

Gaze-Based Interaction Adaptation for People with Involuntary Head Movements (Student Abstract)

Cindy Tong^{1, 3}, Rosanna Chan^{2*, 3}

¹ Nanyang Technological University, Singapore

² Centre for Perceptual and Interactive Intelligence, Hong Kong

³The Chinese University of Hong Kong, Hong Kong
to0001di@e.ntu.edu.sg, yychan@ie.cuhk.edu.hk

Abstract

Gaze estimation is an important research area in computer vision and machine learning. Eye-tracking and gaze-based interactions have made assistive technology (AT) more accessible to people with physical limitations. However, a non-negligible proportion of existing AT users, including those having dyskinetic cerebral palsy (CP) or severe intellectual disabilities (ID), have difficulties in using eye trackers due to their involuntary body movements. In this paper, we propose an adaptation method pertaining to head movement prediction and fixation smoothing to stabilize our target users' gaze points on the screen and improve their user experience (UX) in gaze-based interaction. Our empirical experimentation shows that our method significantly shortens the users' selection time and increases their selection accuracy.

Introduction

Gaze-based interaction technologies such as eye trackers are becoming more popular in our daily lives. In addition to being an access device for games and entertainment, eye trackers are also used by people with physical disabilities to access computers and communication technologies. For example, a human-computer interaction (HCI) user study shows that 90% of the participating children having mild to moderate CP could successfully interact with an experimental system using eye-tracking devices (Cruceat and Butean 2019). An extensive investigation was also performed with young children having dyskinetic CP, which is a complex form of CP (Monbaliu et al. 2017). Although all five participants completed the trial successfully, a parent expressed that the frequent necessary device re-calibration was frustrating (Karlsson, Bech, and Stone 2019).

Most existing eye trackers are based on real-time video analyses (Vedaldi and et al. 2020), where a camera is used to detect the pupil centre and corneal reflection to estimate the user's fixation points. This approach assumes the stability of the user's head, which is often not the case for people having Athetoid or dyskinetic CP and some of the post-stroke

patients (Siniscalchi et al. 2012). These users often have involuntary abnormal movements with patterns different from natural head movements. Therefore, gaze-based interaction still remains challenging for this population group.

In this paper, we propose and develop a new eye-tracking algorithm with involuntary head movement adaptation. Our algorithm (1) adjusts the gaze point on the screen according to the detected head position and pose, and (2) proactively controls the gaze moving speed on the screen. We have performed a preliminary user study with two participants having involuntary head movement. Our results show that the gaze-based interaction performance in both participants was effectively promoted by our adaptation approach.

The Current Work

Problem Description

The current work investigates how computer vision and machine learning techniques can be applied to improve the gaze-based interaction experience for users having involuntary head movements. Specifically, there are two sub-problems that need to be solved:

(1) How can we maintain a smooth and accurate interaction experience even when the user's head is not in the optimal position for the eye-tracker?

(2) How can we stabilise the gaze trace on the user interface so as to facilitate other gaze-based interactions?

Our Approach

Here, we propose the **head movement detection and smoothing algorithm (HMDSA)** as a solution to the above problem. HMDSA contains two key functions, namely, the head pose monitor, and the gaze points speed controller.

Head Pose Monitor $\text{pose}(\mathbf{g}, \mathbf{h}, \theta)$ The purpose of this function is to overcome any gaze point instability due to the user's involuntary head movement. This is particularly important for gaze-based UX, especially when the selection target is small or located at the screen corners. It produces an adapted gaze point \mathbf{g}' based on the 2-dimensional detected gaze point \mathbf{g} , the 3-dimensional detected head position \mathbf{h} and the y-axis Euler rotation angle of the user's head θ .

Gaze speed Controller $\text{speed}(\mathbf{g}, \bar{\mathbf{g}}, \mathbf{h}, \bar{\mathbf{h}}, S, T, \lambda)$ The purpose of this function is to oppose any sudden changes in

*This study was partially supported by the Centre for Perceptual and Interactive Intelligence (CPII) Ltd under the Innovation and Technology Commission (ITC)'s InnoHK. Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Algorithm 1: HMDSA

Input: Eye-tracker detected gaze point $\mathbf{g} = (g_x, g_y)$, head position $\mathbf{h} = (h_x, h_y, h_z)$, Euler rotation angle (y-axis) θ , moving-average gaze point $\bar{\mathbf{g}} = (\bar{g}_x, \bar{g}_y)$, moving-average head position $\bar{\mathbf{h}} = (\bar{h}_x, \bar{h}_y, \bar{h}_z)$.

Parameter: Window timeframe T , speed threshold S , impact factor λ .

Output: Adjusted gaze point $\tilde{\mathbf{g}} = (\tilde{g}_x, \tilde{g}_y)$

```

1: Initialise  $\mathbf{g} = \bar{\mathbf{g}} = (0, 0)$ ,  $\mathbf{h} = \bar{\mathbf{h}} = (0, 0, 0)$ , and  $\theta = 0$ 
2: while user interaction continue do
3:   check if gaze point is detected
4:   if no gaze point detected then
5:      $\tilde{\mathbf{g}} = \bar{\mathbf{g}}$ 
6:   else
7:     obtain  $\mathbf{g}, \bar{\mathbf{g}}, \mathbf{h}, \bar{\mathbf{h}}, \theta$ 
       adapt gaze point based on head pose:  $\mathbf{g}' = \text{pose}(\mathbf{g}, \mathbf{h}, \theta)$ 
       adapt head position based on gaze speed:  $\mathbf{h}' = \text{speed}(\mathbf{g}', \bar{\mathbf{g}}, \mathbf{h}, \bar{\mathbf{h}}, S, T, \lambda)$ 
       compute  $\tilde{\mathbf{g}}$  based on  $\mathbf{g}'$  and  $\mathbf{h}'$ 
8:   end if
9:   update  $\bar{\mathbf{g}}, \bar{\mathbf{h}}$  and output  $\tilde{\mathbf{g}}$ 
10: end while

```

head position caused by involuntary head movements. Having specified the window timeframe T , the gaze speed s can be derived from the current gaze point \mathbf{g} and its moving average $\bar{\mathbf{g}}$ by

$$s = \|\mathbf{g} - \bar{\mathbf{g}}\|/T. \quad (1)$$

With the speed threshold S that limits the gaze speed and the impact factor λ that decides the confidence level of the current measurement, our smoothing function outputs a 3-dimensional vector \mathbf{h}' as the adapted head position where:

$$\mathbf{h}' = \lambda \mathbf{h} + (1 - \lambda) \bar{\mathbf{h}} \quad (2)$$

where $\lambda = 1 - 2/(1 + e^{s/S})$.

User Experiment

Participants, Procedures, and Measurements

We have performed a within-subject HCI experiment to evaluate the effectiveness of our approach (Purchase 2012). Our participants were two patients (1 f, 1m) having involuntary head movements caused by Hypertension and Cervical Strain, respectively. Written consent from both participants has been obtained prior to the commencement of the study. We developed a gaze-based interaction game using Unity and Tobii 5's SDK, where the participants were asked to perform the 1-out-of-8 gaze-based icon selection using the Tobii Eye Tracker 5. Each trial comprises 100 selection tasks. Each participant has performed 2 trials (1 experimental condition and 1 control condition). We have collected and analysed over 144,000 frames of video data that last around 40 minutes. Participants' performance was measured by participant's icon *selection time* and *selection accuracy*, where data was collected from the experimental condition (with adaptation) and the control condition (without adaptation) for statistical analyses.

	Selection Time (s)	Selection Accuracy (%)
	Mean (SD)	Mean (SD)
Experimental	4.79 (1.52)	92 (27.2)
Control	7.72 (4.32)	51 (51.1)

Table 1: Descriptive statistics of user experiment (N = 400).

Results

Descriptive statistics of our performance metrics are provided in Table 1. Analysis of variance (ANOVA) shows statistically significant differences across conditions in both selection time ($p < .001, F(1, 398) = 82.39$) and selection accuracy ($p < .001, F(1, 398) = 103.41$). Furthermore, the effect size is between small to medium for both selection time ($\eta^2 = .17$) and selection accuracy ($\eta^2 = .21$).

Concluding Remarks

Our results show an enhancement of gazed-based interaction brought by our algorithm. Specifically, the decrease in selection time and increase in selection accuracy were both statistically significant with a small to medium effect size. Our limitations include (1) our algorithm has a lower performance when users wear glasses, (2) our head pose monitor function performs better in horizontal rather than vertical head movement. Therefore, adaption for more real-life usage scenarios needed to be developed. Also, the sample size needs to be increased in order to generalise our findings. Overall, we proposed a novel adaptation algorithm to improve the UX during gaze-based interaction, with limitations and future works briefly identified and discussed.

References

- Cruceat, A.-M.; and Butean, A. 2019. Assistive Tools for People with Cerebral Palsy: An Eye Tracker Calibration for Vision and Focus Training. In *21st SYNASC*, 284–289.
- Karlsson, P.; Bech, A.; and Stone, H. 2019. Eyes on communication: trialling eye-gaze control technology in young children with dyskinetic cerebral palsy. *Developmental Neurorehabilitation*, 22(2): 134–140.
- Monbaliu, E.; Himmelmann, K.; Lin, J. P.; Ortibus, E.; Bonouvrié, L.; Feys, H.; Vermeulen, R. J.; and Dan, B. 2017. Clinical presentation and management of dyskinetic cerebral palsy. *The Lancet Neurology*, 16: 741–749.
- Purchase, H. C. 2012. *Experimental Human-Computer Interaction: A Practical Guide with Visual Examples*. New York: Cambridge University Press, 1st edition.
- Siniscalchi, A.; Gallelli, L.; Labate, A.; Malferrari, G.; Palleria, C.; and Sarro, G. D. 2012. Post-stroke Movement Disorders: Clinical Manifestations and Pharmacological Management. *Current Neuropharmacology*, 10(3): 254–262.
- Vedaldi, A.; and et al. 2020. Towards End-to-End Video-Based Eye-Tracking. In *Computer Vision - ECCV 2020*, volume 12357, 747–763. Switzerland: Springer.