Project Description

# Data Description

Several categories of data sets were explored including health, social networking, and stocks in order to narrow in on an interesting data set with real world problems that need solved. This ultimately led to the selection of flight data as there are numerous aspects to flight data ranging from time-space-position information (TSPI) to parametric data recorded from onboard sensors monitoring the aircraft systems. Airlines collect volumes of this data, recorded periodically during each flight of each aircraft within the fleet. This data is critical for establishing competitive advantages and ensuring that all flights reach their destinations safely. One use of this data is condition-based maintenance, where the fatigue of a specific part on the aircraft is measured and compared to its expected lifetime. Another use is in training, where student pilots can observe a playback of an optimal flight and compare their own flights to this gold standard. Since the data sets can be quite large, even for a single flight, visualization techniques are a must have capability in order to effectively analyze and extract the required information.

A set of flights from a King Air B200 simulator has been acquired for analysis. While it not associated with a physical aircraft, the data collected is comparable and its analysis is quite relevant to solving the aforementioned problems. These data sets contain hundreds of dimensions and are recorded every 100 milliseconds during the flight, resulting in fairly large sets of data. This data is recorded during each flight with a single plane often flying several times in one day. Since the depth and breadth of the data is so large, specific subcategories will be chosen in order to isolate useful visualizations.

# Pre-Analysis

Potential methods for presorting or categorizing the data will be explored using some basic techniques such as dimension reduction, clustering, entropy, or establishment of hierarchies in order to narrow in on related and relevant parameters to analyze. It is important to note that these steps are used to prime and enhance the visualization and user analysis. Ultimately the user will be provided controls to adjust and interact with the data to customize the visualization.

# Visualization

The project will primarily focus on the use of glyphs to represent the data by mapping the dimensions to not only the shape of the glyphs, such as in star glyphs, but also to other attributes of the graphic such as the color, position, orientation, 3D representations, and textures. Providing calculations such as the total volume of the glyph or the volume of a specific sector within the glyph will also be explored, as a means to monitor current conditions against a known standard. Other potential user interactions will include the use of scatterplots to aid the user in configuring and setting up the visualization. New glyph shapes and applications will also be explored as a way to enhance the overall analysis and the amount of embedded information that can be concisely represented in the visualization.