

BASIC PATTERN RECOGNITION AND DIGITAL IMAGE PROCESSING USING SAS/AF FRAME

Abhishek Lall
Department of Mathematics and Statistics,
Sam Houston State University,
Huntsville, Texas

Abstract

The principal goal of pattern recognition is supervised or unsupervised classification. The statistical approach to pattern recognition has been most intensively studied and used in practice. SAS provides a host of procedures and packages that one can use to implement basic pattern recognition steps like discriminant analysis, principal component analysis, clustering etc.

This paper presents a GUI built using SAS/AF software that uses all these procedures and provides the user a convenient way to perform these basic tasks. In addition, this it incorporates a fully working model of a decision making software component that has the capability of differentiating a good shrimp from a bad one on the basis of their digital images. For spatial data, this paper provides a convenient GUI to the user, to allow a mapping of

to map data values containing spatial locations .

Introduction

At the heart of this software is the main screen (Figure 1, Appendix) which provides a “point and click” capability to perform some powerful pattern recognition tasks. The first option allows one to extract the RGB values of an image file (in any popular format). It triggers a software module called ImageDig, that accepts any image file and returns the RGB value of each and every pixel in the image ,example (x,y,R,G,B). After extracting the RGB values ,one can perform various pattern recognition tasks mentioned above. This software allows one to perform discriminant analysis, clustering and color distribution analysis. Each option can be chosen by clicking the corresponding radio button which in turn triggers the specific screen.

All screens have an underlying screen control language (SCL) that uses basically use PROC IML, PROC DISCRIM, PROC CLUSTER , PROC UNIVARIATE and PROC FREQ to do the analytical work and PROC GPLOT and PROC G3D to demonstrate the results graphically.

One of the most power tool is the GUI clustering screen (which uses PROC CLUSTER), it provides a user an easy and friendly GUI to perform clustering .

It also hosts a very convenient menu option that triggers a screen which provides a point and click capability for mapping data values to maps. SAS/GIS is used to accomplish the goal.

Example

The best way to explain how this software works is to consider the example module included here in for culling shrimp on the basis of their digital images. The main screen (Figure 2) for this example consists of three options :1) CULLING DEVICE (which triggers a powerpoint presentation on the “shrimp culling ” project) and 2) ANALYSIS (which initiates a screen with options pertaining to some specific pattern recognition tasks associated with shrimp.

When one chooses the analysis option, it pops up a new screen

(Figure 3) which allows the user to either enter the file with RGB values for each pixel of the image or to trigger the digitizer software (imageDIG). This digitizer software is not part of SAS, rather it is an independent software that accepts any image file and returns the RGB value of each and every pixel in the image as mentioned earlier.

The RGB values of the pixels from a digitized image are comprised of vectors of length 3. These comprise complex patterns that can be easily manipulated to retrieve critical patterns using SAS PROCs DISCRIM, CLUSTER, PRINCOMP G3D etc. in attaining key information.

The subsequent screen (Figure 4) consists of a list of additional options. One option allows one to filter the shrimp's image from the background (Figure 5) .The second option allows one to find the length of the shrimp (Figure 6) . The area proportion option allows one to calculate the proportion(Figure 7) of the overall image occupied by the shrimp for use as an additional characterizing variable.

The decisions option (Figure 8) allows one to specify a threshold for length or area proportion and then permits comparison those with shrimp's image.

Mapping data

This convenient menu option allows one to map data values

to a map (Figure 10). It uses SAS/GIS to accomplish a number of tasks, along with PROC GMAP.

Conclusions

This software has a unique combination of basic pattern recognition routines accessible with the click of a mouse button. It uses the SAS/AF software's capabilities to provide a very convenient GUI.

Our future work involves implementing more advanced pattern recognition algorithms in the software.

References

- a) Applied Multivariate Statistical Analysis by Johnson and Wichern
- b) www.sas.com

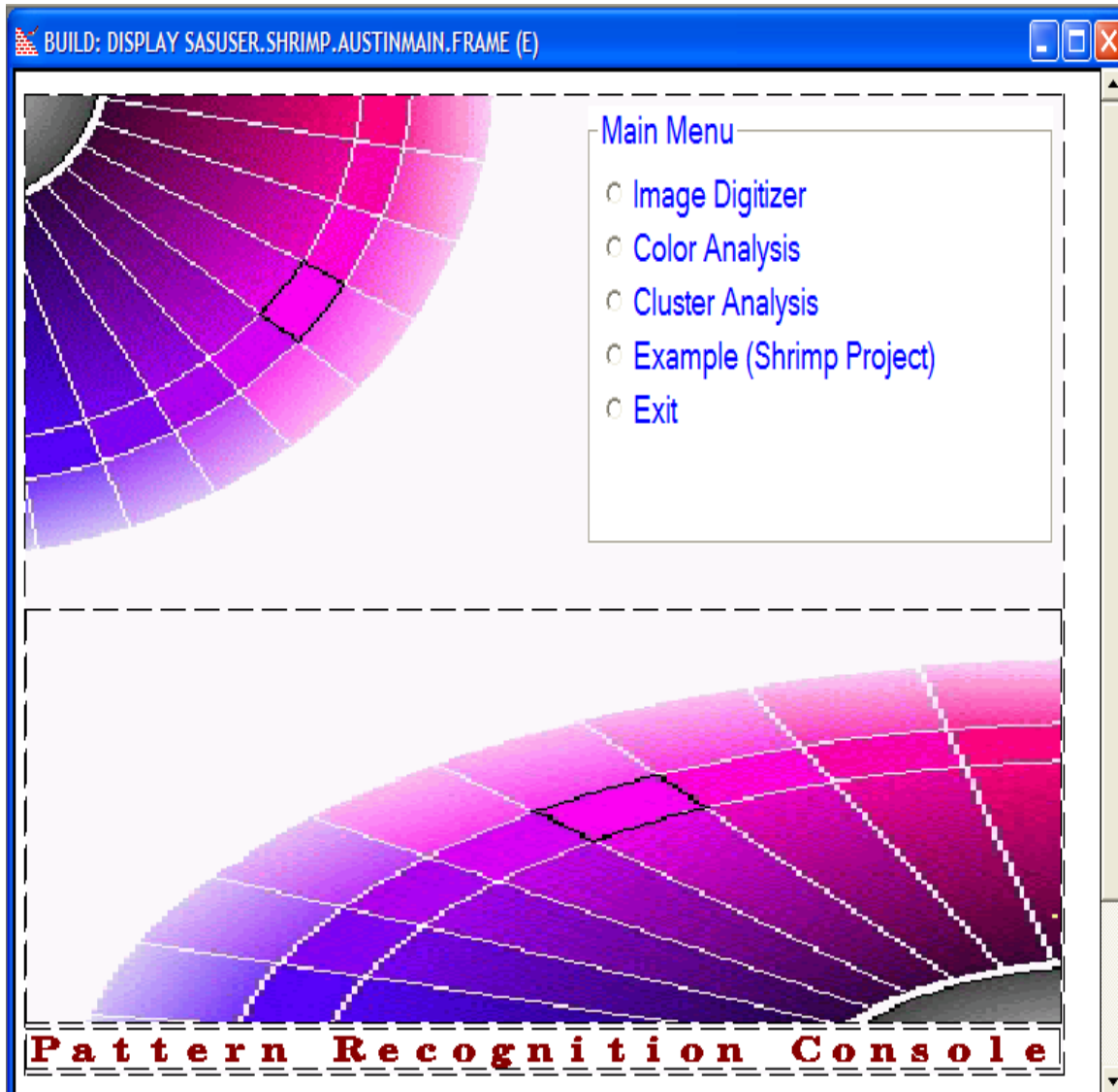


Figure 1 Main Screen



Figure 2 the Shrimp Screen

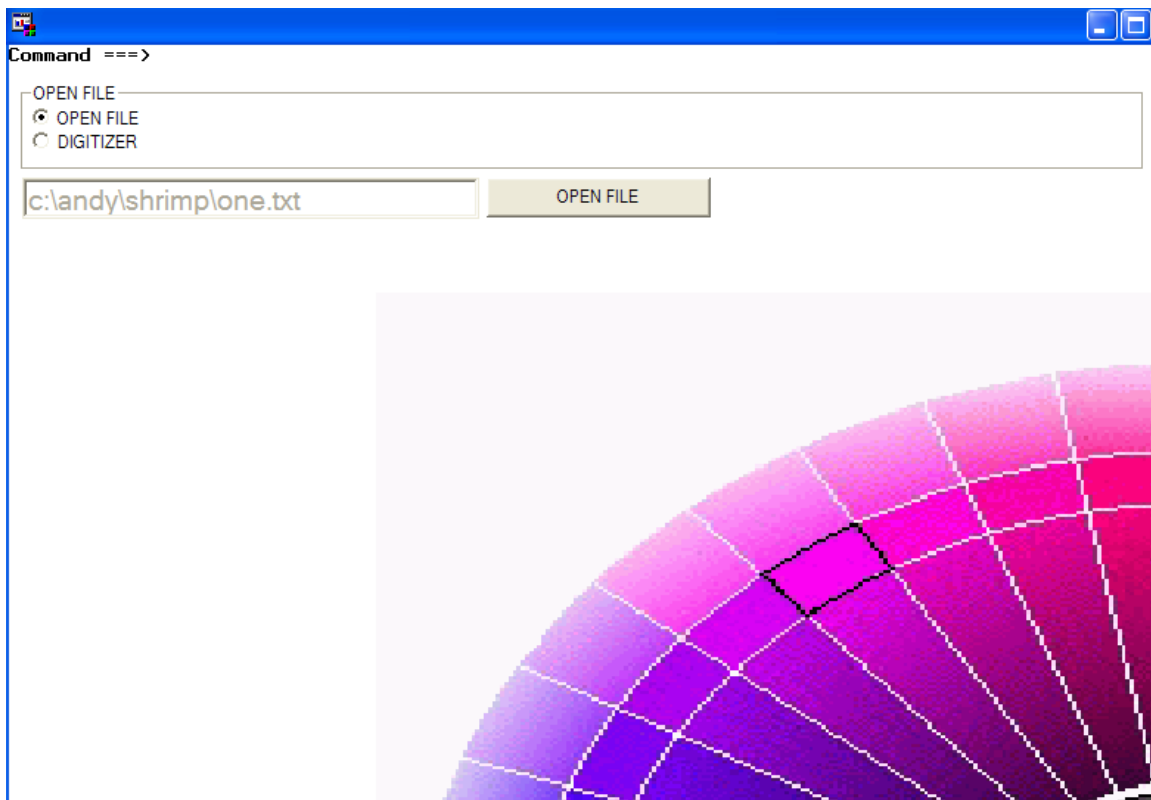


Figure 3 Screen allowing to instigate digitizer software & open the file.

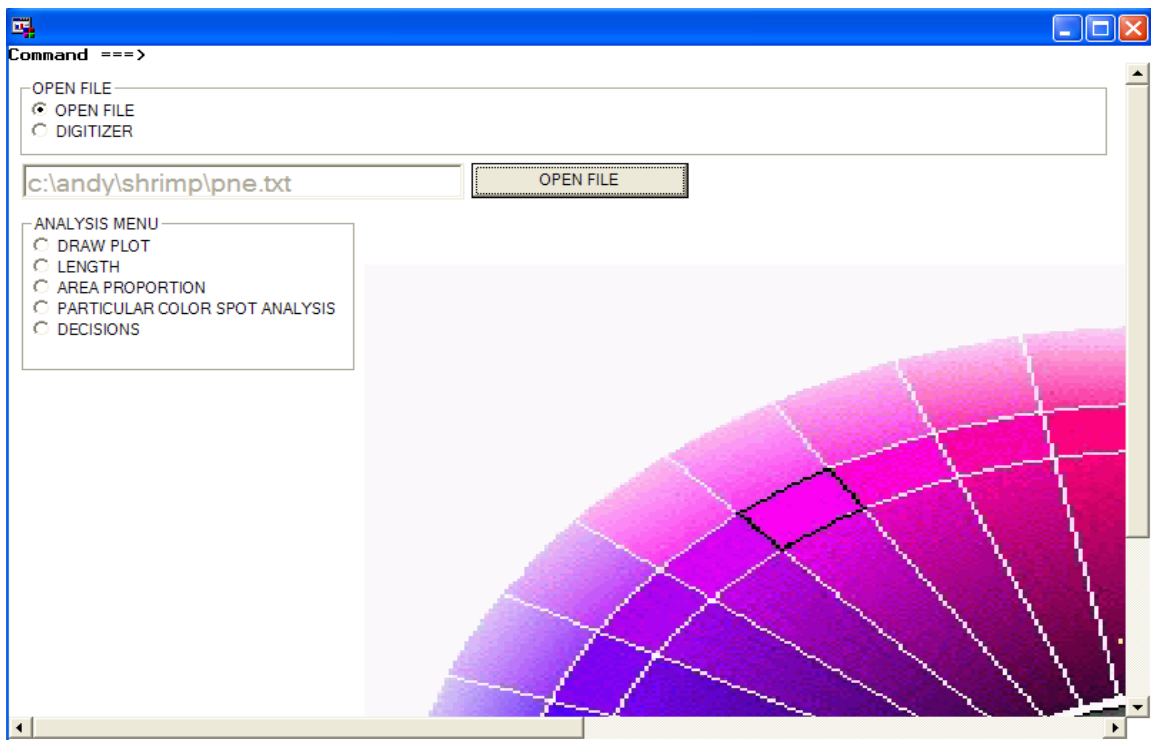


Figure 4 Screen with set of pattern recognition tasks specific to shrimp

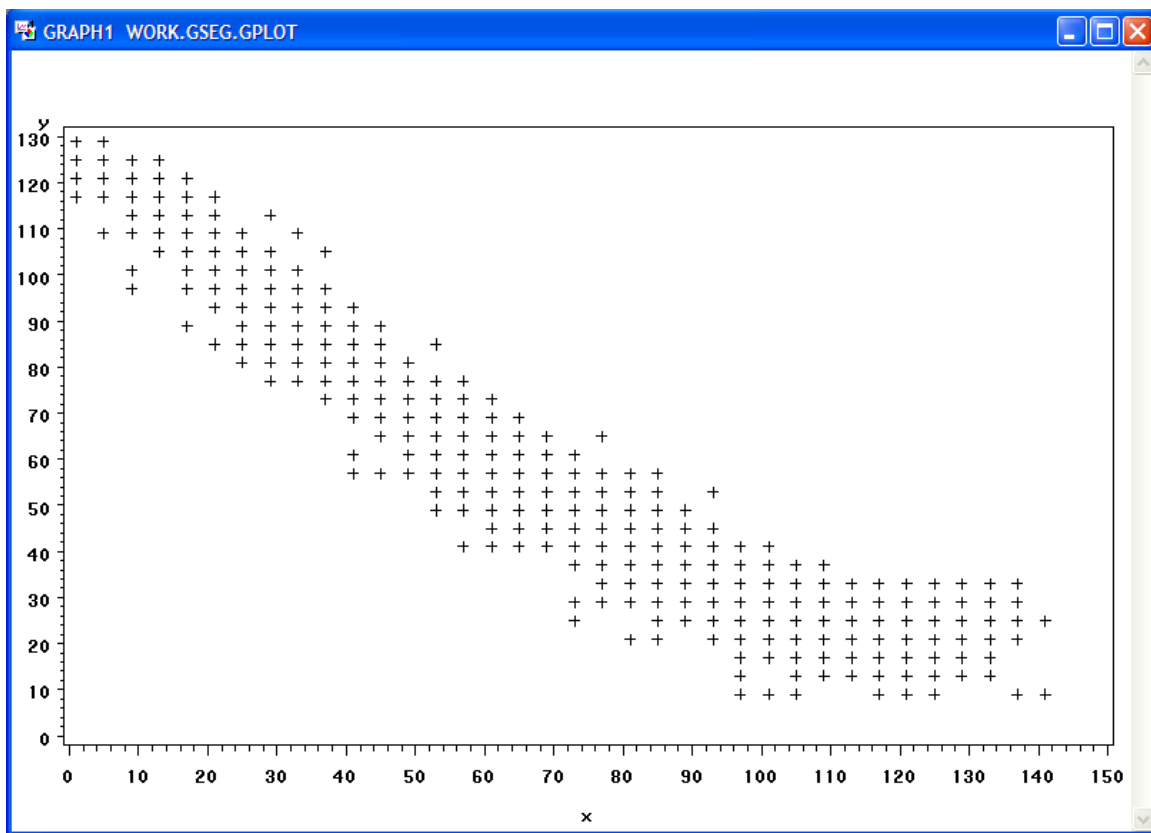


Figure 5 Filtering out shrimps image against the background

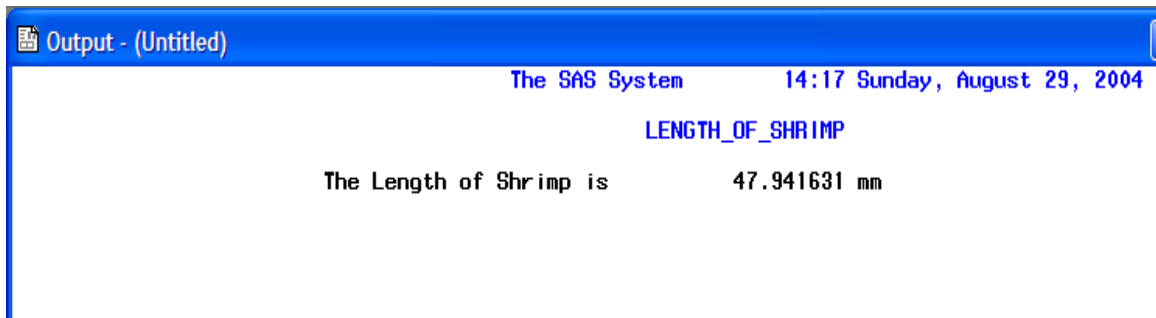


Figure 6 Calculating length of Shrimp

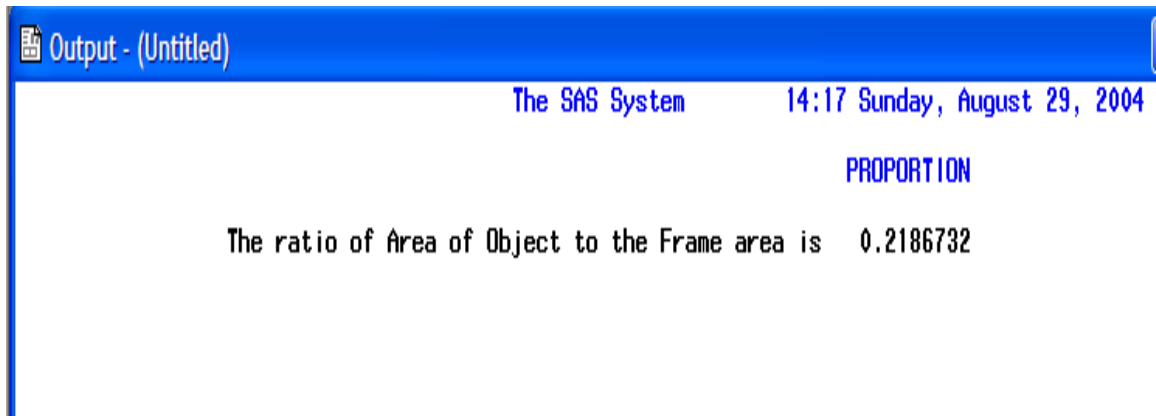


Figure 7 Calculating the area proportion

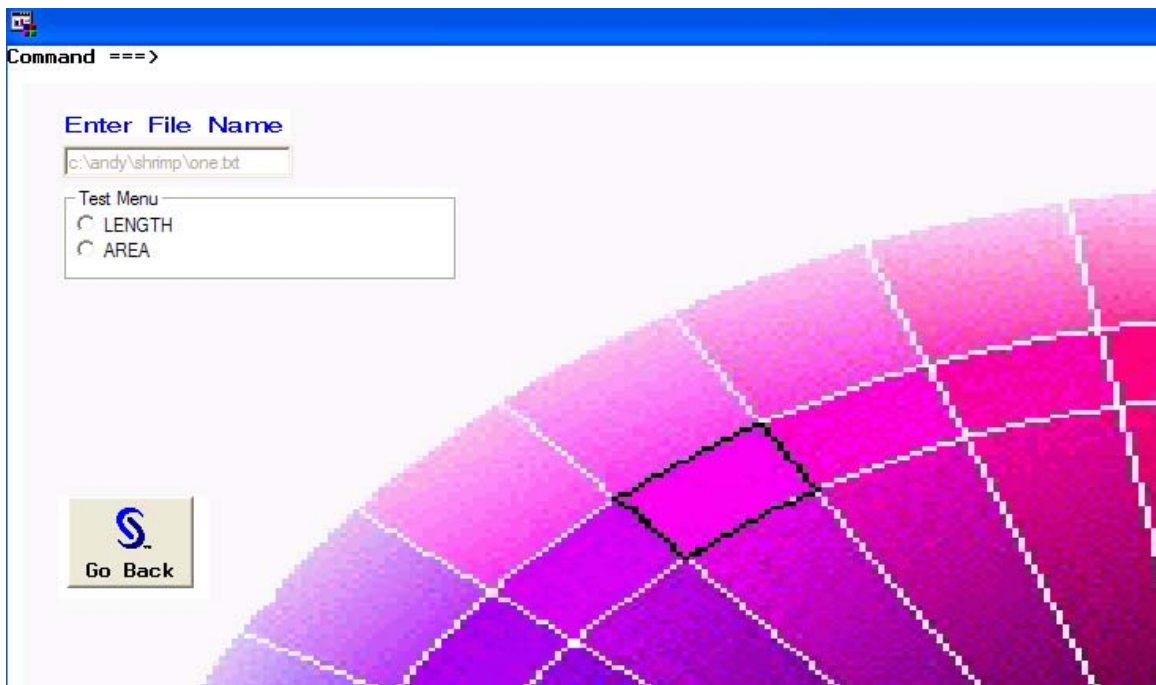


Figure 8 Decisions screen

Output - (Untitled)

The SAS System

14:17 Sunday, August 29, 2004

LENGTH_OF_SHRIMP

The Length of Shrimp is 47.941631 mm

The given shrimp is shorter than the specified measurement

Results