PROJECT REPORT
Pulkit Khandelwal
Mitacs Globalink Research Intern 2015

The project consists of two components:
1) Improvement of the existing SAS web application and addition of new features to the same
2) Analysis of Data to measure accuracy and reproducibility measures
Area of interest: Computer Vision, Digital Image processing, Statistics and Human Computer Interaction
Major tools and languages used: web2py, Matlab and Python

PART1

To locate my code, look for the name "PK" in the comments in default.py and for Views see the corresponding .html file which corresponds to the function name in the default.py code. Also, see db.py for Database Administration related code. You might want to see the comments before every function in default.py
This part of the project dealt with the addition of new features to the existing SAS Web2py application. The following additions and/or improvements were made (in no particular order):

1) Thumbnails have been added.

2) A new table has been added to the Database. Each entry in this table is an image and contain the corresponding details for each segmentation made by the participant when he/she performs an experiment. The entries are as follows: image ID, image Filename, experiment ID, person ID, experiment dataset ID and the last entry stores the JSON string which consists of information such as the coordinates of the points in the strokes, stroke start and end time and other specifications of the image being segmented.

3) Each participant has an option to either log in as an existing user or create a new user profile.

4) A participant can specify the method of input (Points or Strokes) as required.

5) Some basic but important functionalities such as withdrawing from the experiment , viewing the instructions at any particular page (time) were also added.

6) A new experiment could not be created earlier. This bug has been rectified.

7) An experiment could not be carried out as we were not able to go to the next image when we hit the Enter key. This bug has been rectified.

Some other trivial bugs were fixed and functionalities were added which might be noticed in the app. Many of the bugs existed due to syntax error. They were quite trivial but a little bit trickier to find and rectify.

Note: Web2py user manual is a good source for reference. Web2py Application Development Cookbook by Packt Publishing was really helpful as it provides with ready to use recipes. Stackoverflow, google groups and other blogs were also referred to a great extent.

PART 2

The data provided consisted of ground truth and participant marked up images. The aim was to compute the *accuracy* and *reproducibility* measures by performing some statistical analysis. The participant mark up images were divided into two categories; points and strokes.

1) The points and strokes on each of the image were extracted and labelled according to the following : 1=Background and 2=Foreground. Out of a total of 30 participants, 15 used points and the remaining used strokes. The original image and the mark up image were subtracted. This difference image consisted of blue and green strokes. These colours were extracted individually by separating the blue and green colour planes. For points, after obtaining the difference image, the image had yellow and violet points. These colours were extracted by appropriate combination of the primary colours. After extracting the foreground and background strokes/points, the images were saved separately and converted to 1=F and 2=B. These were then added to get the final grayscale image.

2) These labelled images were then used to segment out the foreground from the background in the original images using the Generic *Boykov* graph cut segmentation algorithm.

3) Using *Dice Coefficient* we looked at the mean accuracy measures of the segmentation provided by the points and strokes images. Some statistical analysis was done to verify the hypothesis. These mainly included tests to check whether the dice coefficient values of the strokes and points dataset were normal or not using *Lilliefors test*. Depending upon this result, further tests were employed such as *Two-sample t-test* and *Mann-Whitney U-test*.
The overall idea was to see whether there is any significant difference between the mean accuracy measures between points and strokes.

4) The Generalized *Tanimoto Coefficient* was computed for the segmented images (Only the images which were segmented by at least three users were considered) separately for both the points and strokes segmented images. The statistical tests included the same as above viz., Lilliefors test, Two-sample t-test and Mann-Whitney U-test.
The GTC has to be calculated for each image for both the mark up-type.
The overall idea was to see whether there is any significant difference between the mean reproducibility measures between points and strokes.

Note:
The extraction of the points and strokes from the participant mark up is a very crucial step and should be done with utmost care so that the points and strokes are not missed out.

In the Data Analysis folder you will find the following folders and files:

experiment 2: Ground Truth and participant mark up images
Binary Ground Truth: Binary ground truth images
Strokes and Points: Labelled images(Foreground = 2 and Background=1) grayscale images and segmented, contours images for both Points and Strokes
Results: Code, data, statistical results and graph plots for both the Dice Coefficient and Generalized Tanimoto Coefficient

The following were the results obtained:
Note: The variables used are as mentioned in the MATLAB Help documentation.

DICE COEFFICIENT

STROKES DATA

Number of Samples = 169
Mean of the samples = 0.7740

Lilliefors Test
h =1; p = 1.0000e-03; k = 0.1418; c = 0.0688

POINTS DATA

Number of Samples = 116
Mean of the samples = 0.7369

Lilliefors Test
 h = 1; p = 1.0000e-03; k = 0.1772; c = 0.0828

STATISTICAL TESTS

Two-sample t-test
h = 0; p = 0.0953;
ci = -0.0808
       0.0065
stats =
   tstat: -1.6737
     df: 283
     sd: 0.1840

Wilcoxon rank sum test (Mann-Whitney Test)
p = 0.0334; h = 1;
stats =
    zval: -2.1272
   ranksum: 1.5134e+04

GENERALIZED TANIMOTO COEFFICIENT

STROKES DATA

Number of Samples = 14
Mean of the samples= 0.6751

Lilliefors Test
h = 0 ; p2 = 0.5000; k = 0.1441; c = 0.2257

POINTS DATA

Number of Samples = 10
Mean of the samples = 0.6816

Lilliefors Test
 h = 0; p = 0.2875; k = 0.2026; c = 0.2620

STATISTICAL TESTS

Two-sample t-test
h = 0; p = 0.9326;
ci =   -0.1498
        0.1627
stats =
    tstat: 0.0856
    df: 22
    sd: 0.1820

Wilcoxon rank sum test (Mann-Whitney Test)
p = 0.9766; h = 0;
stats =
    zval: -0.0293
   ranksum: 124