Multimodal Target Prediction Model

Pradipta Biswas and Pat Langdon
University of Cambridge, UK
{pb400, pml24}@cam.ac.uk

Abstract
This paper presents a Neural Network based model that can be used to predict pointing target for both physical and situational impairment. The model takes different trajectory profiles like velocity, acceleration and bearing of movement as input parameters and based on that predicts next pointing target. We reported three user studies – one involving users with physical and age-related impairment using a mouse and the other two involved able-bodied users using head and eye-gaze tracking based systems. We found that the model can accurately predict target in all cases. Finally we proposed an adaptation system using the target prediction model that can statistically significantly reduce pointing times.

Keywords
Target Prediction, Neural Network, Adaptive system.

ACM Classification Keywords
D.2.2 [Software Engineering]: Design Tools and Techniques – user interfaces; K.4.2 [Computing Milieux]: Social Issues – assistive technologies for persons with disabilities

Introduction
This paper presents a neural network based model to predict pointing target that has been tested with three different input modalities and also has been used to develop a cursor adaptation algorithm to reduce pointing time. Researchers already worked on algorithms to reduce pointing time through determining the difficulty of a task using Fitts’ Law [4], increasing target size [6, 10], employing larger cursor activation regions, moving targets closer to cursor location, dragging cursor to nearest target or changing CD ratio[13] and so on. It is certain that these algorithms will perform even better in the existence of a target prediction algorithm so that only correct or most probable targets could be dynamically altered. With this in mind, researchers proposed target or end-point prediction methods.

One of the first algorithms for target prediction was suggested by Murata, which calculates the angle deviation towards all possible targets and select the target with minimum deviation. The results show that pointing time can be reduced by 25% using this algorithm [11]. Asano and colleagues [1] points out that having more than one target on a particular movement direction results in poor performances of the afore mentioned algorithm, especially when dealing with target located far away. They used
previous research results about kinematics of pointing tasks and showed that peak velocity and target distance has a linear relationship. They predict the endpoint with linear regression using the relationship between peak velocity and total distance to endpoint [1]. Lank and colleagues [8] also employed motion kinematics where they assume minimum jerk law for pointing motion and fit a quadratic function to partial trajectory to predict endpoint. Target prediction algorithms will really be beneficial where people may find a regular pointing device difficult to use. Two possible scenarios are users with motor impairment and first time users for a novel interaction device. There are plethora of studies demonstrating difficulty of people with age-related or physical impairment in stopping pointer movements or clicking on small target. Cursor movements vary in characteristics for motor impaired users since they experience tremor, muscular spasms and weakness [7]. The velocity profile includes several stops and jerky movements. This needs to be taken into account when applying target prediction. State space filtering techniques are promising [5] in estimating intended targets as well as smoothing cursor trajectories since it is possible to model the movement, fine-tune or adapt the parameters for different users. Another possible scenario is using novel interaction device involving head or eye-gaze movement. Devices like head or eye trackers are not as easily controllable as a desktop mouse or touchpad and people may struggle to use these devices to precisely control pointing movement. Devices like head or eye trackers have immense applications for both physical and situational impairment (e.g. operating AAC software by disabled users using eye tracker and locking a target by fighter aircraft pilots using head tracker).

However there are not much reported work on target prediction for motor impaired users except Godsill’s work [5] on particle filter and Wobbrock’s work [13] on Angle Mouse. There are also almost no attempt to use target prediction algorithms for novel interaction devices like eye-gaze, gesture or head tracking systems. This paper reports three user studies where we used a neural network based target prediction system for elderly users using a computer mouse and able-bodied users using a head tracker involving Brain-Computer Interface (BCI) and an eye-gaze tracking based system.

**Target Prediction System**

We have implemented a bank of neural networks considering all possible combinations of three movement properties (bearing, velocity and acceleration of movement). We have used supervised learning to detect users’ movement phases from movement features. Initially we trained the neural networks with a multiple distracter task. Later we used the same task to evaluate the networks following our evaluation criteria. Our previous paper [15] already compared this model with a Kalman filter based model and neural network based model performed better. Our training and test set of participants were different. We conducted three user trials involving three different input modalities:

1. The first trial involved people with age related and physical impairment and they used a standard computer mouse
2. The second trial involved a head tracker and Brain Computer interface used by able-bodied users who never used similar system before.
3. The third trial involved an eye-gaze tracker used by able-bodied users who never used similar system before.
We described these trials in the following sections. Besides the target prediction system we also present an analysis of cursor trajectories for the head and eye-gaze tracking systems.

**User Trials**

**Participants**

We have collected data from 23 users using mouse. The users used to operate computer everyday and volunteered for the study. The group of participants included users with a wide range of abilities in terms of visual and motor impairment. Age related impaired users were more than 60 years old and physically impaired users suffer from Cerebral Palsy or Spina bifida.

We trained the neural networks with 13 users among which five have age related or physical impairment (like cerebral palsy). Then we test the system with 10 participants among which five have age related visual and motor impairment. The gender was balanced to nearly 1:1 in both training and test cases.

The head tracking system was evaluated by eight able-bodied users (5 male, 3 females) aged between 20 and 35. They never used any head tracking system before.

The eye tracking system was evaluated by eight able-bodied users (5 male, 3 female) aged between 20 and 45. They never used the eye tracking system before. They did not have any trouble in using the experimental setup.

**Material**

The study was conducted using a 21” screen (435 mm × 325 mm) with 1600 × 1200 pixels resolution and a standard computer mouse. The head tracking was conducted using a Emotiv Epoch Brain-Computer Interface Headset [2] while a Tobii X120 eye tracker[12] is used for eye tracking.

**Procedure**

The task was like the ISO 9241 pointing task with multiple distractors on screen (figure 1). We tried to strike a balance between the complete natural interaction scenario of Input Observer system [3] and the controlled single target task [4] of traditional Fitts’ Law analysis. We developed software to conduct this task which automatically records mouse locations and mouse events on screen every 16 msec.

Users need to click the button at the centre of the screen and then the target button appears with other distractors. We used five different target sizes (20, 40, 50, 60 pixels) and source to target distances (100, 140, 180, 240, 300 pixels). The participants using mouse were instructed to click target for 10 minutes after they were briefed about the procedure.

For the head tracking based system, we used five different target sizes (40, 50, 60, 70, 80 pixels) and source to target distances (100, 160, 220, 280, 350 pixels). The participants were instructed to click target for 5 minutes after they were briefed about the procedure.

We calibrated the eye tracking system with 9 dots before start of trial with each participant. The system was kept on calibrating until the Tobii SDK did not need any further recalibration. Users need to click the button at the centre of the screen and then the target button appears with other distractors. We used five different target sizes (60, 70, 80, 90 and 100 pixels) and source to target distances (160, 220, 280, 350, 400 pixels). We have to change the target size and amplitude than the head tracking system as it was too difficult even for the developer to select target less than 60 pixels wide using the eye tracking system. The participants were instructed to click target for 5 minutes after calibration.

**Results**

We have defined the following three parameters to evaluate the quality of a target prediction algorithm.

1. **Availability**: In how many pointing tasks the algorithm makes a successful prediction.
2. **Accuracy**: Percentage of correct prediction among all predictions.

3. **Sensitivity**: How quickly an algorithm can detect intended target.

Figure 1. Multiple Distractor task

Figure 2 plots the average percentage of availability and accuracy in different systems. In figure 2, different combinations of features are plotted in X-axis while the average Availability and Accuracy is plotted in Y-axis. The model only fires (or turns available) when it detects a change in movement phase. In the availability graphs, green bars show the percentage when it correctly identified target, the red bar shows when it could not. In certain occasions it fails to detect this change of movement phases and so the availability is not always 100% (white bars). Figure 2 also plots the sensitivity of the system. The x-axis shows the fraction of pointing time spent in a scale of 100. The y-axis represents the probability of correct target prediction. We found that velocity and bearing of movements have highest availability for mouse and eye-gaze tracking system while the combination of all parameters have highest availability of head tracking system. Velocity and Bearing also have the highest accuracy for all systems. The sensitivity is found to be rather same for all possible combinations of parameters. Video demonstrations of the multiple distractor task is available at http://youtu.be/QsxDxcccwAw and target prediction system is available at http://youtu.be/p9YOKJ59TiY

**Discussion**

This study proposes a new model for target prediction based on Neural Network, and applied this model in two different scenarios. The model accurately detects change in pointing phase in more than 70% of the pointing tasks for head and eye-gaze trackers and 65% for mouse used by people with age-related or physical impairment. The accuracy of target prediction is nearly 60% for all cases, however it reaches more than 90% as the user reaches near the target as shown in the Sensitivity graph (figure 2). This model is better than previous results [14]. For example, Ziebart’s models [14] achieve more than 50% accuracy after crossing 70% of pointing time while our models have more than 70% accuracy at similar stage. Future work will investigate other statistical models for target prediction and modalities of interaction like gesture recognition system.

**Adaptive System**

The main purpose of developing the target prediction system is to reduce pointing times. After validating the target prediction model, we tested its performance again as an adaptation algorithm. We have developed an adaptation system that enlarges a target whenever it is predicted as a probable target. Recent work by Hwang [6] already found similar technique can reduce pointing times for older adults in a single target task though Lee [9] did not find a significant reduction of pointing times in multiple distractor task. We used the similar multiple distractor set up and head tracker as our previous studies described above. In this study, whenever the target prediction system predicts a target, we increased its size by 1.5 times. We do not change the colour or any other property of the target.
Participants
We have evaluated this adaptation system by six able-bodied users (4 male, 2 female) with an age range between 22 and 45.

Material
We used the same screen and head tracker hardware and software for this study.

Procedure
It was same as the multiple distractor task described above, participants were not trained with the adaptation system before the trial, they were using it for first time.
Results

We recorded 181 pointing tasks and there were 5 instances of wrong selection. A SIZE × ADAPTATION ANOVA on pointing times found a significant effect of adaptation \((F(1,27) = 31.53, p<0.01, \eta^2 = 0.54)\) while a DISTANCE × ADAPTATION ANOVA also found a significant effect of adaptation \((F(1,31) = 30.05, p<0.01, \eta^2 = 0.49)\).

Discussion

This study shows that the target prediction system can reduce pointing time when coupled with a target magnification system. Although this study only evaluated head tracker but future work will include other modalities of interaction. It can be noted that the differences between adapted and non-adapted pointing times increases for smaller targets and longer source to target distances. We can assume that presence of such target prediction and adaptation system will not only reduce pointing times for existing eye-gaze and head tracking interfaces but also allow better utilization of screen space as target sizes can be reduced and more targets can be used in a single screen.

References