# Navi Campus: Quantitative Methodology for Evaluating the User Interface of a Navigation App Using Eye Tracker and Smartphone

Jesus Zegarra Flores
Research and Innovation
Capgemini Engineering
Illkirch, France
email:jesus.zegarraflores@capgemini.com

Sabine Cornus
Faculté des Sciences du Sport
Université de Strasbourg
Strasbourg, France
email:cornus@unistra.fr

Emma Charbonnier
Research and Innovation
Capgemini Engineering
Illkirch, France
email:emma.charbonnier@capgemini.com

Laurence Rasseneur
Faculté des Sciences du Sport
Université de Strasbourg
Strasbourg, France
email:rassene@unistra.fr

***Abstract*— Navi Campus is a mobile app developed between Capgemini engineering and the Strasbourg university. Currently, navigation apps can have some limitations; for instance, the GPS coordinates of the destination do not correspond to the real destination in a university campus and they are not adapted to people having disabilities. Our version of the app, which overcomes the problems mentioned before, is in advanced stage and we want to evaluate the use of the app in a quantitative way using a mix of data (IMU, GPS and cameras) coming from the eye tracker Tobii and the app. The idea of the eye tracker is that we can detect the ocular and head movements in addition to the information from the mobile app. In the first part of the article, we will show the kinds of data we have chosen from the app and from the eye tracker to have quantitative analysis from the speed, the step frequency, and the time that a person looks at the phone. For validating our quantitative method, these firsts tests have been conducted with three people with different profiles in three different paths of similar characteristics from the Strasbourg Campus University. This methodology seems promising for analyzing the efficiency of the app establishing the relationship between the quantitative information and the behavior of the users.**

***Keywords-User activity; GPS; navigation; eye tracker; mobile app.***

## I. INTRODUCTION

In the Strasbourg university campus in France, common mobile navigation application like Google Maps are not precise enough to locate the building's entries and to provide wheelchair-accessible paths. These are the reasons why the mobile application Navi Campus was developed [1]. This application provides outdoor navigation to help freshman students and visitors to go to a specific building on the campus from inside or outside the campus. Indeed, Navi Campus provides timetables from buses and tramways around the campus. The application is now at an advanced stage of development, and this paper introduces a quantitative method to evaluate the performances of the application.

Indeed, qualitative and quantitative studies allow to evaluate the feeling and behavior of the users [2], [3] about the interface. This evaluation is an integral part of the development of an application, in order to improve the quality of the tool and its usability. For navigation applications, the interface is even more important as misunderstanding information can lead the user to get lost. From this perspective, we conduct a first quantitative study of the interactions between the user and the Navi Campus app.

This article describes the context of the experimentation in Section 2. Section 3 proposes a methodology choosing the data for analyzing the use of the application and the behavior of the user, based on eye tracking sensor and smartphone IMU. Section 4 presents the results and Section 5 presents the conclusions and perspectives.

## II. CONTEXT OF THE EXPERIMENT

### A. Environment of the study.

In Figure 1, three paths are defined across the University Campus of Strasbourg, which is a calm environment with no cars and uncrowded.

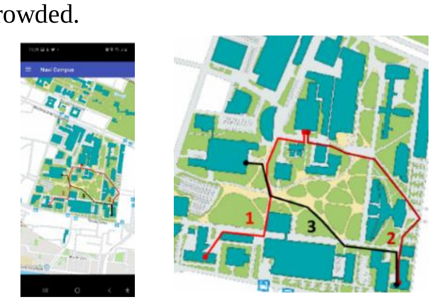


Figure 1. Three paths for the tests using Navi Campus

The first path is 400m long from "Le portique" to "SUAPS" and includes 3 sharp turns and 1 crossroad. The second path is 450m long from "SUAPS" to "Pangloss" and includes 2 sharp turn and 1 smooth turns. The third path is

440m long from “Pangloss” to “UFR Mathematic Info” and includes 3 sharp turns and 1 smooth turn.

### B. *Subjetcs and Material*

Three subjects, who use smartphones every day, are asked to use the app to go through all three paths in the same order. The indication given was just to follow the information given by the app. Subjects have different knowledge of the campus:

Subject A knows neither the campus nor the application.

Subject B knows the campus but not the application.

Subject C knows the campus and the application.

While they are walking using the phone horizontally on the hand, the app records IMU data from the phone with a sample time of 20ms and the GPS data with a sample time of 1s. The participants also wear the Tobii Pro Glasses 2 [4] eye tracker device, which includes IMU to record head motion with a sample time of 10ms. Tobii eye tracker also provides gaze motion and video of what the user sees, which can be used to determine where the user is looking on the interface. After performing a calibration task of the eye tracker device, the subjects completed the paths and the data can be analyzed. Unfortunately, gaze data available is not exploitable as the eye tracker don’t work well outside, due to high brightness.

### C. *Segmentation of the paths*

In order to sharply analyze the behavior of the users, considering the information giving by the phone, the paths are automatically segmented in turns and straight lines. To do that, the walking orientation of the subject is analyzed using the recorded GPS coordinates (after doing the path), to determine when the user has turned or not.

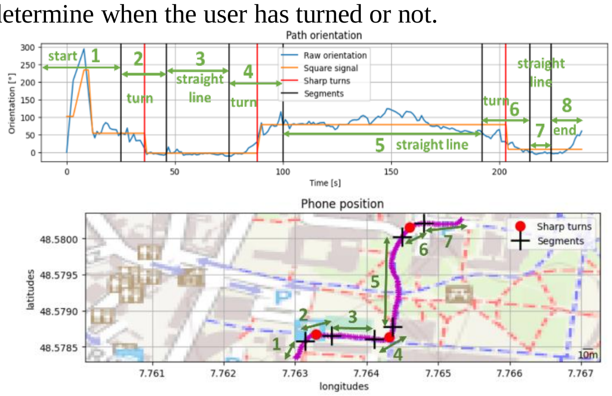


Figure 2. Example of segmented path (first path)

After approximating the orientation signal to a square signal, sharp turns can be extracted using a threshold 60°. Turn segments starts 15 meters before the turn and finishes 15m away. False turns are ignored when the user is not moving. In Figure 2, three sharp turns, 3 straight lines and start and end segments are detected (example of first path).

## III. METHODOLOGY

After synchronizing the phone and the eye tracker data in post-processing, in each path segment, the speed, the steps frequency and the time when the user looks the app are identified using algorithms with Python to compare the behavior of all subjects in straight lines and turns.

### A. *Speed analysis*

The user speed is calculated from the GPS coordinates, as the distance between two consecutives points calculated with the Haversine formula divided by the elapsed time. As the GPS is not very precise, some abnormous points can occur. Therefore, speeds higher than 8 km/h are ignored (Figure 3). The mean speed of each segment is then calculated and compared between segments.

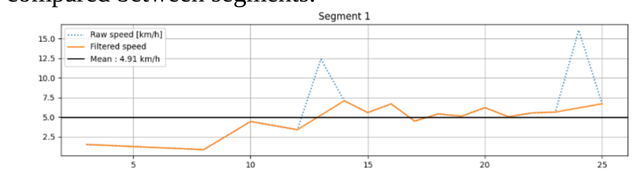


Figure 3. Example of filtered speed.

### B. *Step frequency*

If the user steps frequency is low, it can be read as the user hesitating about the path to follow or an obstacle like people riding bicycles on his/her way. During the tests, the subjects hold the phone horizontally even when they are not looking at it, that is why the detection is based on the phone acceleration rather than the eye tracker acceleration, as the head motion induce too much acceleration variations.

To identify if the user is walking at a low steps frequency, the compound acceleration is calculated.

$$a_{comp} = a_x + a_y + a_z \tag{1}$$

To reduce noise, an average filter (window of 40 samples) is applied. As the sample time is not stable, no frequency filter is employed. The average signal is used to define a variable threshold for steps detection. 1.7m/s² is an empiric value determined in indoor environnment, that gives less than 10% error. When the compound acceleration is higher than the threshold, a step is potentially detected (Figure 4 (a)). If this potential step is made of at least 4 samples, it is a valid step (Figure 4 (b)). The delay between two consecutives steps allows to determine a local frequency (Figure 4 (c)).

$$freq_{local}[i] = \frac{1}{time_{step_{i+1}} - time_{step_i}} \tag{2}$$

The mean steps frequency for each segment is used to know if the user is walking at a low frequency. When the local step frequency is lower than 0.5*mean steps frequency, a low frequency event is raised.

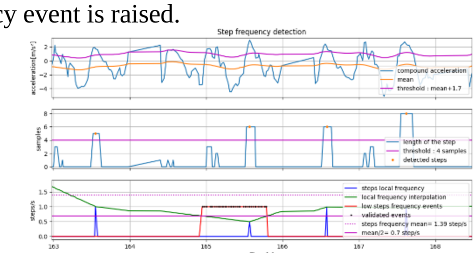


Figure 4. User low steps frequency detection . (a) Potential step detection, (b) step validation, (c) low steps frequency detection

As the accelerometer can miss some acquisitions, a confidence indicator is used to determine whether the low frequency event is due to low frequency or miss acquisions. The low frequency events due to miss acquisition are ignored.

 ISBN: 978-1-61208-916-4

### C. Looking at the phone

When the user looks at his phone, he/she puts his/her head down and the phone appears on the eye tracker video. To detect these events, two algorithms are used.

#### 1) Head motion

It is obsverved through the glasses' IMU. When the user looks down, the gyroscope senses a rotation along the lateral axis X and the accelerometer a translation along the 2 axis Y and Z (see Figure 5).

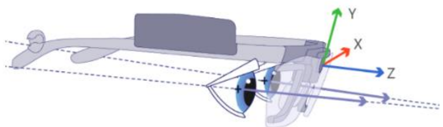


Figure 5. Glasses parameterization. [4]

To detect rotation along X, the signal is filtered with a average filter (window of 50 samples) to reduce noise. If the average variance (window of 50 samples) is higher than the variance of the whole segment, a rotation of the head is detected (see Figure 6).

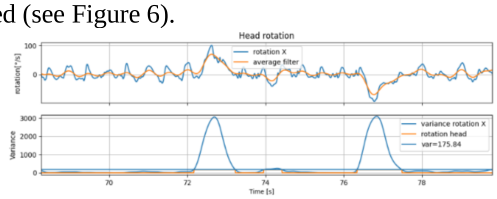


Figure 6. Detection of the head rotation

For the translation, the compound acceleration is defined as $a_{comp} = a_y + a_z$. The noise is reduced with an average filter (window of 200 samples). This mean is coupled to a square signal shown Figure 7 (top).

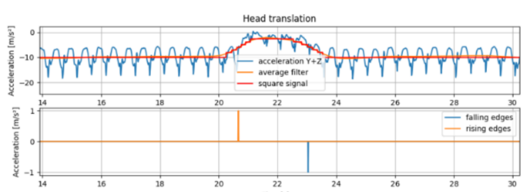


Figure 7. Detection of the head translation

When a rising edge of the square signal is followed by a falling edge, and both occur at the same time as a rotation event, an head down event is identified. For example, Figure 8 shows 11 head down detection in the first path of subject C.

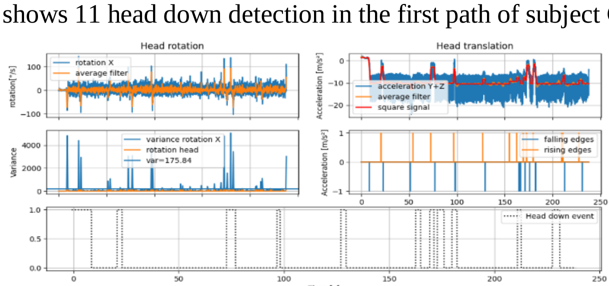


Figure 8. Detection of the head down event

#### 2) Phone detection

When the user looks down, it does not always mean that he/she is looking at the application. A validation of the phone is made using the video. Each 400 milliseconds frame of the eye tracker is analyzed by a neural network mask RCNN [5] in mode transfert learning using the dataset coco to determine whether or not the user is looking at his phone.

As the conditions of the study were sunny, the brightness disturbs the detection ; in some cases, the phone is not detected or identified as other objects (knife, tie, skateboard, snowboard). Some exemples are presented in Figure 9.

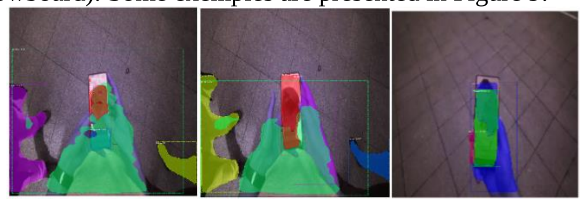

Figure 9. Exemples of identification as a cellphone (left), a tie or knife (center) and a snowboard (right)

The missidentifications can occurs quite often for some users. Indeed, as the subject A did all his paths on bright conditions, the phone is highly misstaken as a tie or a knife.

TABLE I. APPEARANCE PERCENTAGES OF EACH CLASS CONFUSED WITH THE CELL PHONE CLASS FOR ALL PATHS

| **Percentage appearance** | **Class detected** | | | | |
|---|---|---|---|---|---|
| | ***Cell phone*** | ***Tie*** | ***Knife*** | ***Snowboard*** | ***Skateboard*** |
| Subject A | 43,976 | 25,000 | 23,494 | 3,916 | 3,614 |
| Subject B | 55,631 | 17,342 | 0,450 | 18,468 | 8,108 |
| Subject C | 75,824 | 0,000 | 3,297 | 8,791 | 12,088 |

To avoid losing too much data by accepting only the cell phone class, we define a superclass phone, regrouping the class *cell phone*, *tie*, *knife*, *snowboard* and *skateboard*.

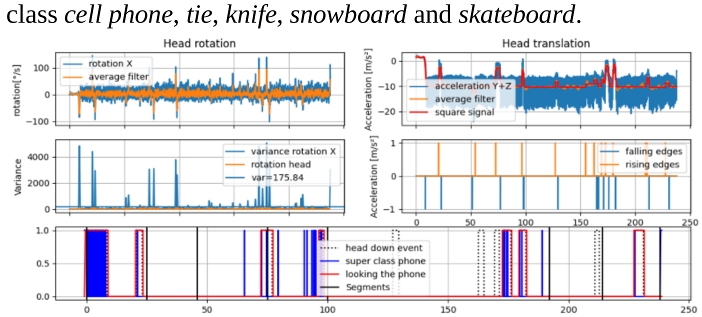


Figure 10. Detection of the user looking down to his phone

When the user looks down and an object of the super class is detected, the user is looking at his/her phone (see Figure 10).

## IV. RESULTS

As all the subjects arrive at their destinations with no major problems, the application is globally efficient. To analyze the subjects' behavior, the previous algorithms are applied to the data, and the results are summarized on table 2 for the first path. The duration of the events (low steps frequency and time looking at the phone) is formulated as percentage to the segment duration.

TABLE II. INFORMATION FOR THE FIRST PATH IN EVERY SEGMENT FROM THE THREE SUBJECTS

| Path 1 | *Subject* | *Segment duration [s]* | *Mean speed [km/h]* | *Mean step frequency [step/s]* | *Low step frequency* | | | *Looking the phone* | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | *Ratio [%]* | *Mean time duration [s]* | *Mean slow step frequency [step/s]* | *Mean time duration [s]* | *ratio [%]* |
| 1: Start | A | 29.0 | 4.89 | 0.83 | 32.24 | 9.35 | 0.06 | 6.12 | 63.36 |
| | B | 35.49 | 2.59 | 0.11 | 90.88 | 32.26 | 0.02 | 17.08 | 96.24 |
| | C | 25.0 | **4.91** | 1.04 | 37.28 | 9.32 | 0.07 | 2.36 | **9.46** |
| 2 : Turn | A | 25.0 | 4.88 | 1.12 | 0 | 0 | N.A. | 0 | **0** |
| | B | 24.0 | 5.3 | 0.63 | 30.04 | 7.21 | 0.03 | 1.91 | 15.95 |
| | C | 21.0 | **5.95** | 1.43 | 0 | 0 | N.A. | 0 | **0** |
| 3 : Straight line | A | 37.0 | 4.9 | 1.43 | 0 | 0 | N.A. | 1.96 | 10.58 |
| | B | 43.0 | 5.57 | 1.35 | 6.51 | 2.8 | 0.54 | 2.9 | 6.74 |
| | C | 29.0 | **5.6** | 1.28 | 4.14 | 1.2 | 0.04 | 2.3 | 7.92 |
| 4 : Turn | A | 32.0 | 4.08 | 1.44 | 0 | 0 | N.A. | 2.84 | 17.78 |
| | B | 27.0 | 4.92 | 1.2 | 0 | 0 | N.A. | 5.99 | 22.17 |
| | C | 25.0 | **5.1** | 1.01 | 0 | 0 | N.A. | 1.5 | 6 |
| 5 : Straight line | A | 114.0 | 4.68 | 1.36 | 0 | 0 | N.A. | 2.72 | 19.09 |
| | B | 91.0 | 5.65 | 1.34 | 1.27 | 1.16 | 0.19 | 4.81 | 15.85 |
| | C | 92 | **5.78** | 1.39 | 0.91 | 0.84 | 0.59 | 2.79 | 6.06 |
| 6 : Turn | A | 28.0 | 4.55 | 1.25 | 0 | 0 | N.A. | **2.61** | **27.91** |
| | B | 23.0 | 5.03 | 0.78 | 2.78 | 0.64 | 0.36 | **3.53** | **30.66** |
| | C | 22.0 | **5.23** | 0.89 | 0 | 0 | N.A. | 0 | **0** |
| 7 : End | A | 48.0 | 4.32 | 0.83 | 9.19 | 4.41 | 0.25 | 2.74 | 22.86 |
| | B | 28.0 | 5.22 | 1.04 | 4.93 | 1.38 | 0.33 | 0 | 0 |
| | C | 24.0 | **5.83** | 0.8 | 4 | 0.96 | 0.02 | 3.5 | **14.58** |

For all paths, the user who knew best the environment and the app (C) is generally the fastest to get to the destination and the one who looks in average the least the application. On the segment 6, subjects A and B have almost the same percentage to look at theirs phones, but the mean time is higher for the subject B. An explanation can be that the subject A looks often at the phone and quickly, as subject B looks to the phone but more time.

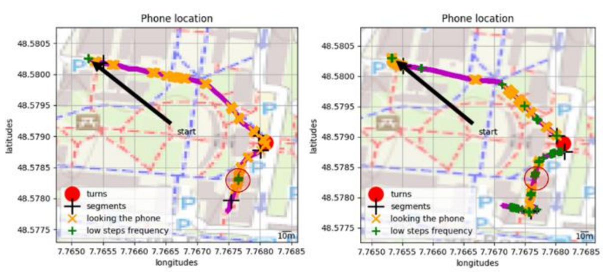


Figure 11. The person slows down (+ in green) and looking the phone detection (x orange) when pop up notification appears

In the second path, the notification of the destination appears suddenly in front of another building's entry, which disturbs the users (Figure 11). Indeed, subjects A and B slow down and look at the phone to be sure if it is the good entry. The third path does not show any abnormality.

## V. CONCLUSIONS AND FUTURE WORK

The proposed evaluating method seems to be robust as it works with several users' behaviors on different paths and segments. The notification, which appears suddenly, can disturb the user in ambiguous situations. Moreover, on the first path, the user B who knew the environment would not have taken necessarily the same path as shown by the app.

Using these first results, we have started to have first interpretation of the behaviors of the subjects using the app.

The phone detection could be improved by training a specific neural network for this issue. The eye tracker camera sensors for gaze and video are sensitive to brightness, it is recommended to test at the end of the day.

We will test our methodology in more complex environments to validate it. Moreover, some experiments will be done with disabled people to choose the appropriate sensors to analyze the user interface.

Finally, a qualitative questionnaire could be associated with this quantitative study to better understand the behavior and decisions during navigation and to have a feedback after the navigation experience.

## ACKNOWLEDGMENT

We would like to thank the entire Research medic@ team from Capgemini Engineering, particularly Julien Balbierer and Eric Bournez.

## REFERENCES

[1] J. Zegarra Flores, F. Pereme, J.-P. Radoux and L. Rasseneur, "Navi Campus : an enhanced GPS navigation app for University Campuses", 12th ITS European Congress, abstract TP0815, Strasbourg, June 2017, pp. 37

[2] R. Fryers, T. Holzer Saad and J. Dinsmore, "Report Defining The Needs of Stakeholders For a Wayfinding Platform For Individuals With Intellectual Disabilities and Their Carers," Trinity College Dublin, Trinity College Dublin, March 2018, pp. 1-29.

[3] I. Maly, J. Balata, O. Krejcir, E. Fuzessery and Z. Mikovec, "Qualitative measures for evaluation of navigation applications for visually impaired", 6th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), 2015, pp. 223-228

[4] Tobii AB, User's manual Tobii Pro Lab, Version 1.130, 12/2019., pp 82, http://www.vinis.co.kr/TPL_manual.pdf

[5] K. He, G. Gkioxari, P. Dollàr and R. B. Girshick, "Mask R-CNN", *IEEE International conference on computer vision ICCV*, pp. 2980-2988, 2017.