SANDIA REPORT

SAND2002-0702
Unlimited Release
Printed March 2002

A Visual Empirical Region of Influence Pattern Recognition Tool for Leave-One-Out Data Analysis

R. F. Martinez

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865)576-8401
Facsimile: (865)576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.doe.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800)553-6847
Facsimile: (703)605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: http://www.ntis.gov/ordering.htm

SAND2002-0702
Unlimited Release
Printed March 2002

A Visual Empirical Region of Influence Pattern Recognition Tool For Leave-One-Out Data Analysis

R. F. Martinez
Laser Optics and Remote Sensing Dept.
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-1423

Abstract

In previous research at Sandia National Laboratories new pattern recognition (PR) algorithms based on a human visual perception model were developed. We named these algorithms Visual Empirical Region of Influence (VERI) algorithms. This document describes a graphical user interface tool developed to control the VERI algorithm and a visualization technique that allows users to graphically animate and visually inspect multi-dimensional data after it has been classified by the VERI algorithms. The VERI Interface Tool was written using the Tcl/Tk Graphical User Interface (GUI) programming language, version 8.1. Although the Tcl/Tk packages are designed to run on multi-platforms, we have concentrated our efforts to develop a user interface for the ubiquitous DOS environment. The VERI algorithms are compiled, executable programs. The commands to execute VERI algorithms are quite difficult to master when executed from a DOS command line. The algorithm requires several parameters to operate correctly. From our own experiences we realized that if we wanted to provide a new data analysis tool to the PR community we would have to make the tool powerful, yet easy and intuitive to use. That was our motivation for developing the VERI Interface tools. This document illustrates step-by-step examples of how to use the interface. This interface runs VERI in Leave-One-Out mode using the Euclidean metric. For a description of this type of data analysis, and for a general Pattern Recognition tutorial, refer to our website at: http://www.sandia.gov/imrl/XVisionScience/Xusers.htm.

Contents

Figures

Introduction

This paper discusses and exemplifies using the VERI algorithms in Leave-One-Out mode via a user-friendly interface. Screenshots of the interface windows have been captured and how to use them is described.

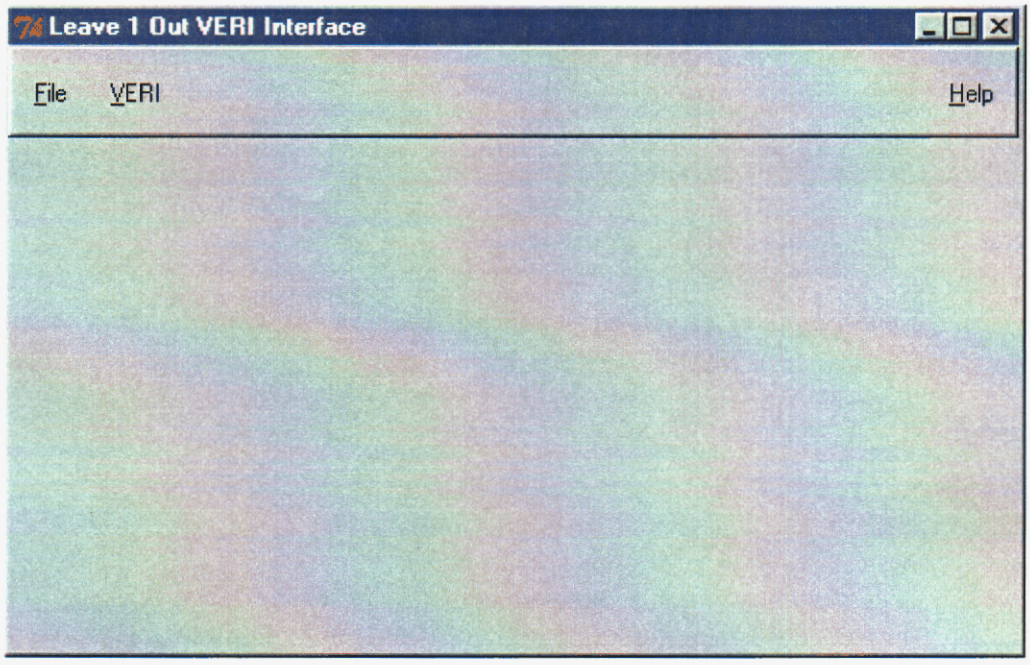

Figure 1. Leave-One-Out Main Interface Window

Figure 1 shows the top-level Tcl/Tk[1] window named "Leave 1 Out VERI Interface". It is created each time the Leave-One-Out interface is executed. There are two pull-down menus in the menu bar, the File menu, and the VERI menu. They are used to set VERI command parameters and for data visualization. A help menu also exists. In the help menu, the user is referred to user documents supplied with the interface.

The interface can be executed after the necessary software and hardware is acquired and installed on your PC system. There is currently no charge to acquire the interface. For information on how to acquire the VERI Interface package, contact the author at rfmarti@sandia.gov.

11.4296875	0.005347252	1.80006E-05	1
10.4375	0.019611359	0.001790166	1
9.3359375	0.029211998	0.001452088	1
9.0703125	0.0508461	0.00472486	1
5.23046875	0.01540947	0.000308871	2
13.6796875	0.035961151	0.002737999	2
4.3984375	0.036808014	0.003687263	2
18.8984375	0.000670433	6.69956E-05	3
18.4609375	0.000799179	0.000872612	3
17.96875	0.001009941	0.000922203	3
17.8671875	0.000728607	0.001067042	3
-5.66015625	-0.013422012	-0.000935555	5
-5.33984375	-0.018762589	-0.001707077	5
-5.390625	-0.018527031	-0.00110054	5

Figure 2. Sample contents of VERI spread sheet file named test.ss

Data Format

The VERI program reads the data it will analyze from a computer file. The user must manipulate their data to create files with a specific format. The VERI program is designed to read data from files saved in a spreadsheet format (hence the .ss in the file names). Figure 2 shows an example of what the contents of a properly formatted spreadsheet file named test.ss may look like. Spreadsheet format requires that the data on each row be separated by white space such as character spaces or tabs, and ended with a carriage return. Each row of data in a spreadsheet file is a spectrum of N measures of data (dimensions) acquired in a test setting, followed by an integer class value for that data. Each row contains all the dimensions of one data point. In Figure 2, there are fourteen data points, each point has three dimensions. There are four classes of data (1,2,3,5). All the data in a class must be stored contiguously.

A class of data refers to measurements acquired when the same types of inputs are being measured by the system. For example, to measure gas compounds, a test bed of sensors may be developed whose signal is proportional to the amount and type of gas accumulated by the sensors. In this system, all data measured (with varying concentrations) when acetone vapors are accumulated by the sensors are in one class. Likewise all data measured when methanol vapors are accumulated by the sensors are in another class.

The number of dimensions and number of points per class in figure 2 are only for illustrating proper file formatting. In most cases this would be considered a sparse data set. Not enough data has been measured to properly represent the classes of data.

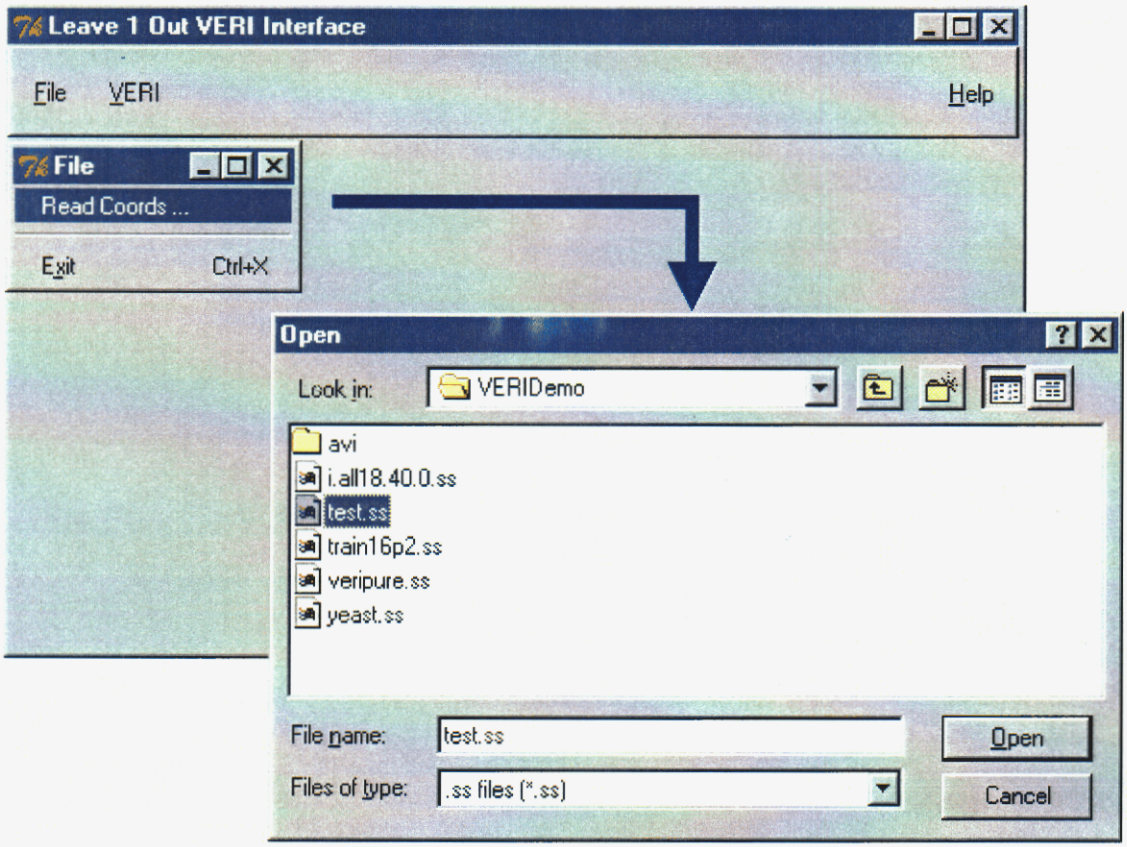

Figure 3. Reading data with the "Read Coords" sub menu in File menu.

The "File" menu is where the interface accesses the spreadsheet data files. Figure 3 shows an example of opening the "File" menu and selecting the "Read Coords" sub-menu. The user is browsing the computer hard disk to find a file to analyze. The file test.ss has been selected. The Tcl/Tk packages interface with built-in Windows functions and dialog boxes to read in file names. Double clicking on a file name or single clicking and pressing the "Open" button will read in a file name and close the dialog box. The interface defaults to list names of files of type ".ss files (*.ss)". These are spreadsheet files with names ending with the characters ".ss".

The "File" window can be opened and posted. Posting a window causes it to remain open until closed by the user. The "File" window is posted by selecting the menu name "File" in the main menu bar with the left mouse button and releasing the button when the dashed line at the top of the File pull down window is highlighted. The dashed line is not illustrated in figure 3. Clicking on the standard Windows close icon (X) in the upper right corner of the window will close the posted "File" menu.

Selecting the "Exit" sub-menu from the "File" menu shown in figure 3 will exit and quit the VERI interface. It can be selected at any time when then interface is running.

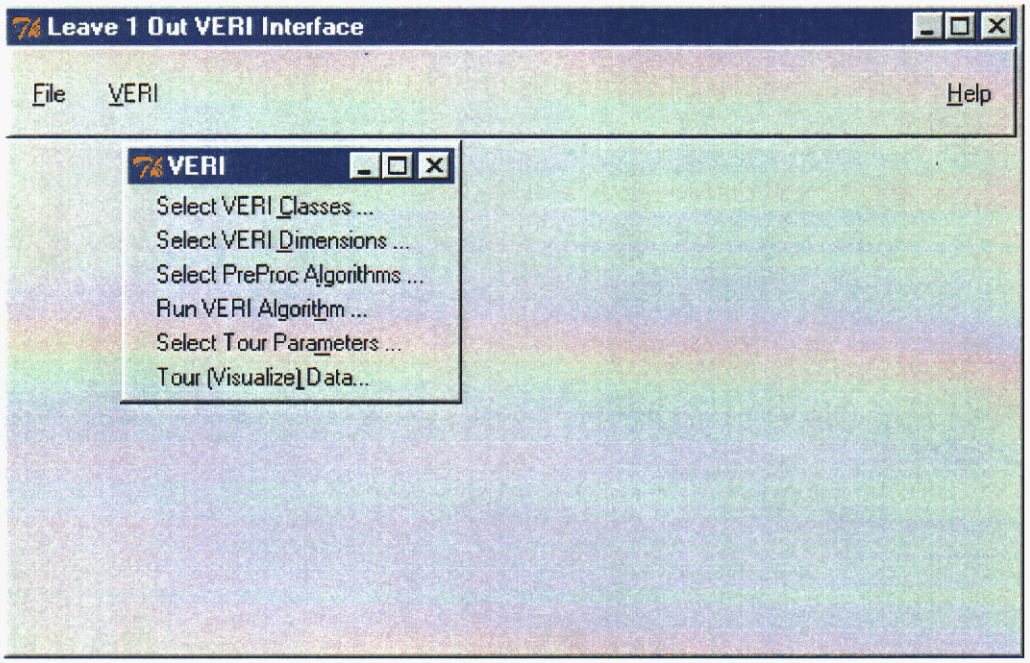

Figure 4. The VERI Pull Down Menu -- Runtime and Visualization parameters.

VERI Command and Control

Figure 4 shows the VERI pull down menu. The VERI window shown can be opened and posted in a manner similar to the way the "File" menu is posted. The VERI menu has sub-windows from which to select the remaining parameters for the VERI command line. There is also a sub-window to run the VERI algorithm with the supplied parameters, and sub-windows for data visualization. Each sub-window and the effects it has on the VERI results will be described next.

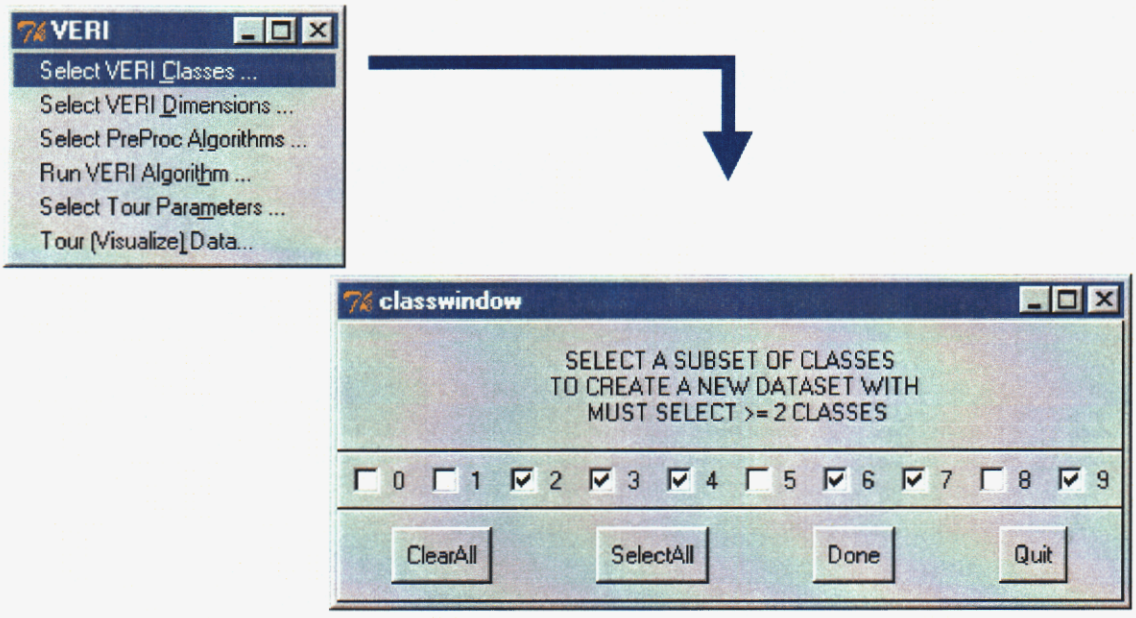

Figure 5. VERI Classwindow – select classes for VERI leave-one-out analysis.

Figure 5 illustrates the process of opening the "Classwindow" and selecting which classes of data to use in a VERI analysis. In this example, the data file has ten classes, numbered zero through nine. At least two classes must be selected. The classes can be selected individually by clicking in the box to the left of a class number with the left mouse button, or the buttons below the class selections can also be used. The operations performed by the buttons are self explanatory, but will be explained for the sake of thoroughness:

ClearAll – clears all currently selected classes
SelectAll – selects all available classes
Done – Saves selected class values and exits Classwindow
Quit – Does not save class values and exits Classwindow

The VERI results are generated for each class analyzed. The results usually vary with the number of classes analyzed. The more classes analyzed, the more likely it is that the VERI algorithm will find spatially overlapping points from multiple classes in the results. This information is useful in sensor development.

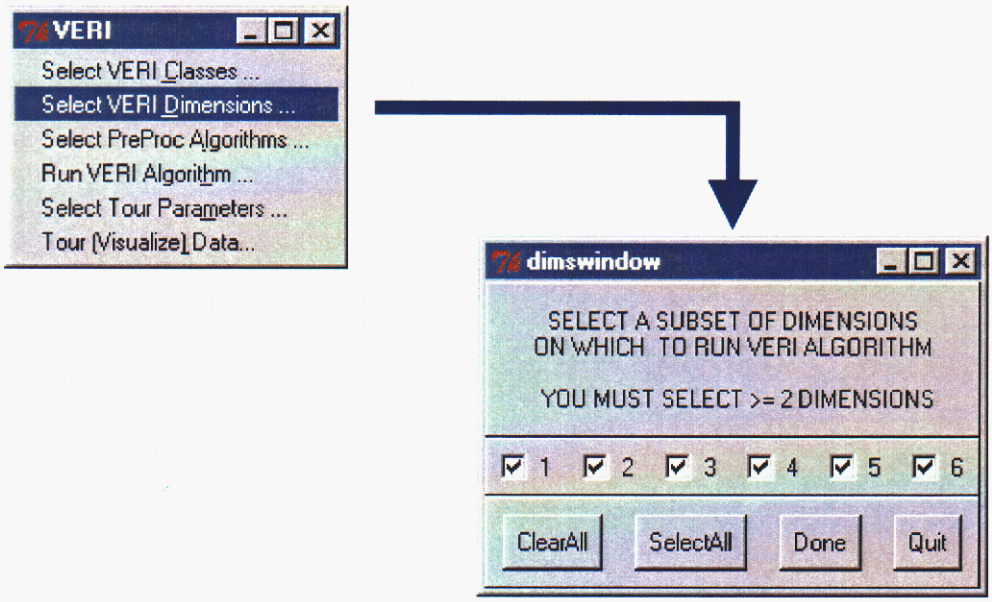

Figure 6. VERI Dimswindow – Select dimensions of data for analysis.

The "Dimswindow" shown in figure 6 is used to select which of the N dimensions of data you want to include in a VERI analysis. Different combinations of dimensions will generally produce different VERI results. Selecting all dimensions (as shown in this figure) is not necessarily the best selection. In pattern recognition it is often quality over quantity of data that matters the most to calculate the best results. For example, if you are keenly aware of the test station where the data you are analyzing was acquired, you may know innate characteristics of the sensors used to measure the data. You may know which sensors produced drifting, noisy or non-monotonic data. Likewise, you may know which sensors produced stable measurements. It is characteristics such as these that may be useful to help select the dimensions of data to use for your analysis. The user may also notice dimensions that reoccur in some of the best results generated and select them as part of an optimal subset of dimensions.

The process of maneuvering and selecting dimensions in "Dimswindow" is similar to the process described for selecting classes in the "ClassWindow" illustrated in Figure 5.

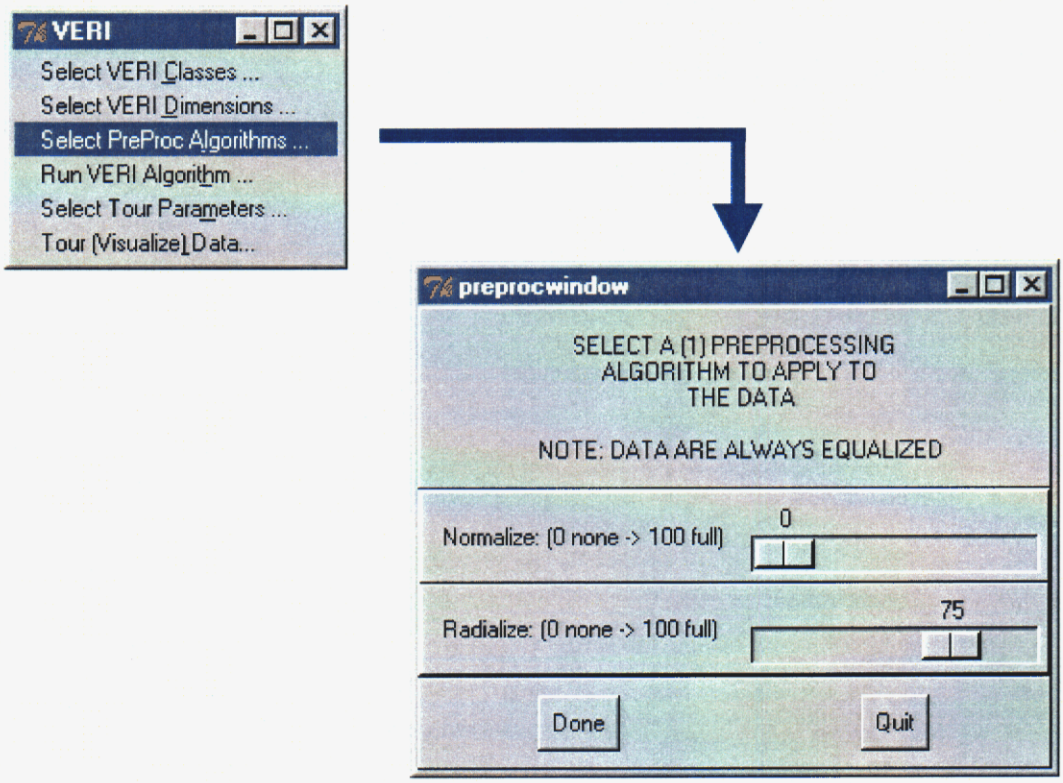

Figure 7. VERI Preprocwindow – processes to spatially separate data.

We have incorporated data processing algorithms as part of the interface to try and improve data separation and the VERI results. The best results are achieved when classes of data are well separated from each other and when the points in a class uniformly fill a volume. Using these algorithms attempts to accomplish these goals, but their use is optional.

Figure 7 shows the algorithms available under the window named "Preprocwindow" which have proven useful in spatially separating classes of data. Two algorithms are available, but only one can be applied per run of the VERI algorithm. It is best to think of the data points as multi-dimensional vectors with a magnitude and direction when using and describing these functions.

The Normalize function scales the magnitude of the data points in the direction of a unit sphere. As examples, a 25% normalize function scales each vector to a magnitude of 25% between the original magnitude and a magnitude of 1.0, a 50% normalize function scales each vector to a magnitude of 50% between the original magnitude and a magnitude of 1.0, and a 100% normalize function scales each vector to a magnitude of unit length 1.0.

The Radialize function attempts to introduce inter-class data point separation by moving points radially away from the origin. After the Radialize function is applied, the direction of each vector is maintained, but their magnitude is increased. The data points are further away from the origin after the Radialize function is applied. If you have points in several classes that spatially occur close to the origin, the Radialize function can serve to separate the points and classes.

The N dimensions of data that are selected for analysis are always equalized. Data that have been equalized occur in the range $-1.0 <= Eq <= 1.0$. Equalization eliminates sensor signal scaling and helps to spatially distribute the data in a uniform volume where the maximum distance from the origin to a data point is $\sqrt{N}$.

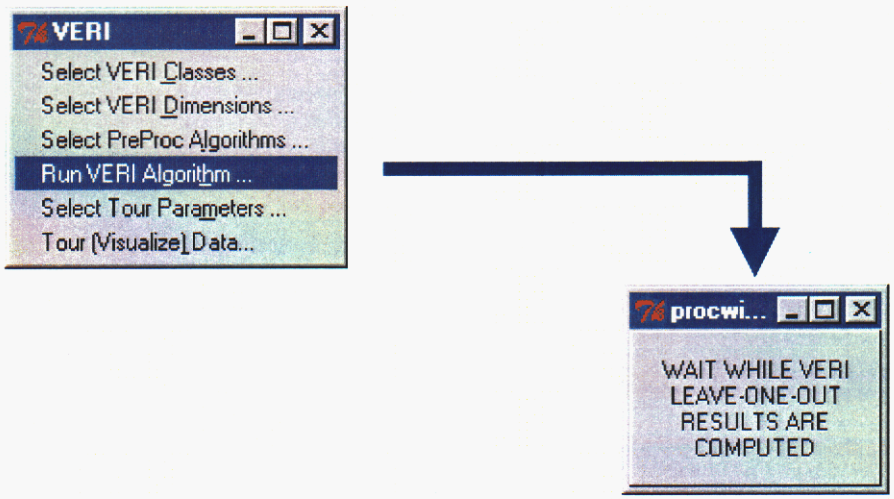

Figure 8. VERI Procwindow – displays when VERI is executing.

The VERI algorithm can be executed after a data set and all the necessary parameters have been selected. Figure 8 shows the "Procwindow" that is displayed while the VERI algorithm is executing. Files containing the analysis results are created on your hard disk in the same directory as the input data file. It may be necessary to delete the result files periodically to free up space on your disk.

The "Procwindow" disappears when the code is done executing. At that point you can go on to the next step, to Select Tour Parameters. In that window you will determine how the data looks when visualized.

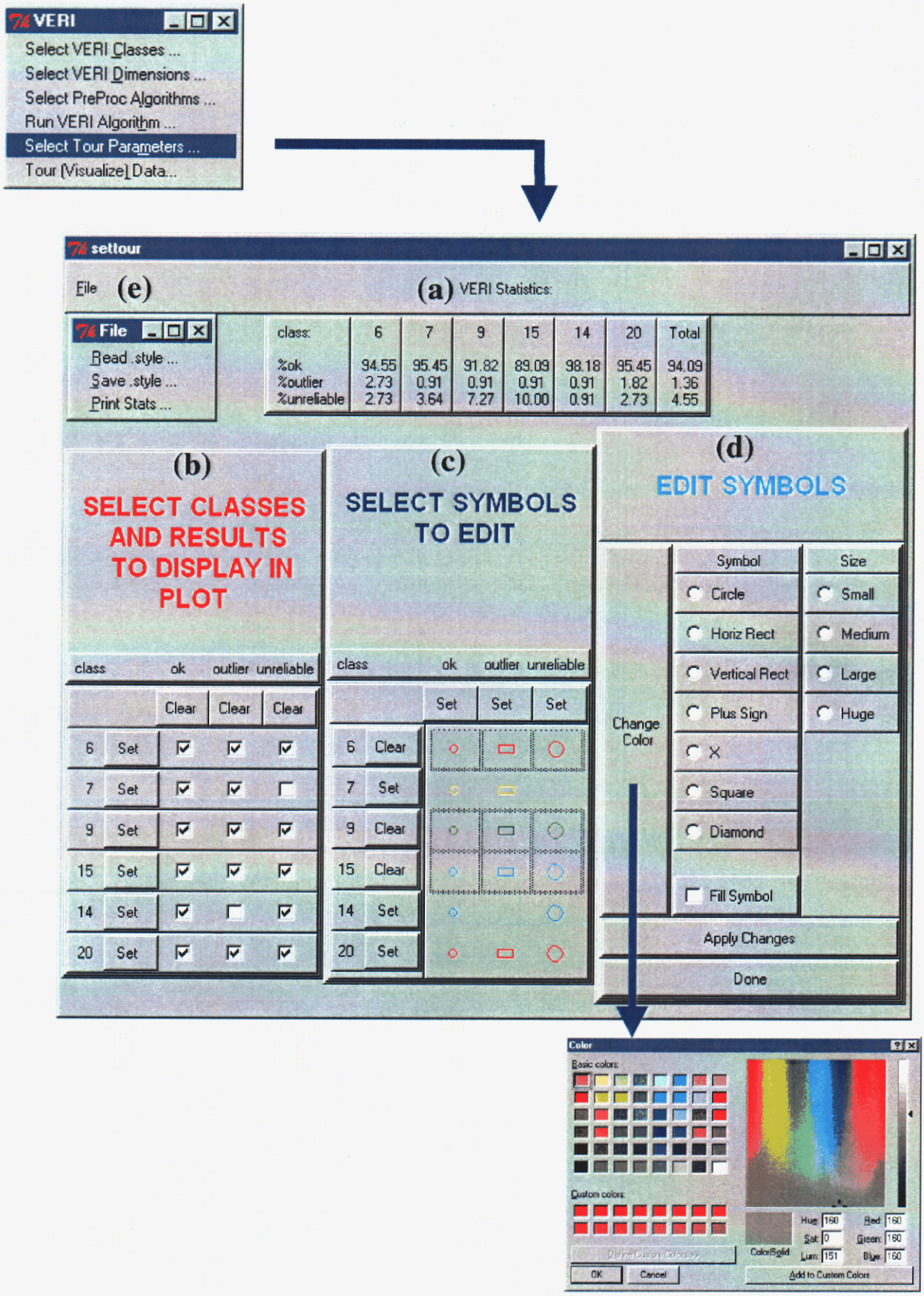

Figure 9. VERI Settour Window – VERI Statistics and Visualization Parameters

Configuring Data Visualization

After the VERI algorithm has executed, the pattern recognition statistics generated by the algorithm can be analyzed and plotting characteristics for data visualization can be selected. Figure 9 shows the Settour window where these features are accessible. This is a window with many options available. For clarity, important areas of the figure have been labeled, and each is described in detail in the following text.

VERI Statistics

Figure 9(a) is the Settour VERI Statistics Window where the pattern recognition statistics generated by the VERI leave-one-out algorithm are displayed. The VERI statistics are displayed for each class of data the user selected to investigate. The statistics generated are %ok, %outlier, and %unreliable. There is also a Total column. The values in the Total column consist of the averaged statistics generated from the evaluated classes.

If you currently or in the future plan to use the VERI Interface and are not familiar with the meaning of these statistical values, please visit our website to learn more about these values. The definition of these pattern recognition statistics and other valuable pattern recognition information is in the VERI User's Guide at our website: http://www.sandia.gov/imrl/XVisionScience/Xusers.htm**. We strongly encourage colleagues who are interested in the VERI algorithms to read the information at our site. A knowledgeable VERI user is a happy user.**

Class Display Control Panel

Figure 9(b) is the Settour Class Display Control Panel. This is where a user selects which classes and which type of classified results from those selected classes to display in the visualization. The user selects which type of classified results to view (ok, outlier, or unreliable) for each class of data evaluated. The selections can be made one (class, classification result) pair at a time. Figure 9(b) shows that (class 7, unreliable) and (class 14, outlier) are not selected to display. Selecting individual class/result pairs is accomplished by pointing the mouse cursor to the button corresponding to the pair of interest and clicking on it with the left mouse button. Rather than selecting each (class, classification result) pair individually, a Set/Clear button exists to control the display of all classified results from each class. The button next to each row affects only that row of settings. There are also buttons to Set/Clear columns, i.e., to Set/Clear all classes for each type of classified result. For example, selecting the Set/Clear button for outlier-classified results will toggle the outlier setting for all classes.

There are many interesting views one may want to observe while visualizing data. This panel offers the user total control of what data they will plot. Sometimes it is interesting to view all the data points that occur with the same classification. For example, you may want to see what the ok-classified results look like for all classes. Other times you may select to view all classification results of a subset of the classes, and of course it is interesting to view all classes with all classification results displayed.

Symbol Editing Selection Panel

Once the visualization data has been selected the user can edit the default symbols set to display in the animation. Figure 9(c) is the Settour Symbol Editing Selection Panel. Editing is not done in this panel, only the selection of which symbols to edit. Symbols are selected in this panel similarly to how (class, classification result) pairs are selected in the Settour Class Display Control Panel described in figure 9(b). Rows or columns can be selected with the Set/Clear buttons, or individual symbols can be set/cleared by placing the mouse cursor over a symbol and clicking on the left mouse button. When a symbol is selected, a gray box encloses it. Figure 9(c) shows all the plotting symbols for classes 6, 9, and 15 are selected for editing. Data that was selected for display in the Class Display Control Panel (Figure 9(b)) will display plotting symbols. Only displayed symbols can be edited.

Settour Edit Symbols Control Panel

The symbols that are selected for editing are modified using the Settour Edit Symbols Control Panel shown in figure 9(d). Settings made in this panel affect all symbols selected in the Symbol Editing Selection Panel. In figure 9(d) the Change Color button has been pressed, and the color selection window has popped up. The user can select from an array of standard colors, or can select a custom color. After a color has been selected, the color selection window must be closed. The color of the text "Change Color" changes to the selected color. Other symbol properties that can be changed include the symbol shape, whether or not a symbol is filled with color, and the size of the symbols when plotted. Changes are not made to the selected symbols until the "Apply Changes" button is selected. The "Done" button closes the Settour Window.

The Settour Window can remain open after all initial configurations have been made. This is useful when the data is being visualized. Additional changes can be made to affect the animation in real time. For example, the user can alter which classes are animated or can modify symbol characteristics. This is a "power-user" feature. However, this is one thing to keep in mind when operating in this mode. The VERI statistics in figure 9(a) are valid for the last time the algorithm was run before the Settour Window was open. If you run the VERI algorithm more than once, you must close the Settour Window and reopen it to refresh the statistics.

File Menu

Figure 9(e) is the Settour File Menu. In this menu are the options to read style files, write style files, and print VERI statistics to hard copy. A style file contains information on the characteristics of symbols used for data visualization. Plotting information such as symbol shape, fill, size and color are stored in these files. A style file can be generated when the user has set up plotting characteristics (described in figure 9(b)) that they would like to use again in another data visualization at a later time. A style file is created by selecting "Save .style" in the File pull down menu and entering a file name in the dialog box. Similarly, a style file can be loaded into the interface by selecting "Read .style" in the File pull down menu. When the dialog box opens, the user can browse and load a previously created style file. After the selected style file has been loaded, the plotting characteristics of the classes of data that are currently selected for display will be set from the values save in the file. If the style file has plotting symbol information for C classes, it will load the information into at most C classes for the new data being examined. It is useful to create a few style files and have them ready for use. The user can create a useful style file after a data file with a large number of classes has been configured for visualization. That same style file can be used to set up animation parameters for the same data set at later time, or it can be used to configure data sets with a smaller number of classes.

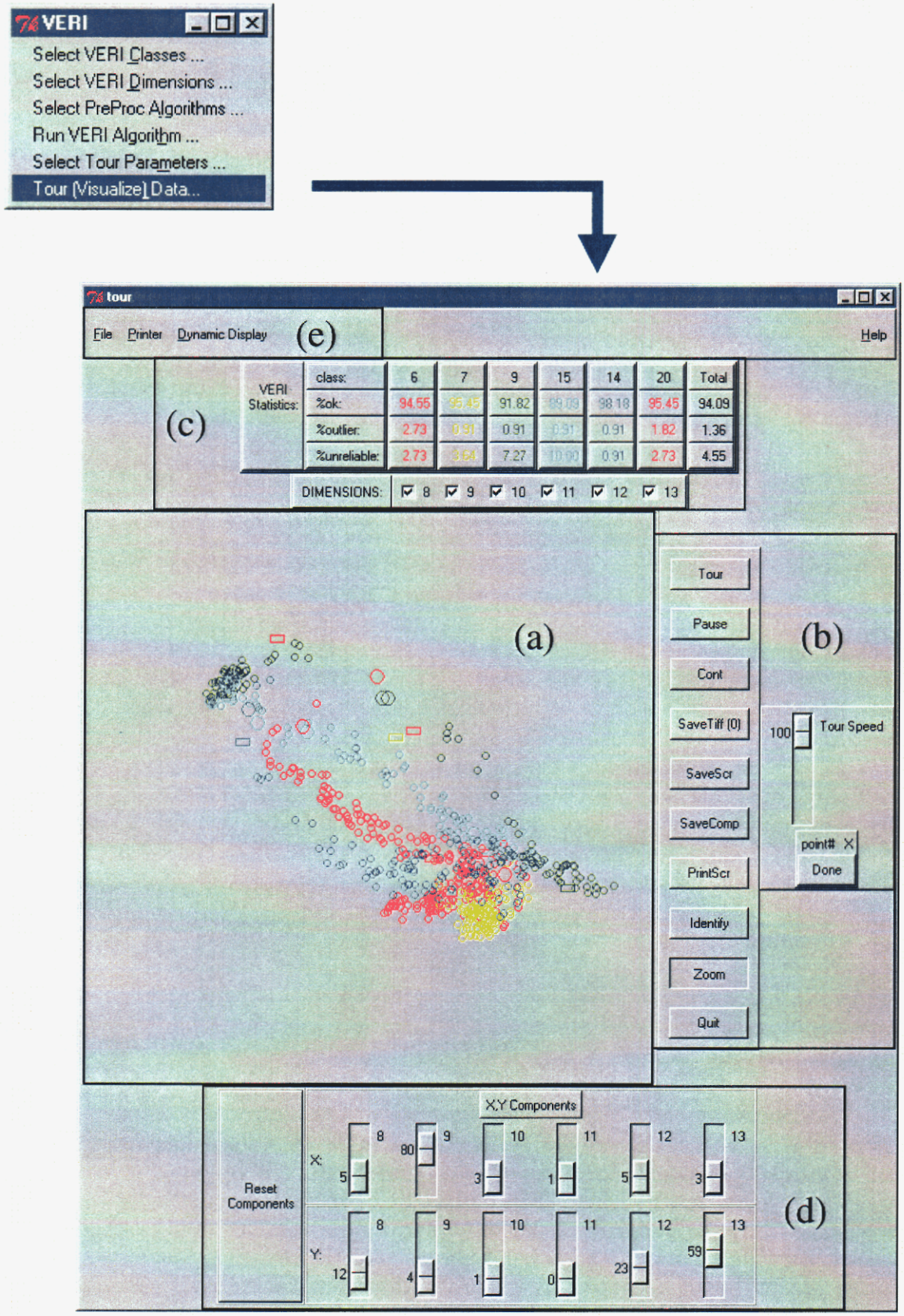

Figure 10. VERI Tour Window – Visualize Multi-Dimensional Data

Visualization

Data visualization is an important feature provided by this version of the VERI Interface. The Tour Window shown in Figure 10 allows a user to create animated mappings of multi-dimensional data. This data animation feature spatially displays the classification results in a series of random visual perspectives allowing the user to visually analyze the multi-dimensional data from many different views. The classification results and symbols shown in the animations are based on selections the user made in the Settour Window shown in figure 9. For more than two dimensions the animation is an approximation, the displayed points are derived from the users multi-dimensional data. The interface computes a random series of visual perspectives with which to display the animation. There is a trajectory or path between each pair of perspectives. The animation or Grand Tour of the data consists of allowing the interface to display the data through the trajectories of a series of visual perspectives. Analyzing the data visually is in agreement with the results of the VERI algorithm. Since the algorithm was developed from research done on human vision pattern recognition[2], it is fairly intuitive to comprehend the animated classification results.

In the Tour Window the user can visually inspect the VERI classification results, they can see how points classified and why. A user can use a visual tour to gain insight into spatial properties of data. Features such as the degree of class separation or overlap can be observed. Classes that spatially overlap can indicate the sensors used to acquire the data do not distinguish well between sampled inputs. Class dispersion, class shapes, and class densities are also important features to identify and analyze in data. These features can help determine whether more data is needed to represent a class or if the data dimensions used are suitable for representation of the classes.

Figure 10 shows a snapshot of the Tour Window while data is being visualized. Like the Settour Window, the Tour Window is full of many capabilities. It has been labeled and each section will be described in turn.

The Animation Canvas

The area in Figure 10(a) is the animation canvas. It is here where the VERI results are dynamically plotted for visual inspection. The animation displays the data from different visual perspectives. The classes of data are plotted and labeled according to the specifications made in the Settour Window. It is important to remember that the classification results shown in this animation were generated using all the classes of data the user specified for analysis. If not all the classes are shown in the animation, the results may be difficult to interpret. With that in mind, it is still useful to examine subsets of the classes of data. Sometimes a better view of data can be acquired when not all classes are displayed. This figure shows a snapshot of an animation with points from six classes of six dimensional data. There are points that classified ok (small circles), outlier (rectangles), and unreliable (large circles). The red and yellow classes of points (classes 6 and 7) fill small volumes while the other classes (classes 9,15,14,20) occupy stringy volumes.

Plot Control

The buttons in Figure 10(b) control or refer to the data plotted on the animation canvas. All these buttons are selected by moving the mouse cursor over the button and pressing the left mouse button.

The "Tour" button is pressed to initiate the animation or Grand Tour on the canvas. In a tour, the data is animated through a series of display trajectories. The goal is to display the data in a wide range of views allowing the user the opportunity to look for features or characteristics in the classification-coded plot. Earlier in describing the Settour Window mention was made of how changes could be made to the plotted classes and plotting symbols in real time. Those changes do not occur until the displayed data comes to the end of a trajectory.

The "Pause" button pauses the tour and displays the last calculated trajectory position of the data symbols. When the tour is paused, a set of scale widgets appears at the bottom of the Tour Window. These are described in Figure 10(d). The "Cont" button continues the visualization tour.

There is a feature available for users who want to save images of the tour projections to a hard drive or other external storage device. Selecting the "Save Tiff" button toggles the ability to save tiff image files to a drive. When the Tour window is initialized, the default state of the "Save Tiff" button is "Save Tiff 0", which indicates files are not being saved. When this button is selected, it brings up a dialog box that allows the user to select a location and an initial file name for saving the tiff files. When the dialog box is closed the tour proceeds and each image projected on the animation canvas is saved to a tiff file. The button then has the label "Save Tiff 1" to indicate tiff files are being saved. The saved files are named by using the entered name and concatenating sequential frame numbers to it. A maximum of ten thousand frames can be saved, but be warned, the files occupy a lot of disk space. To stop saving tiff files, select the "Save Tiff" button again and its label will return to "Save Tiff 0". The saved tiff image files can be manipulated externally. They can be modified, added to documents, or animated with a video presentation tool.

If you do not want to save a contiguous series of frames, an option exists that allows the user to save only one screen to a file. This is nice when you have "Paused" the tour with a view you like. The "SaveScr" button allows the user to save the currently displayed plot. When the "SaveScr" button is selected, a dialog box comes up and allows the user to enter a file name. The data is saved in a postscript file.

Similar to the "SaveScr" button is the "SaveComp" button. The coordinates for multi-dimensional data are calculated using a component (percentage) of the data from each dimension. The "SaveComp" button allows the user to save the components used to calculate and plot the data currently displayed on the animation canvas. These saved component values can be reloaded to show other data with the same view. The procedure of reloading the components is described shortly.

The "PrintScr" button allows the user to print the currently plotted data to a postscript printer. This is best performed when the tour is paused.

Often when a tour has been paused the user may want to identify individual data points. For users who are intimately familiar with the data being examined, having the capability to identify data points can provide insight to their data acquisition or data manipulation techniques. For example, the user may observe that several points in a class are outliers. By identifying these points one may notice that every fifth point is an outlier. That information may help the user learn about a fault in their data acquisition station or in their data processing techniques. The "Identify" button brings up a window and displays point numbers of points under the mouse cursor. Selecting the "Done" button exits the identify mode. The points are numbered according to their position in the spread sheet data file.

The "Zoom" button is a multipurpose button. Selecting "Zoom" with the left mouse button causes the users view of the data to zoom in. Selecting "Zoom" with the middle mouse button causes the users view of the data to zoom out. Selecting "Zoom" with the right mouse button restores the view to its pre-zoomed state.

The "Tour Speed" scale widget is used to slow a tour from the default speed. If a data set is small the data points are plotted quickly and tour may move around the animation canvas faster than the user may prefer. The slider on the scale slows the tour speed.

When all visualization work has been done, selecting the "Quit" button will exit the Tour Window and return the user to the Main Interface Window.

VERI Statistics and Dimensions

Figure 10 (c) shows a color-coded view of the VERI statistics computed for the data being visualized. The statistics of each class are displayed with the same color-coding used on the classes plotted in the animation. Statistics of classes that were used in the VERI analysis but whose points are not displayed in the tour are shown in a gray background. In this illustration all the analyzed classes are displayed in the tour.

A set of check boxes is displayed to remind the user of which dimensions of data are being displayed in the tour. These are not modifiable, to change these could seriously impact what a user sees in a visualization tour. With dimensions different from those used in the VERI analysis, the VERI results would not represent any view of the data shown in animation.

X,Y Components

Sometimes it is helpful to tweak a paused view of a data set displayed on the animation canvas. The scale widgets in Figure 10(d) allow a user to adjust what component each dimension of a data point vector contributes to the calculation of the plotting coordinates for points in the plot. Selecting the "Reset Components" button restores the original component values as well as the displayed plot.

Pull Down Menus -- Extra Operators

There are three pull-down menus available on the Tour tool bar. The titles are shown in Figure 10(e). These features available under these menus are less used but important nonetheless. To avoid clutter on the Tour Window, each pull-down menu is shown and described below.

Figure 10(e.1). Tour Window ReadComp Menu

Figure 10(e.1) shows the options available under the "File" menu. This menu has the "ReadComp" sub-menu. It allows the user to read in a set of stored vector components. This operation reads in files that are stored with the "SaveComp" button. When "ReadComp" is selected, a Windows dialog box allows the user to enter a components file name from which to read data. This is useful when a tour has been paused. A set of saved components can be read and used to display a tour with the same visual perspective as was used in a previous tour. This plot can be printed and compared to another printed plot with the same components.

Another way of using this feature is to run two copies of the VERI Interface with different data sets. The user can select "ReadComp" from each of the interfaces, read in the same components file, and display different data sets with the same view. Once the same view has been established, it is possible to set the components scale widgets (figure 10(d)) at similar values in the two interfaces and continue to visualize the data sets with similar perspectives.

Figure 10(e.2). Tour Window Printer Menu

Figure 10(e.2) shows the "Printer" pull-down menu. This menu allows the user the options to "Select Printer" and "Printer Setup". Selecting these options brings up Windows dialog boxes associated with printer selection and configuration. These settings determine where a hardcopy print of a paused tour is printed when the "PrintScr" button is selected.

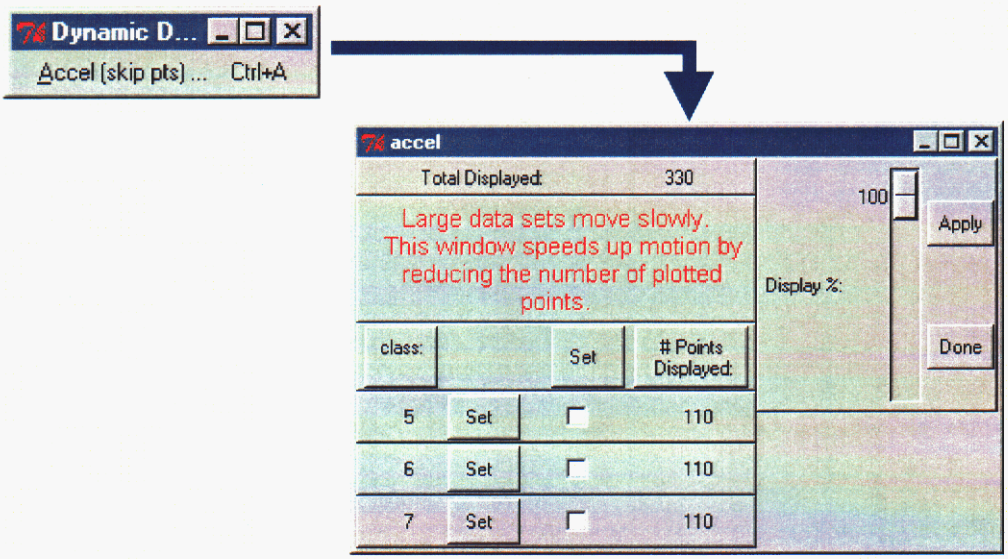

Figure 10(e.3) Accel Window – Select subset of points to display in tour.

The motion we perceive in an animation can slow down to a speed that makes it difficult to view when data sets with a large number of points and dimensions are visualized. With the advances in fast CPU's and video boards this is not the hindrance it once was, but a method was included to improve the speed of a slow animation. The "Dynamic Display" menu allows the user the option to speed up the animation. From the "Accel" sub-menu the user can opt to display a percentage of the data points from the classes being examined. The classes are selected in a manner similar to how classes are selected in the Settour Window. All classes can be affected with the Set/Clear button, or each class can be selected individually. Adjusting the scale widget changes the number of points displayed in each of the selected classes. Different percentages of reduction can be applied to different classes of data. The "Apply" button applies the currently selected reduction to the currently selected classes. Each class shows how many points will be

displayed. The subset of points shown by each class is randomly selected during the animation, allowing all points from a class equal opportunity to be displayed sometime during the tour.

The Accel window can be operated while the animated tour is occurring, the tour does not have to be paused. The number of data points shown by each class can be adjusted until the user is comfortable with the rendering speed of the animation.

A "Total Displayed" value is shown at the top of the window. When the user is finished reducing the number of points to display, selecting "Done" will close the "Accel" window and return the user to the Tour window.

Help Menu

There is a "Help" menu available on the far right hand side of the Tour Window menu bar. The information available in the Help menu should now be replaced with the description of the Tour Window found in this document. Future versions of this interface will provide links to this document and a FAQ list at our website.

Conclusions

This document is an instructional guide for a new pattern recognition data analysis tool. It describes an interface package developed to control the Visual Empirical Region of Influence (VERI) algorithm operating in Leave-One-Out mode. It also describes a powerful method provided for visualizing the classification results generated by the algorithm. The interface package was developed using the Tcl/Tk GUI programming language version 8.1. Screen shots of the interface's operations are used to describe its modes of use. Examples and advice on how to use this tool to analyze data are presented throughout the text.

A copy of the VERI Interface package can be obtained without charge by contacting the author at rfmarti@sandia.gov.

References

1 Tcl and the Tk Toolkit, John K. Ousterhout, Addison-Wesley, 1994

2 G.C. Osbourn and R.F. Martinez, EMPIRICALLY DEFINED REGIONS OF INFLUENCE FOR CLUSTERING ANALYSIS, Pattern Recognition, Vol. 28, No. 11, 1793-1806(1995)

Distribution:

	MS	Name, Org.
15	1423	R. F. Martinez, 1118
1	1423	G. C. Osbourn, 1118
1	1423	J.W. Bartholomew, 1118
1	1423	G. N. Hays, 1118
1	1425	R. C. Hughes, 1744
1	1425	W.G. Yelton, 1743
1	1425	C. E. Davis, 1744
1	9018	Central Tech Files, 8945-1
5	0899	Technical Library, 9615
2	0612	Review and Approval Desk, 9612