

TRAJECTORY CLUSTERING FOR THE CLASSIFICATION OF EYE-TRACKING USERS WITH MOTOR DISORDERS

A. Clemotte¹, H. Arregui², M.A. Velasco¹, L. Unzueta², J. Goenetxea²,

U. Elordi², E. Rocon¹, R. Ceres¹, J. Bengoechea³, I. Arizkuren³, E. Jauregui³

¹ CAR UPM-CSIC, Ctra. Campo Real, km 0.2, 28500, Arganda del Rey, Spain; erocon@csic.es

² Vicomtech-IK4, Paseo Mikeletegi, 57, 20009, Donostia, Spain; lunzueta@vicomtech.org

³ Irisbond, Paseo Mikeletegi, 57, 20009, Donostia, Spain; e.jauregui@irisbond.com

Abstract

This paper presents a pilot study completed in the framework of the INTERAAC project. The aim of the project is to develop a new human-computer interaction (HCI) solution based on eye-gaze estimation from webcam images for people with motor disorders such as cerebral palsy, neurodegenerative diseases, and spinal cord injury that are otherwise unable to use a keyboard or mouse. In this study, we analyzed cursor trajectories recorded during the experiment and validated that users with different diseases can be automatically classified in groups based on trajectory metrics. For the clustering, Ward's method was used. The metrics are based on speed and acceleration statistics from full filtered tracks. The results show that the participants can be grouped into two main clusters. The main contribution of this work is the evaluation of the clustering techniques applied to eye-gaze trajectories for the automatic classification of users diseases based on a real experiment carried with the help of three clinical partners in Spain.

Keywords: Eye-gaze estimation, Motor disorder, User-type classification, Trajectory clustering

1 INTRODUCTION

People with neurodegenerative diseases (ND), spinal cord injury (SCI) or cerebral palsy (CP) are generally not able to handle standard computer peripherals (keyboard or mouse) due to their severe motor impairment. These people require alternative communication systems in order to interact with their social environment and ambiance. The main objective of INTERAAC¹ is to develop a (1) low-cost and (2) multi-platform communication tool adapted to these three profiles of motor disability. As opposed to other video-oculography systems that are based on infrared cameras, the developed system will be based on algorithms for head pose and eye-gaze estimation from webcam images. INTERAAC faces two main challenges: (1) developing a cutting edge pointing device that

Head pose detection



Figure 1: Capture of a participant with PC during a work session in ASPACE-Cantabria.

would outperform current eye-gaze estimation systems in cost and features and (2) making the solution accessible to users with severe motor impairments.

The first step for the development of such tool is to group and describe the target population as detailed as possible in terms of their eye, head and trunk control as well as gaze strategies [3]. This is especially problematic for INTERAAC due to the wide range of motor disorders and disability profiles that we can expect to find among our users. Clinical scales such as Manual Ability Classification System (MACS) [2] have already been associated with the performance of people with motor impairment and traditional computer peripherals [1] but they are of no use for non-hand-held pointing input devices such the eye-gaze estimation systems. The classification will allow the assessing of the developed solution with each cluster independently. We hypothesized that our participants can be classified into groups depending on what kind of strategies they use to interact with the computer, so each group will be affected differently by the new development.

The aim of this work is to explore a methodology to group a population with a very diverse profile of motor impairments into a manageable number of descriptive groups or clusters. We hypothesized that our participants can be classified into two main groups. This paper presents the preliminary results of an experimental study done

¹<http://interaac.esy.es/es/proyecto/>

in three specialized centers for the classification of users with CP, SCI, and ND based on information from cursor velocity and acceleration, when the cursor is controlled with an eye-tracker. Indirectly, the classification will provide with the main objectives parameters or metrics that most quantify the performance of the new proposed solution for each of the impairment clusters.

2 BACKGROUND

There are solutions in the literature that analyze velocity and accuracy measures as well as cursor trajectories and other aspects of task performance in order to classify users.

Gajos et al. [4, 5] worked on the assessment of users' capabilities in order to create custom graphical computer interfaces (GUI). They analyzed pointing, dragging, list selection and multiple clicking and measured aspects such as the widget manipulation time, interface navigation time or error rate.

Their algorithm used this information for the dynamic personalization of the interface based on the user's preferences [4] or the user's motor skills [5] by minimizing two cost functions (of ability and preference).

Hurst et al. [6, 7] assessed pointing performance in order to develop systems that can automatically adapt to users' current needs in real-world computing environments. They developed a user's classifier that was able to discriminate between expert or novice users [6] between impaired and non-impaired users. In the classification process, they used different metrics to describe movement (the distance of movement), targeting (e.g. times the cursor enters the target), clicks (accidental clicks) and the breaks taken by the user [7].

In contrast to these studies, our approach is a technique to classify users based on metrics that rely on the interaction with the computer, in order to identify patterns, beyond the functional clinical scales.

3 METHODS

For this study, we designed an experimental protocol in collaboration with 3 centers specialized in motor disorders. In the experiments, we asked a total of 14 participants with severe motor impairments to control the cursor with movements of their eyes and to complete several series of pointing tasks. The locations of the cursor and the targets on the screen and the participants' point of gaze were recorded and analyzed.

3.1 Participants

Our clinical partners recruited a total of 14 participants with cerebral palsy (6) in ASPACE-Cantabria, spinal chord injuries (4) in Hospital Nacional de Paraplégicos de Toledo (HNPT) and several neuromuscular disorders (4) in ADEMGI. During the trials, a caregiver and/or therapist assisted the participants if needed. None of the participants had vision, hearing or cognitive problems that might cause a misunderstanding of the task or prevent them from finishing it properly. The local ethical committee gave approval to the study. All participants were informed beforehand and signed a written informed consent to participate.

During the following sections the users are identified by their type of disease: six users with cerebral palsy (CP_i), four users with spinal cord injuries (SCI_i), and 4 with neurodegenerative diseases: one user with ataxia (ATX), one user with amyotrophic lateral sclerosis (ALS), one user with multiple sclerosis (MSC) and another one with neuromuscular disorders (NMD).

3.2 Instruments

The participants used the IRISBOND PRIMMA eye-tracker² (IRISBOND, Spain) to translate the eye-gaze estimation into cursor coordinates. In order to perform a click action, the participants had to maintain the cursor within a pre-defined area during a pre-defined *dwelt time*.

A webcam recorded the whole work session and 2 inertial measurement units (IMU) registered the orientations of the participant's head and the computer screen. Thus, this setup allows comparing the accuracies of new head pose and eye gaze estimation techniques to be developed within the INTERAAC project, with the data registered by the other devices.

3.3 Procedure

All participants were sitting in their wheelchair at an optimal distance from the eye-tracker that guaranteed the best measures (around 60 cm). The routine of the participants' accommodation included moving the participant, the screen, and the eye-tracker, see Figure 1.

At the beginning of the work session, each participant went through a calibration process. A quantitative threshold was established to check the reliability of the calibration before the execution of the task. Only participants that exceeded this level continued with the tests.

²<http://www.irisbond.com/productos/irisbond-primma>

After the calibration, each participant practiced in a short training session with the ‘Look to Learn’ software³ (Sensory Software) in order to get used to the interface and the rest of the working environment (screen, camera, etc.). No data was collected during this phase.

Finally, they were asked to complete a series of pointing tasks with the FittsStudy software [11]. FittsStudy allowed us to control the size of the targets, the distances between them and the order of appearance, i.e. the difficulty of each pointing task. All the information from the pointing tasks and gaze estimation was recorded for further analysis.

In general, the work sessions took approximately 40 minutes. This time includes the preparation of the participants, the calibration process, possible repetitions and some breaks between trials to avoid the apparition of fatigue.

Calibration process

In the calibration, the participants were asked to look at several targets on the screen (5, 9 or 16 depending on their attention capabilities). The positions of these targets were known, but their order of appearance was randomized. As the participant looked at these stimuli the eye-tracker estimated the spatial relationship of the pupil and the corneal reflex.

At the end of the process, the algorithm provided us with the quality of the calibration. That information was used in order to decide whether starting the tests or repeating the whole process (in case the calibration output was below our threshold, 1.5 degrees).

Pointing task

The proposed test consisted in looking at 4 series of 10 concentric circles on the screen, as Figure 2 shows. This gives a total of 40 clicks for each participant. The target circle where participants should click was differentiated by a darker shade of blue.

In order to control the task difficulty of each series of circles, we used the concept of *index of difficulty* (ID). The Shannon formulation used by MacKenzie (1989) represents the ID as:

$$ID = \log_2 \left(\frac{A}{W} + 1 \right) \quad (1)$$

where A is the amplitude of the movement needed

³<https://thinksmartbox.com/product/look-to-learn/>

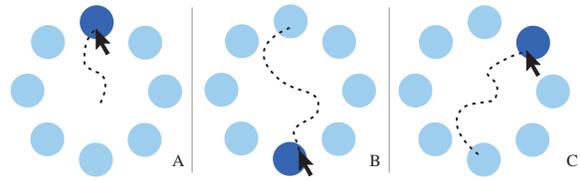


Figure 2: Representation the pointing tasks. The sequence A, B, C shows 3 consecutive positions of the target (dark blue).

Table 1: Sequence of the presets index of difficulty (ID). The corresponding amplitude of movement and width of the target is measured in pixels.

Sequence	ID (bits)	A (px)	W (px)
1	1.41	500	300
2	1.80	500	200
3	2.58	500	100
4	3.46	500	50

to reach the target and W is the size of the target. Fitts’ law establishes that there is a linear relationship between the movement time (the time needed to reach and click on it) and the value of ID ; the more difficult the task is, the longer it takes to be completed.

This index allowed us to set a range of difficulty between 1.41 and 3.46 bits that would ensure successful targeting and that is very similar to the ranges of difficulty that the participants would find in real-world desktop applications. The values of ID ’s were distributed in increasing order, see Table 1. For each of these ID ’s we defined the values of A and W . This sequence was repeated for all participants.

In order to click on the targets, the participants were asked to complete a *dwel time*. A click action is ordered when the dispersion of the cursor positions in both horizontal and vertical axes during the last second is below than a preset threshold of 68 px (or 2 cm in the screen). The minimum time between two consecutive clicks was set to 1 s in order to avoid the rebound effect (i.e. the detection of undesired multiple clicks).

Participants received two kinds of feedback during the tasks: whereas the participant had visual feedback of the cursor position throughout all the session, failed tasks produced a distinctive sound.

3.4 Methodology for the analysis

Dataset

Tracks can be understood as continuous trajectories sampled at 40ms periods after participants completed the pointing task. These tracks con-

sist of the chronologically ordered set of X-Y coordinate pairs from: left eye, right eye, combined eye (CT, combination between left and right eye) and, filtered eye (filtered trajectory of the CT to clean out saccades and jumps). The whole testing produced 14 tracks from differences participants: 4 from ADEMG, 6 from ASPACE and, 4 from HNPT.

In order to obtain some features of the pointer behavior on the screen, instantaneous speed and acceleration are estimated from position differentials captured by the eye-tracking system from filtered tracks (FT). The hypothesis is that these speed and acceleration patterns can be used to classify different users. In other scenarios, these metrics are widely used when classifying drivers from their driving trajectory patterns [8].

Instantaneous speed and acceleration are then summarized with the following set of statistics: $avg(speed)$, $max(speed)$, $median(speed)$, $max(accel)$, $min(accel)$ and $median(accel)$. Average acceleration and minimum speed values equal 0 for the full set of tracks, and thus, are discarded. These track features are expected to represent the way each user tackles the eye movement towards the target point on the screen.

PCA and Clustering

Once each track is characterized by a set of metrics, an exploratory analysis of these variables is done to evaluate their representativeness. The technique chosen for this task was the Principal Components Analysis (PCA), which is a statistical procedure for dimensionality reduction. It uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables (principal components). One of the applications of this dimensionality reduction is representing multidimensional data in 2-D plots. We decided to reduce the data to 2 and 3 dimensions and explore the data produced. We established a threshold of 90% for the PCA algorithm.

The next step is performing the clustering itself, based on the features that have been proposed. The objective of the analysis is to discover similar behaviors among users in order to group them in clusters. Any clustering algorithm requires a dissimilarity method for calculating dissimilarities between trajectories [9]. In our work, Ward's hierarchical agglomerative clustering [10] has been used comparing two dissimilarity or distance metrics (Euclidean and Manhattan distances) in order to evaluate if the clusters obtained, if any, are consistent in both metrics. To check the quality of the clusters, from a quantitative point of view,

Table 2: Composition of the three main principal components (93.3% of variance explained)

Variable		PC1	PC2	PC3
Speed	Avg.	0.130	-0.668	-0.500
	Median	0.480	-0.284	-0.327
	Max.	-0.465	-0.343	-0.059
Accel.	Median	0.477	-0.368	-0.347
	Max.	-0.500	-0.307	-0.124
	Min.	0.243	0.355	-0.389

an additional partitioning around medoid (PAM) was performed.

4 RESULTS

4.1 Trajectory sample

Figure 3 shows one pointing task, with the trace of the filtered track in green and the target in blue, obtained during a HNPT user test. In addition, is depicted the velocity and acceleration of this track. This a sample of a pointing task performed by one participant. In the speed graph, we can see that a click took about 200 milliseconds, which is what usually lasts a saccade.

4.2 PCA Analysis

The PCA analysis shows that using the two first principal components, components are able to explain 78.3% of variability. A graphical representation of the different users is given in Figure 4, showing some subtle group patterns, especially ASPACE users (CP). Three linearly independent components are able to explain a significant proportion of the variability of the dataset: 93.3%. Table 2 explains the contribution of each variable. With a threshold of 90%, this second reduction to 3 components is acceptable.

4.3 Trajectory clustering

Hierarchical clustering methods have been tested for the features selected, representing the cluster classification by a dendrogram. Euclidean and Manhattan distances are chosen to compare hierarchical clustering results, as shown in Figure 5. The clusters are not the same but, in both cases, users with CP are all grouped together. Moreover, in the second case these users are clearly separated from rest. In both cases, the existence of two clusters is quite clear.

PAM was performed with parameter $k=2$: the average silhouette width obtained is 0.44 (Manhattan) vs 0.37 (Euclidean), which means that a clustering structure may exist but it is relatively

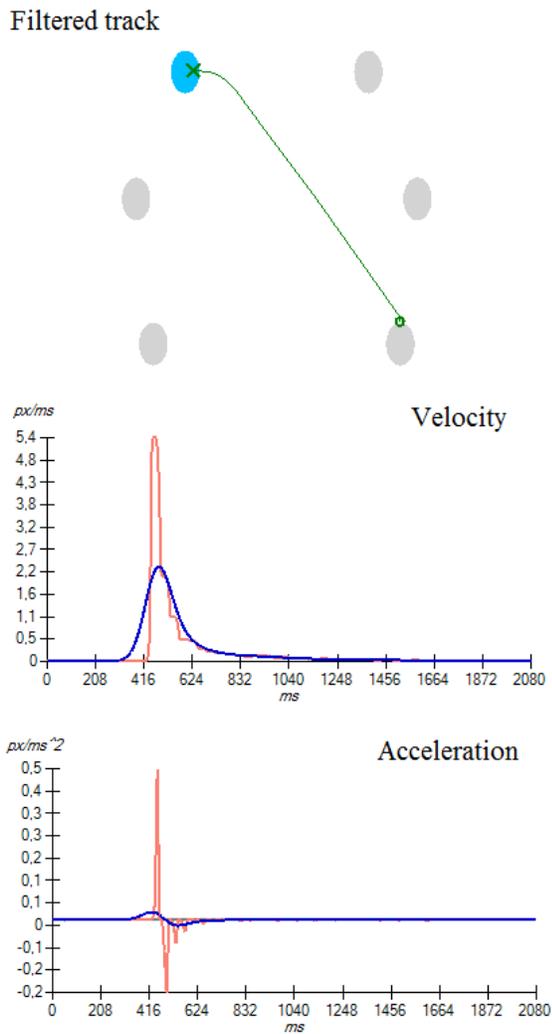


Figure 3: A partial track of a HNPT participant. The amplitude of movement (distance between circles) is 500 pixels and the target width (diameter of the circles) is 50 pixels. Below the partial track, speed and acceleration of this movement is depicted. From the velocity chart we can see that the click take it to the participant approximately 200millisec.

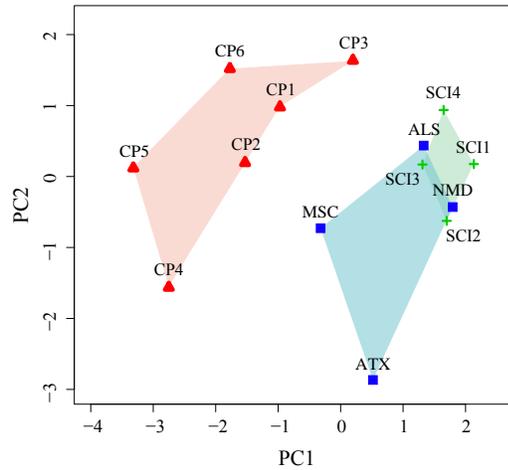


Figure 4: Representation of the users projected over the two main principal components. Pluses are SCI_i , triangles represent are CP_i and squares correspond to ATX, ALS, MSC and NMD.

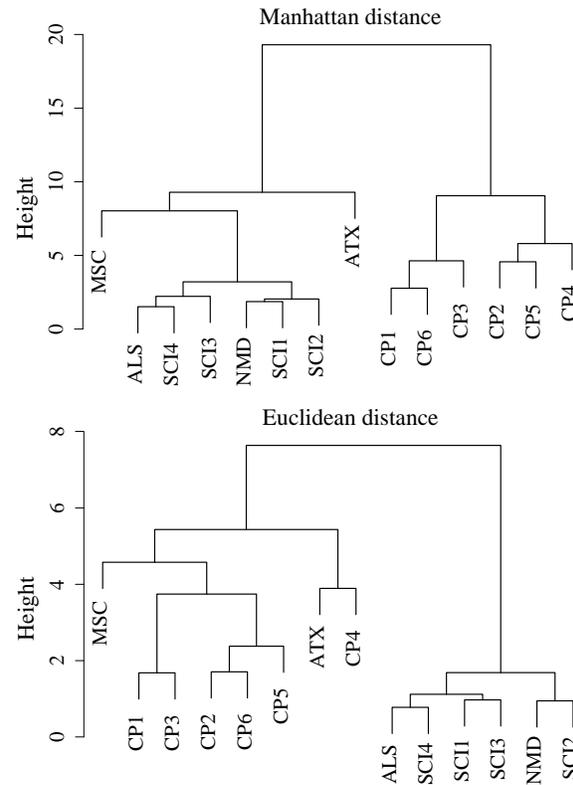


Figure 5: Dendrogram showing clustering results with Manhattan and Euclidean distance metrics and Ward method. Two main clusters can be found.

low. The two clusters given by the Manhattan distance-based PAM differs from the agglomerative clustering in only one individual (CP3).

5 CONCLUSION

One of the objectives of the INTERAAC project is to discover groups or clusters that could somehow be related to the diseases of the participants. The hypothesis is that not all users have the same strategy to control the computer so that each user group will be benefited differently by new developments.

In this paper, we presented a pilot study of an automatic user-type classification procedure based on the analysis of point-of-gaze trajectories recorded by an eye-tracking system. The tests were carried out with three different groups of users, all of them with severe levels of motion disorders. The results showed that, within the 14 users, at least two groups could be deduced where cerebral palsy and spinal cord injuries were separated. It was concluded that simple speed and acceleration metrics of the eye-tracked trajectories can be used to group users, which would help in tuning automatically further parameters of the eye-tracking system and, therefore, improving the interaction with the computer.

Further work will include: replicating the analysis with partial tracks between objective points instead of full tracks, evaluating the trajectories together with inertial measurements given by head-mounted sensors and, add a metric that measures the difference between the track of each eye.

Acknowledgments

We would like to thank HNPT, ASPACE-Cantabria, ADEMG, and especially the participants of the experiments for their support. This work has been funded by the Spanish Ministry of Economy and Competitiveness, under the call Retos-Colaboración 2015 of the the National Programme for Research Aimed at the Challenges of Society 2009-2016 (RTC-2015-4327-1). A. Clemotte wants to thank Itaipu Binacional.

References

- [1] C. Davies, A. Almanji, and N.S. Stott. A cross-sectional study examining computer task completion by adolescents with cerebral palsy across the Manual Ability Classification System levels. *Developmental Medicine & Child Neurology*, 56(12):1180–1186, 2014.
- [2] A.-C. Eliasson, L. Krumlinde-Sundholm, B. Rösblad, E. Beckung, M. Arner, A.-M. Ohrvall, and P. Rosenbaum. The Manual Ability Classification System (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. *Developmental Medicine and Child Neurology*, 48(7):549–554, 2006.
- [3] L. Fridman, J. Lee, B. Reimer, and T. Victor. "Owl" and "lizard": Patterns of head pose and eye pose in driver gaze classification. *arXiv preprint*, pages 1–9, 2015.
- [4] K. Gajos and D.S. Weld. Preference elicitation for interface optimization. In *Proc. ACM Symposium on User Interface Software and Technology*, page 173, New York, USA, 2005.
- [5] K.Z. Gajos, D.S. Weld, and J.O. Wobbrock. Automatically generating personalized user interfaces with Supple. *Artificial Intelligence*, 174(12-13):910–950, 2010.
- [6] A. Hurst, S.E. Hudson, and J. Mankoff. Dynamic detection of novice vs. skilled use without a task model. In *Proc. SIGCHI Conf. Human Factors in Computing Systems*, page 271, New York, USA, 2007.
- [7] A. Hurst, S.E. Hudson, J. Mankoff, and S. Trewin. Distinguishing users by pointing performance in laboratory and real-world tasks. *ACM Trans. Accessible Computing*, 5(2):1–27, 2013.
- [8] N. Lin, C. Zong, M. Tomizuka, P. Song, Z. Zhang, and G. Li. An overview on study of identification of driver behavior characteristics for automotive control. *Mathematical Problems in Engineering*, 2014:15 pages, 2014.
- [9] H. Munaga, L. Ieronutti, and L. Chittaro. CAST: A novel trajectory clustering and visualization tool for spatio-temporal data. In *Proc. Int. Conf. Intelligent Human Computer Interaction*, pages 169–175, New Delhi, India, 2009.
- [10] F. Murtagh and P. Legendre. Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *Journal of Classification*, 31(3):274–295, 2014.
- [11] J.O. Wobbrock, K. Shinohara, and A. Jansen. The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proc. Conf. on Human Factors in Computing Systems*, page 1639, New York, USA, 2011.