TNM098 - Advanced Visual Data Analysis

Lab2 - Analysis of spatiotemporal (eye-tracking) data

Authors:
David Tran - davtr766
Oscar Nord - oscno829

I. Introduction

The second lab of the course *TNM098 Advanced Visual Data Analysis* consisted of a data-file that contained data from an eye-tracking session lasting about five minutes. The objective was to identify regions of interest and how they changed over time as the session proceeded. The data-file included the following aliases: Time, gaze duration and position in *X* and *Y* coordinates on the computer screen.

Questions of interest:

- How many regions can be identified?
- Which regions are only used for part of the analysis procedure?
- How many are heavily used and when?
- What are the frequent transitions between the areas of interest? and how do those transition patterns change over time?

II. Approach

The first step to answer the questions above was to convert the given *tsv* file to *csv*. This lab is more suitable for a visual representation approach and therefore, *R* is used to easily plot and manipulate the data with various functions [1].

In this lab, data preprocessing was necessary in order to plot the gaze points correctly. Since every tuple represent one gaze point and have *x* and *y* as aliases, one have to manipulate the *y* variable in order to plot correctly by converting the y-coordinate with the formula:

$$y_{\mathrm{i}} = y_{\max} - y_{\mathrm{i}}$$

where *i* are from 0 to number of tuples in the data file. Now when plotting the gaze points, the y-axis will be correct.

After the preprocessing of the data, R is then used extensively to plot the gaze points based on the available aliases and DBScan [2] is also used to effectively cluster the data. The steps that was made in this lab are the following:

1) Plot all gaze points from the data file.
2) Plot all gaze points from the data file with over 800ms in gaze duration.
3) Plot heat map based on the recording time stamp.
4) Create an animation showing how the eyes moves over time.
5) *DBScan* clustering with minPts = 15 and eps = 6.

III. Result

Figure 1 illustrates all the gaze points in a plot and gives an overview of the area.

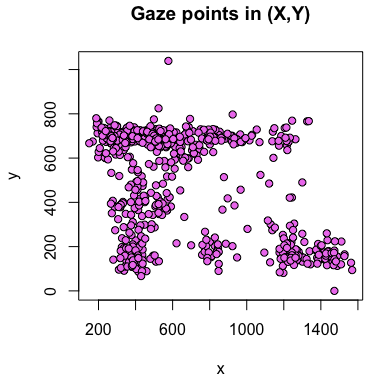

Fig. 1: A visual representation of all gaze points from the given data file.

The next plot in figure 2 shows exactly the same plot but now filters all gaze points that have less than 800ms in duration and thus shows the interesting areas that the eye seem to spend more time looking at. To complement this, a heat map plot was created and can be seen in figure 3 along with a *DBScan* clustering plot in figure 5. In order to see how the

eye moves over time, the library *googleVis* [3] is used to animate the gaze points. Illustrations of how the animation interface looks can be seen in figure 4.

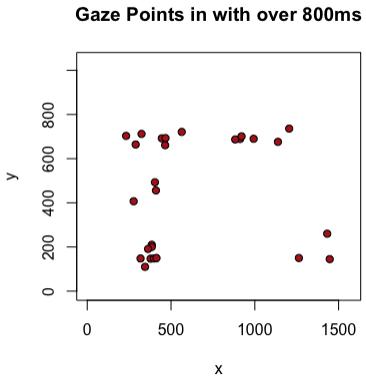

Fig. 2: A visual representation of all Gaze Points from the given data file with over 800ms in gaze event duration.

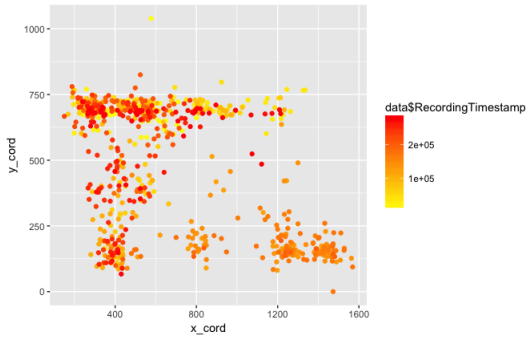

Fig. 3: A heat map of the Gaze points over time. The axis corresponds to x and y coordinates and the right-hand side illustrates the recording time stamp and the color corresponds to the respective gaze point.

IV. Discussion

When analyzing the result it clearly shows eight distinct clusters where the gaze event duration exceeds 800 milliseconds with a clear concentration to the left part of the screen. Animation in figure 4 shows that the user concentrates on the upper left part of the screen for the major part of the recorded session. The participant's focus fluctuate the most during the first half of the five minute tracking-session and the concentrates to the left.

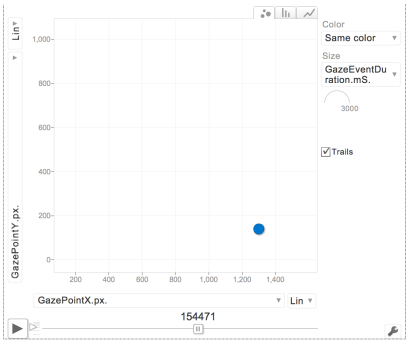

Fig. 4: Animation interface of gaze point movement over time. The axis of the window is based on the x and y position of the gaze point. The timeline is located at the right-hand side of the play button. The circle becomes larger if the event duration of the points are large.

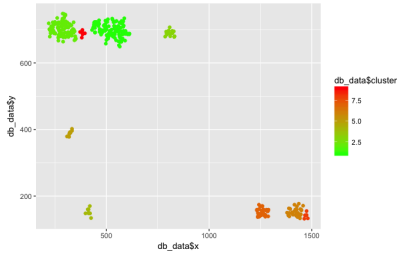

Fig. 5: DBScan clustering with minPts = 6 and eps = 15 which illustrates cluster with minimum of six points within the radius 15. The plot indicates that there are eight clusters with these parameter values.

References

[1] The R Project for Statistical Computing, Accessed: 2018-04-22 https://www.r-project.org/

[2] DBScan, Accessed: 2018-04-22 http://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.htn

[3] , Introduction to googleVis 0.6.2, Accessed: 2018-04-22 https://cran.r-project.org/web/packages/googleVis/vignettes/googleVis.pdf