Subject: plotting and large data files
Posted by David Ritscher on Thu, 29 Oct 1998 08:00:00 GMT
View Forum Message <> Reply to Message

How does one handle plotting of data sets with many more points than one can plot on the screen or print on a laser printer?

I often work with large data sets, and am hitting up against some limitations in IDL and PV-Wave's plotting routines. In a typical plot I wish to plot 60,000 points per line. Obviously, this is more than a computer screen can render. There is the plotting keyword Nsum (corresponding to !P.Nsum), which averages several points together for each plotted point. If the raw data is very smooth,this would be useful. Unfortunately, I'm looking at things like ECG signals, where there are 100 or so little peaks to be seen if I plot the raw data, which are lost if I use the Nsum keyword.

The solution for display of data where one can't afford to lose any detail is a bit awkward:
1. calculate the exact plotting area, in pixels, of the current graphics window
2. for each horizontal pixel (column), figure out which of the raw data points will map into this pixel (column)
3. take the min and max of this range of the raw data points
4. for eaching pixel column on the screen, plot both a min and a max point

This should (and seems to) produce a result identical to plotting all the raw points. Information is never lost with this approach (including noise information). Steps 2 and 3 were done in C for speed.

Finally, now, my questions:

1. Is there any way to do something similar with a laser printer? Different laser printers have different mappings of pixels, sometimes even different mappings in the two directions (for example, higher resolution in the long axis of the paper). How would I perform the above steps 1. and 2. with a laser printer?

This would be very useful, since I like to print one half hour of a patients data per page. Although I'm only plotting one channel, this still results in PostScript file sizes of about 10 MB. This tends to choke up resources.

2. Would other users find this capability a useful thing to add to IDL and PV-Wave basic functionality? As an alternative to the Nsum keyword, the keyword control I would want would be to have the raw data automatically downsamples the correct amount to correspond to the actual pixels, and with a capability of controlling how this

downsampling is done:
  * averaging
  * sampling (i.e., take first sample of each of the raw data sequences)
  * min and max (plot a high and low for each column, as above)
  * user-defined function

David Ritscher
--
Cardiac Rhythm Management Laboratory
Department of Medicine
University of Alabama at Birmingham
B168 Volker Hall - 1670 University Boulevard
Birmingham AL 35294-0019
Tel: (205) 975-2122 Fax: (205) 975-4720
Email: david.ritscher@bigfoot.com