



Figure 13: Scagnostics display panel of AutoVis. The data are statistics on 520 baseball players. The center panel shows the scatterplot matrix (SPLOM) of the scagnostics computed on these data. The left column shows two scatterplots considered by AutoVis to be exemplars for all the 120 possible scatterplots for the data set. The right column shows scatterplots that AutoVis considers to be anomalies.

Images

AutoVis recognizes a variety of graphic file formats. For image data, AutoVis computes the color distribution of the image itself.

JPEG

Figure 11 shows the result of dragging a JPEG file into the AutoVis window. We have selected a painting of Wassily Kandinsky. On the left are kernel density histograms of the HSB components of the image. On the right is a rotatable (through mouse movements) 3D scatterplot of the HSB components.

Unrecognizable files

Figure 12 shows the result of dragging a file with a format unknown to AutoVis. In this case, AutoVis provides a color-highlighted hex dump of the file. The file in this

example is a recent Microsoft Word document. We are working on decoding many variations of Word files, but parsing their structure from Java or other non-Microsoft environments is a non-trivial enterprise. This example illustrates the principle that the software should provide some representation of the data instead of crashing from failing to parse correctly.

Scagnostics

Figure 13 shows the scagnostics pane of AutoVis. The data are derived from statistics on 520 baseball players. The scatterplot matrix (SPLOM) in the center panel shows the nine scagnostics from Wilkinson *et al.*²⁶ computed on these data. On the left are two scatterplots that AutoVis considers to be exemplars. These plots were computed through a *k*-means cluster analysis of the scagnostics on all possible pairwise scatterplots. The algorithm continues to