Multidimensional Visualization Technique Viewer (MVTView)

David B Heise
Case Western Reserve University
EECS 466 - Computer Graphics, Spring 2004

MVTView is a program that takes multidimensional data and allows the user to choose one of several multidimensional / multivariate visualization techniques (MVT) to view the data. MVTs are employed and needed more often than most people realize. This project aims to implement a small range of available techniques as a means of comparison and evaluation for potential uses. The two basic categories of MVTs are symbolic and non-symbolic[1] . Symbolic representations are also referred to as icons or Glyphs. These employ symbols to represent the data'smulti-dimensionality. Examples of Symbolic MVTs include: Stars, Sun Rays, Polygons, Piles[2], Kleiner-Hartigan Trees and Castles, and Chernoff Faces[3]. Non-Symbolic MVT rely on formulas or more traditional graphing techniques to visualize the data. Examples include parallel graphs, Andrew's plots, and physics based models[4,5].

Another possibility, which has the ability to incorporate both representations and is currently generating the buzz in this field of study, is the immerse visualization probe[6]. This technique involves surrounding the user (to different degree's depending on the implementation) by the data and allowing navigation and manipulation of this data in real to semi-real time. In general these techniques are fairly simple, straight-forward and easy to understand. They require little to no training, but an explanation of the axes is required in order to correctly analyze the data. Each technique has its advantages and disadvantages that make it a better or worse choice for any given application. Careful consideration of the specific data can allow for further customizations of all of these techniques. For example if the data has specific targets and limits involved, color coding of the data according to its relative location to these constraints would be a valuable addition.

Background:

Most data is multidimensional in nature, however other techniques are usually implemented that do not allow for visualization of all the available data. These other methods usually involve statistical methods showing simple numbers, which are very effective but not visually stimulating and can be easily misread or misinterpreted. Statistical methods are highly accurate but they are lacking in their use of the human system of perception. By modeling the data into glyphs (i.e. shapes and icons), or other visual techniques that model things that the human mind is accustomed to viewing and analyzing.

A Simple Example:

We start with a simple survey of people that includes: weight, height and age. These three items make the three dimensions of our data. To view this data is simple, we plot each of these three variables respectively along the x, y, and z axes of a standard three dimensional graph:

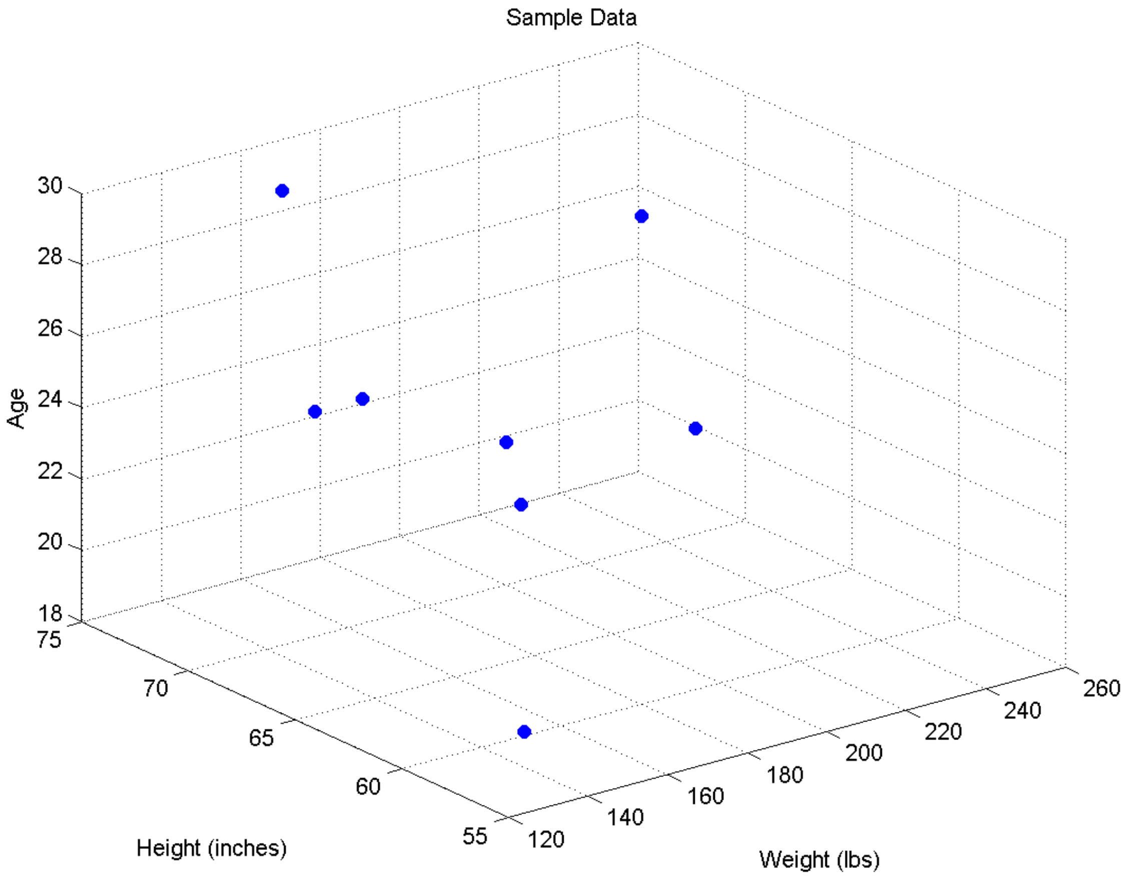

Figure 1 - Basic 3D Scatter Plot

Now the problem arises when our survey contains more than three questions (which is often the case). Lets take a more complicated survey to illustrate the problem, like a course review. A course review make ask any number of questions requiring the student to rate the question on some arbitrary scale that represents how well the student agrees/disagrees with the statement. Lets say that our survey has eight questions and is given to thirty students. This makes it a 8D problem with 30 data sets. It could be construed as one 9D problem, but we are specifying the last dimension (the student) and the comparison dimension. We'll use this example to illustrate some MVT, after having normalized (scaled between 0 and 1) the data.

Parallel Graphs:

The first MVT that we explore is the parallel graph. Here we take each dimension of the data and plot that dimension on its own axis, putting the different dimensions/axis parallel to each other in the same plane. For continuity we use lines to connect all the points in a data set.

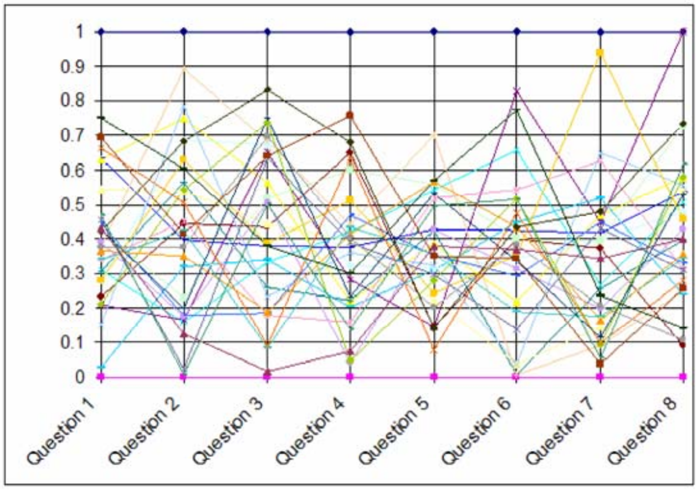

Figure 2 - Simple Parallel Graph

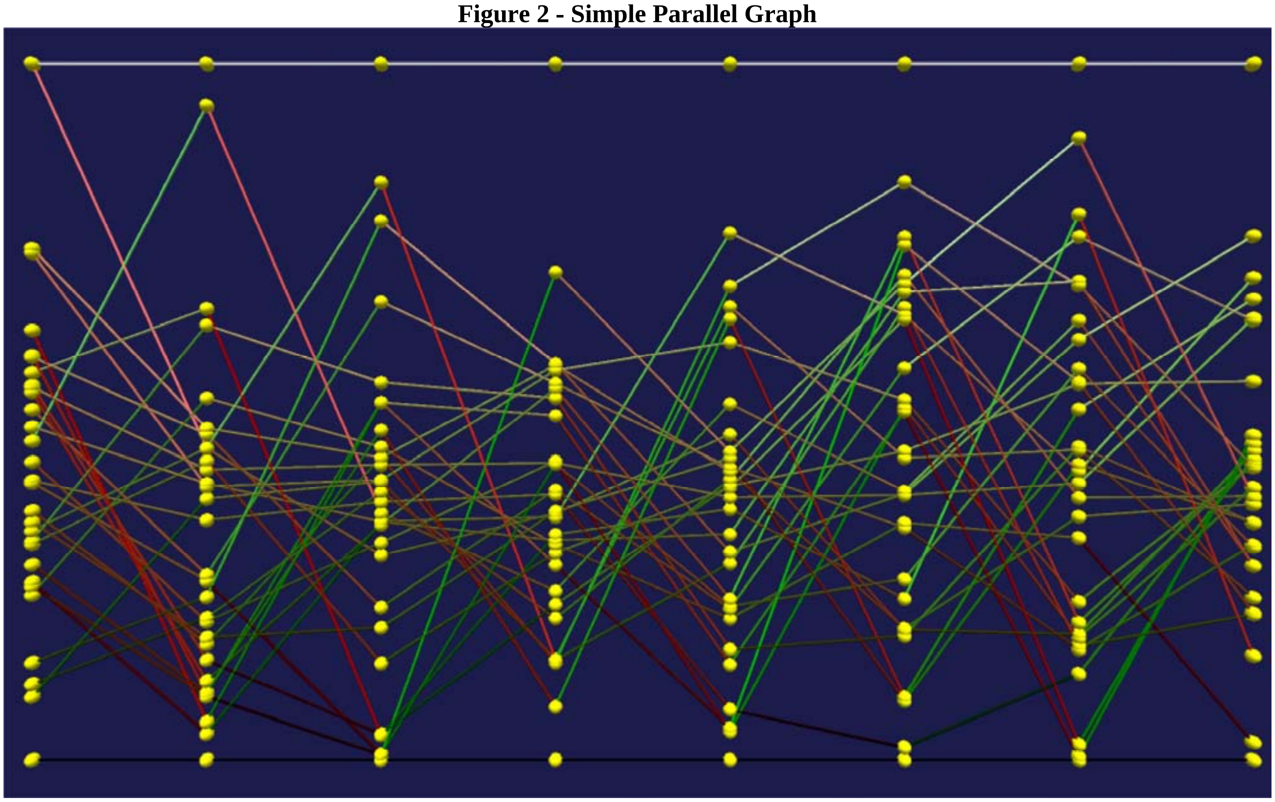

Figure 3 - 3D Parallel Graph by MVTView

Parallel graphs are useful in that they allow for an arbitrary number of dimensions and datasets to be shown. However, it is apparent, even with as few as thirty datasets, that parallel graphs quickly become cumbersome even though useful data can still be easily extracted visually. For instance clusters, minima, and maxima in a single dimension are still quite apparent, but following a single dataset through the graph can be difficult. The first problem that is apparent is that with so many data points, it can be hard to see which values go where and what it all means.

Stars:

Stars are a form of glyphs, or shapes. Like parallel graphs, stars take each dimension of data and plot it on a separate axis, however instead of plotting these axes parallel to each other, they are radially spread out from a central point. Two variations of the star are the sun ray (where each axes is extended to the full length and the data values are connected from axis to axis), and the polygon (where no axis are drawn and the lines are curved). Also each dataset is a separate star, although all datasets on a single star is possible.

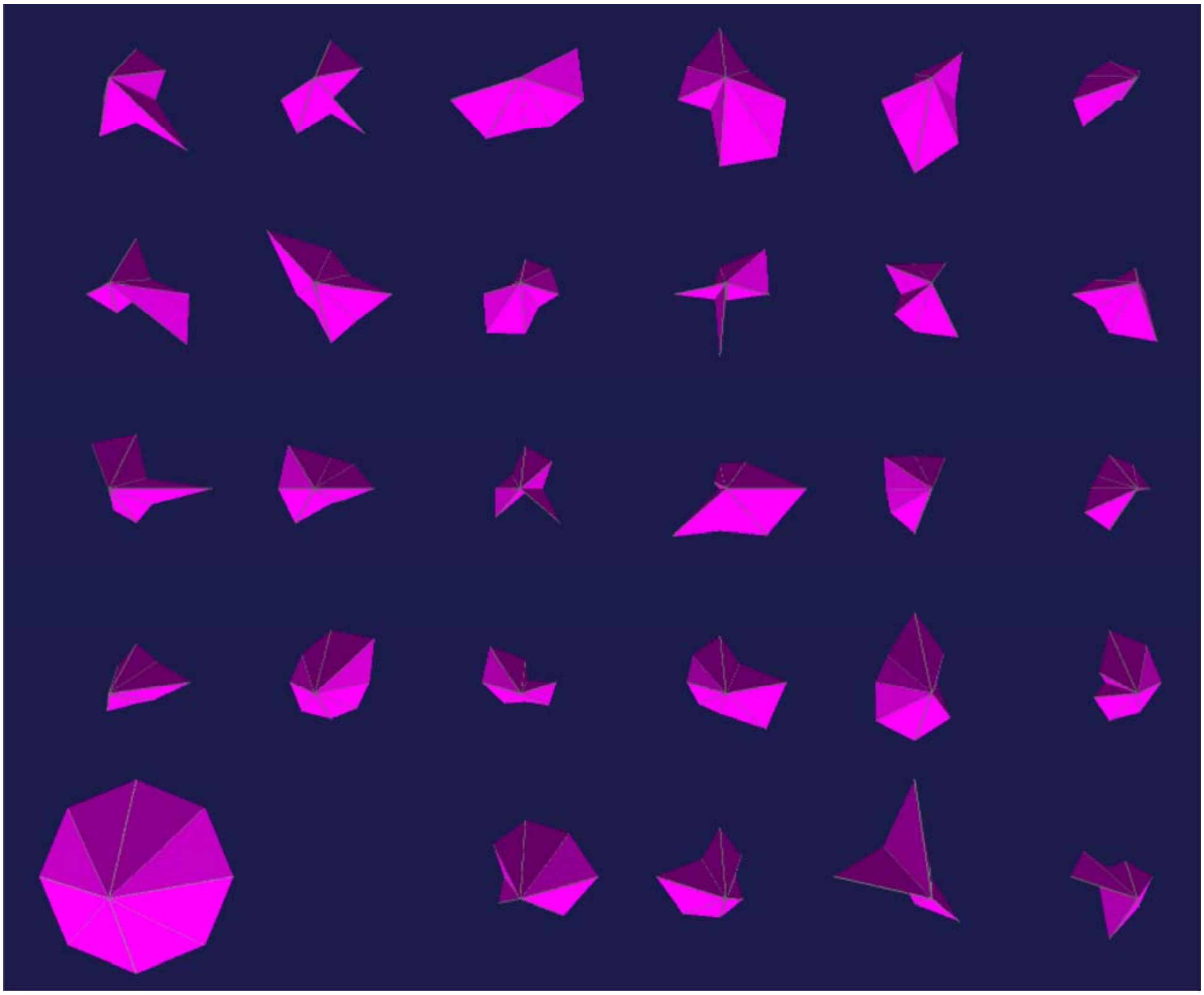

Figure 4- Stars

We can see from this graph that the stars do not suffer from the same overpopulation problem that the parallel graphs do. Each dataset is represented separately and it is very easy to see a not only view individual datasets, but the population as a whole easily spotting outliers in single or multiple dimensions.

Piles:

Piles are a method of stacking icon/glyph based visualization techniques. In our example we create piles of stars, which are called columns. These columns make it easier to view relationships between datasets on a single dimension. However they introduce other problems of overlapping and ordering of the data and the dimensions to allow for optimal viewing of the column.

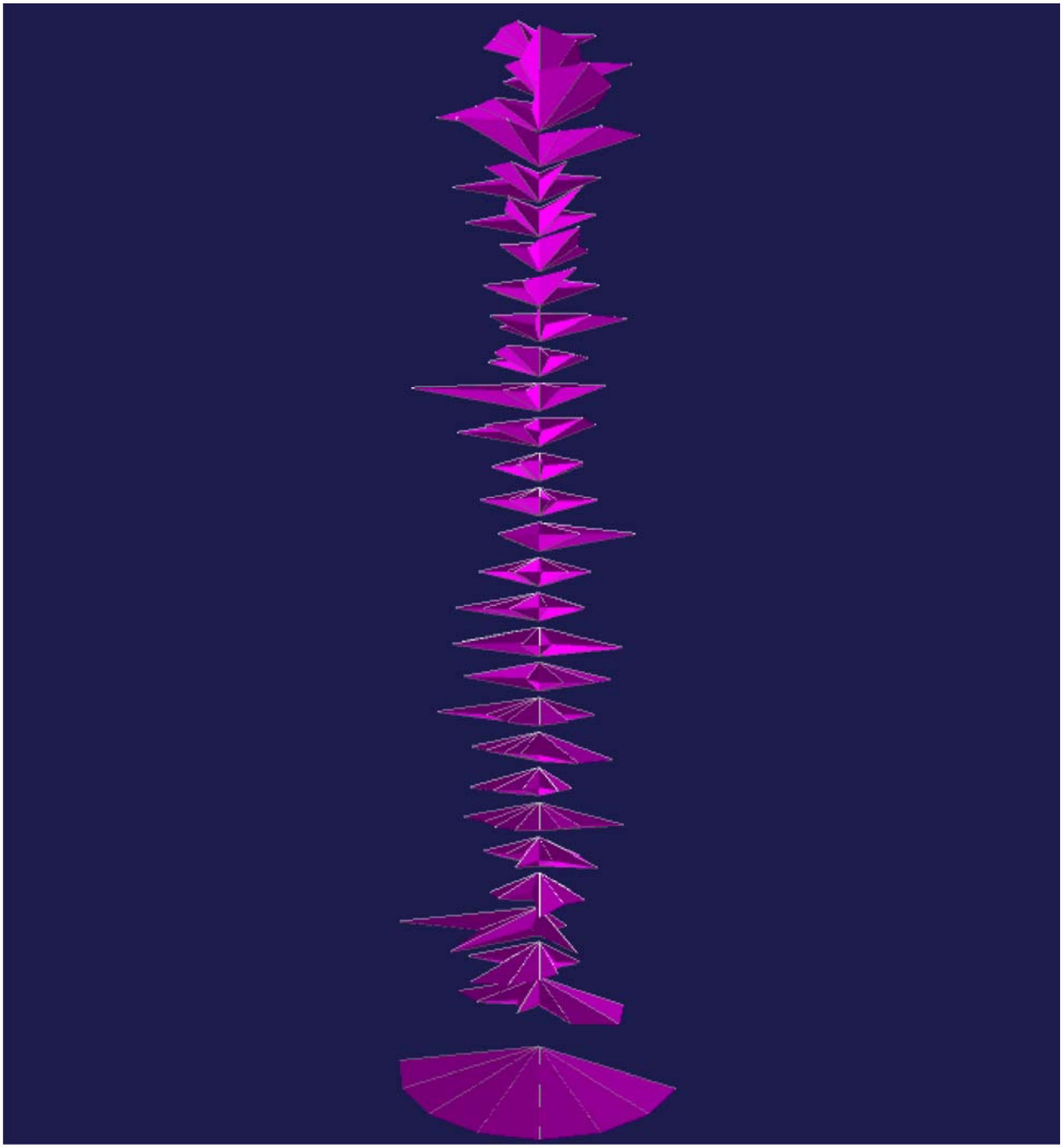

Figure 5 - Piled Stars

Chernoff Faces:

Chernoff faces use a facial image as the glyph modified by the data. Certain facial features are adjusted within acceptable range of realistic to semi-realistic values. The number of features to adjust are only limited by the same factors that would make one face different from another, however certain features are much more distinguishable than others. In order of importance: area of face, shape of face, length of nose, location of mouth, curve of smile, eyes (location, separation, angle, shape, and width), pupil location, and eyebrows (location, angle, and width)

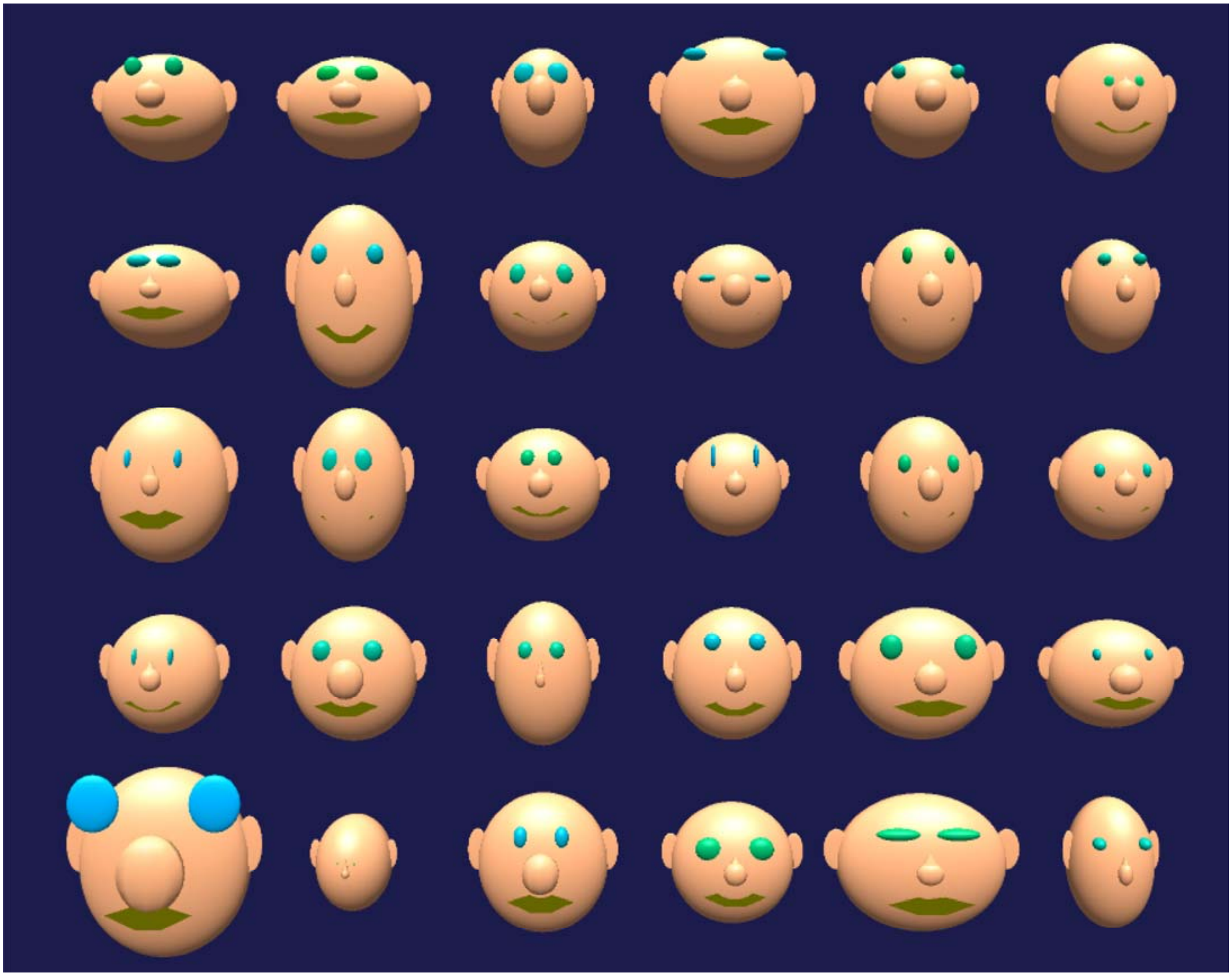

Figure 6 - Chernoff Faces

Force Maps, Vector Maps, and other physics based methods:

These physics based models of data visualization rely on physical equations of force, motion, or any other multidimensional equation to base the clustering of the various datasets. In our example here we've chosen to use a simple vector map with the addition of color.

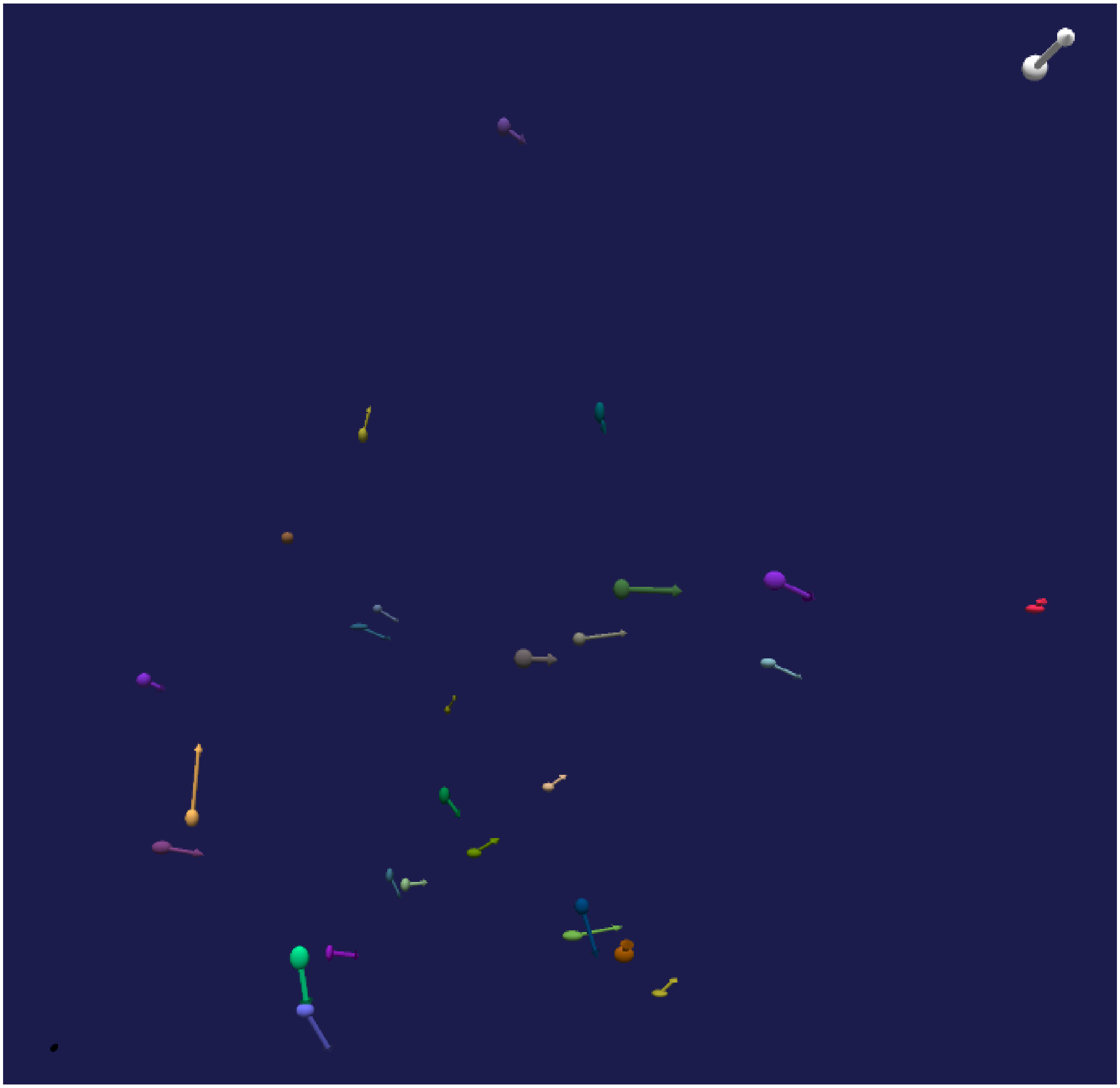

Figure 7 - A Simple Vector Map

Colors can be a extra dimension to the visualization of the data, however usually the colors are encoded in such a way that they either represent specific things, are scaled between a limited set of colors, or are a representation of dimensions shown in other ways. In our vector map, we've chosen to color the vector in a RGB fashion corresponding to the XYZ direction of the vector.

References

1. E Cluff, R P Burton and W A Barret; "A survey and categorization of multidimensional presentation techniques." Journal of Imaging Technology, 1991, 17(4), 142-153.
2. John D Melville, R P Burton; "Piles for Hyperdimensional Graphics"; Computers & Graphics, 1997, 21(1)
3. H Chernoff, "The use of faces to represent points in k-dimensional space graphically." Journal of the American Statistical Association 1973, 68, 361–368.
4. Holger Theisel, M Kreuseler; "An Enhanced Spring Model for Information Visualization"; Eurographics '98, 1999, 17(3)
5. Ming C. Hao, U Dayal, D Cotting, T Holenstein, M Gross; "Accelerated Force Computation for Physics-Based Information Visualization"; 2003
6. James R Miller, E A Gustavo; "The Immersive Visualization Probe for Exploring n-Dimensional Spaces"; IEEE Computer Graphics and Applications 2004, ?(1), 76-85