad ways to do it, one is using some

graphical attributes to present our 4th dimension data, and the other one is using some non-graphical attributes to do so.

J. Using graphical related attributes to present the 4th dimension

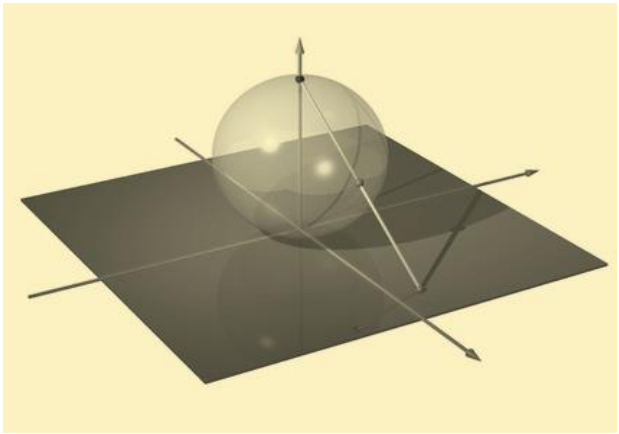
This is a very natural way, because when we recognize information, we are so used to use our eyes. Basically, we have several methods to do so.

H. Color

Colors play a great role in data visualization, where they are used to measure, label and enliven data. In multi-dimensional data visualization, for example in 4-D, we can use the first 3-D to construct a set of 3D objects, then we use different color for different points to indicate the difference in the 4th dimension data. One thing to point out it, the data point can either be in a 2-D plane, or be a 3D or higher dimensional space. Color is only specified by three dimensions: value, hue and Chroma. Hue is the name of color, such as red, green or blue. Value is the perceived darkness or lightness of the color. Chroma shows its colorfulness. High Chroma colors are saturated, low Chroma colors are grayish. From our experience, we know that if we use many colors, the view become badly for us. We could not easily figure out data points among different colors.

I. Size of the point

Much like the color above, size is an attribute of data point is also important in data visualization. One thing I want to mention: that the "point" here is actually an "object", not as the same as "point" in mathematics. Same as above, the data point can either be in a 2-dimensional plane, or be a 3 or high dimensional space. Several main drawback of size include: it may be difficult to implement in conjunction with any of the projection that introduces a decrease in the size with apparent distance. And, the size of each data point will greatly decrease the total number of data points in one view.



J. Using non-graphical related attributes to present the fourth dimension (Sound)

As the popularity of computers and multimedia in computer is another way is recently under research: usage of sound to present the high dimension data. Some principles are: the sound should not be interfere with the visualization process; it would be simple as most of us can understand with combination of the visual information, but not too easy that becomes unsatisfactory and annoying; should be some interactive, users can have some control.

K. Use animation

Actually, a lot of higher dimensional visualization (especial 4Dimension) programs use animation to achieve their purpose. In the animation, with moving (changing) 3 dimension visualization, the 4th dimension – time, is also represented. Examples include [10] [11] [12].One of my opinion: if none of dimensions is time, this method can also be acquire by carefully selecting one dimension/attribute and transform it as the "time "in the dataset (how other dimensions change according to this attribute?).

II. MY 4-DIMENSION DATA VISUALIZATION ALGORITHM: DYNAMIC COLOR PLANE

My specified 4-D data visualization algorithm (called "dynamic color plane "a multi-dimensional data visualization algorithm) takes the benefit of human

eyes' is complex functions toward colors. We use a color which is based visualization algorithm. And we take the advantage of animation in our algorithm.

We use a dynamic color plane to move through the 3D data point's space, and dynamically changing the colors of the points that interact with the plane. We can change moving direction of this plane, allowing the user to study any relations between the points and the colors. By doing so, the user may effectively figure out the distribution of the colors, and the tendencies of the data, which may be very helpful to the user (viewer).

This algorithm is better than existing color based algorithms, for the reasons:

In some conditions as will be mentioned in next section, my algorithm acts as the general visualization algorithms that make use of colors. So my program would be better as these visualization programs. In other state when the dense data points prevent visualization effects in the generally algorithms, my algorithm will be better than other because we take use of:

1) Clustering technology.

2) Our described dynamic color plane algorithm moving.

By the way, general algorithms that assist visualization can also be intergraded into my algorithm.

CONCLUSION

I'll begin this section with expression of a hope for the future, rather than a positive prediction of what will happen. New advances in data visualization have merged from scientific research, rooted firstly in studies of visual perception and then in human cognition. These studies have explored the capacities and limitations of both: produce data visualization methods and applications that take advantage of our most powerful potential and work around many of the limitations that hinder us. As such, data visualization is well equipped to assume a central role in the business intelligence, for it is intelligence that it is tailored to foster.

One of the problems that infestation data exploration and analysis involves large amounts of data that can be strenuous to visualize in ways that don't overwhelm the user or hide what is important behind a wall of clutter. This is a problem that has been receiving a great deal of recent attention by research community. Methods are being traverse and sophisticated algorithms are developed to tame the quantity of data either by decreasing the amount in ways that avoid loss of meaning, or by decreasing visual clutter in the visualization itself through novel approaches for the positioning of data objects, better uses of color and other visual attributes such as transparency. So this work is ideal for being included in business visualization software that is already effective.

References

[1] Chui M, Brown B, Manyika J , Bughin J, Roxburgh C, Byers AH, Dobbs R. Big Data: the next frontier for innovation, productivity and competition. June Progress Report. McKinsey Global Institute (2011).
[2] A Deep Catalog of Human Genetic Variation. 1000 Genomes: http://www.1000genomes.org/.
[3] The 1000 Genomes Project: Via M,Burchard EG, Gignoux C.
[4] Internet Archive Wayback Machine (2015).http://archive.org/web/web.php.
[5] Comparing content in web archives: differences between the Danish archive Netarkivet and Internet Archive.Nielsen J. In: Two-day Conference at Aarhus University, Denmark. 2015
[6] Lemur Project: The ClueWeb09 Dataset. (2015). http://lemurproject.org/clueweb09.php/.
[7] Laney D, Beyer MA. The importance of "Big Data", Stamford: Gartner; 2012.Google Scholar
[8] Mark Chignell, Kamran Parsaye, "Intelligent Database Tools & Applications".(QA 76.9 Dbm.PS)
[9] Shaun Bangay "Visview: A system for the visualization of Multi-dimensional data", in "Visual Data Exploration and Analysis V". (TA1505 Pse 3298).
[10] Laboratory for Scientific Visual Analysis, http://www.sv.vt.edu/classes/ESM4714/exercises/exer9/exer9.html,http://www.sv.vt.edu/index.html
[11] Dimension examples, from http://www.stanford.edu/group/4D/sections/4dexamples.
[12] Helena Mitasova, Bill Brown, Lubos Mitas,Irina Kosinovsky, Dave Gerdes,John Isaacson,Terry Baker, "Interpolation and Visualization from 3 dimension and 4 dimensionscattered Data Using GRASS GIS" from http://www2.gis.uiuc.edu:2280/modviz/viz/vol1.html