

# Evaluation of Vision-based Head-Trackers for Assistive Devices

S.P. Guiness<sup>1</sup>, F. Deravi<sup>1</sup>, K. Sirlantzis<sup>1</sup>, M.G Pepper<sup>2</sup> and M. Sakel<sup>2</sup>

**Abstract**—This paper presents a new evaluation methodology for assistive devices employing head-tracking systems based on an adaptation of the Fitts Test. This methodology is used to compare the effectiveness and performance of a new vision-based head tracking system using face, skin and motion detection techniques with two existing head tracking devices and a standard mouse. The application context and the abilities of the user are combined with the results from the modified Fitts Test to help determine the most appropriate devices for the user. The results suggest that this modified form of the Fitts test can be effectively employed for the comparison of different access technologies.

## I. INTRODUCTION

An equipment, or product, that can be used to increase, maintain, or improve the functional capabilities of individuals with disabilities can be defined as an Assistive Device [1]. A survey conducted by the World Health Organization (WHO) estimates that 1 billion people around the world live with some form of disability. Head tracking systems form a key access component for assistive devices when standard access devices such as keyboards and mice are no longer appropriate. This is especially the case for people suffering from neuro-disabilities such as brain injury, stroke, Amyotrophic Lateral Sclerosis (ALS), Cerebral Palsy (CP) and Multiple Sclerosis (MS). In this paper we present a methodology based on the Fitts Test for the effective evaluation and comparison of different assistive devices. In particular, we compare three head pointing systems using vision-based head tracking. One is a system under development by the authors, Headtracker, and the remaining two are existing systems SmartNav and CameraMouse. These three systems are compared with a standard mouse. The results of these evaluations are used to measure their relative performance and to investigate the feasibility of the adapted Fitts Test [2] as an evaluation methodology for head tracking technology. The evaluation can also be used to help determine the assistive device best suited for an individual and used to find the most suitable resolution of the screen for the individual.

The rest of the paper is organized as follows. A review of the background information on head tracking, Fitts Law and Fitts Test is presented in Section II. In Section III, the evaluation methodology is discussed. In Section IV, the experimentation and the procedures followed for the experiments conducted are discussed. In Section V, the results

are analyzed and discussed. Finally, Section VI contains the conclusions of the investigation.

## II. BACKGROUND

### A. Head Tracking

Different technologies have been used for head tracking. Anson et al [3], compared three head tracking systems using either ultrasonic, infrared or gyroscopic sensors. The first device, the HeadMaster Plus device was an ultrasonic-based system and used a light headset containing four microphones and a control box, which sends ultrasonic sound towards the user. The microphones are used to triangulate the source of the sonic sound. The Tracker 2000 was the second device. It used an infrared camera and required that the users wear a reflective dot on their head. The third device was Tracer. The Tracer system consisted of a solid-state gyroscope, which was contained within a headset. The gyro tracked the rotational movement of the head which was then converted to a cursor movement on the screen. Recently, a number of vision-based assistive devices have been reported ([4]–[8]). Shan et al [8] attribute this to the fact that personal computers have become more powerful and also camera system for computers are nowadays readily available, and this has contributed to advances in computer vision. CameraMouse [6] is a vision based tracking system, which can track features on the user's body. The face or facial features are mostly commonly used and the users have to select the feature they want to track. In Cloud et al [4], the performance of CameraMouse was evaluated for different users, features tracked (nose, lower lip, interior of left eye and the thumb) and with different applications such as BlinkLink [8]. It was observed that the nose, being the most prominent feature of the face, was the most effective feature for tracking. Pereira et al [7], proposed a head tracking system that uses a low cost infrared camera and the user must wear a reflective dot. The effectiveness of the head tracker was tested using Fitts Law [2] as recommended by the ISO 9431-9:2000 [9]. In a study conducted by Ashdown et al [5], head movement is used to enable a user to switch between multiple monitors. The proposed method used three cameras - two on top of the monitors and one near the keyboard. The proposed algorithm generated a 3-D model of the user's head. It was found that the mean distance moved by the cursor using the head in pixels was reduced by 32% i.e. it requires less effort to move the screen cursor using the head tracker than a standard mouse. However, it was also seen that the mean time to complete a task increased by 24% and that it was more difficult to move the pointer between points that are close together.

<sup>1</sup>S.P. Guiness, F.Deravi and K. Sirlantzis are with Faculty of Science, School of Digital Arts and Engineering, University of Kent, Canterbury, Kent, CT2 7NZ, United Kingdom {spg23,f.deravi,k.sirlantzis} at kent.ac.uk

<sup>2</sup>M.G Pepper and M. Sakel are with the East Kent University Hospitals Trust, Canterbury, Kent, United Kingdom {Matthew.Pepper, Mohamed.Sakel} at ekht.nhs.uk

## B. Fitts Test

Fitts test [2] was developed in 1954 to model human movement. The result of the experiments showed that the rate of performance of the human motor system is approximately constant over a wide range of movement amplitudes. Mackenzie et al [10], adapted the Fitts Law for assessing Human Computer Interfaces (HCI). This work was later embedded in an International Standard for HCI, ISO 9431-9:2000 [9] providing guidelines for measuring the users performance, comfort and effort. Performance of the device was measured by making the user perform tasks using the device. There are six types of tasks - one-direction, multi-directional, dragging, free-hand tracing (drawing), and, hand input, grasp and park (homing/device switching). ISO 9431-9:2000 [9] requires that the input device be tested for at least 2 different Index of Difficulty (ID). Index of Difficulty (ID) is a measure of the difficulty of the task [11] and is evaluated using (3). In this paper, Douglas et al [12], investigated the validity and practicality of the ISO framework using both multi-directional and the one-direction Fitts Tests for two devices namely a touchpad and a joystick.

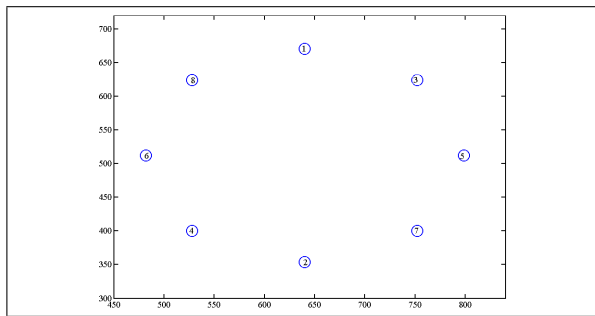


Fig. 1. Example of a Multi-Directional Fitts' Test

## C. Experiment Design

Anson et al [3], conducted an investigation to compare the efficiency of three commercially available head tracking devices. To test the efficiency of these devices, each participant was asked to produce a series of drawings using each device. The drawings were presented to each participant in a predetermined order. To prevent the participants from memorizing the drawings, the order of using the different devices was balanced to control for learning effect. For each participant, the time taken to draw and the number of errors were recorded. It was found that the HeadMaster Plus produced the most consistency. According to the feedback provided by the participants the two fastest devices i.e. the HeadMaster Plus and Tracer were the most uncomfortable to use and the Tracker 2000 system was the one preferred by most of the participants. In a study conducted by Williams et al [13], the efficiency of three different alternate pointing devices was compared. The devices used in the study were a standard mouse as a base line, a head orientation system and an EMG based system. Fitts Test was used to measure the performance and the throughput of the devices. In the study, it was found that with the head orientation system the movement both to

a single direction targets (i.e. targets at the cardinal points - North, South, East and West directions) and in diagonal movements were much better than with the EMG system. In the case where an EMG-system was used the diagonal movements followed a square shape path indicating that the participants were moving the cursor first in one direction and then the other, rather than simultaneously controlling the two cursor directions. In Pereira et al [7], a head tracker was developed, and evaluated with ten individuals with cervical spinal cord injury. The device developed used an infrared camera and a reflective dot. The evaluation involved a Multi-Directional Fitts Test with 16 different orientations. The participants had to undergo 12 sequence of test for two ID 2 bits and 5 bits. The results for the 12 attempts showed that the mean throughput movement time for ID 2 bits and the mean movement time for ID 5 bits of the proposed system were  $0.75 \pm 0.12$  bits/second,  $3.02 \pm 0.44$  bits/seconds and  $5.77 \pm 1.12$  bits/seconds respectively. There was also no significant difference in the mean movement time of the first and last attempts. From the results, it was seen that the device was easy to use and the movement time showed that the device adequately emulated the movement of a mouse.

## III. METHODOLOGY

The performance measurement is carried out by using a modified Fitts Test. The Fitts Test used is based on the Multi-Directional task with circular targets. Figure 1, displays potential target locations that could be used. The red circle represents the Home location. The Home location is placed centrally as it is assumed to be the default location where users would position their head pointer. Using the Home location approach in the Fitts Test would double the number of data collected i.e. 80 (16x5) data points as opposed to 40(8x5) data points with an eight point Multi-Directional task and thus increase the accuracy of the data being captured. The targets are presented randomly to the participants by our application. The participants have to move the cursor to the central home location and once the dwell click of 500ms is registered, a target location is displayed to the participant. The participant has to once again move the cursor to the location of the target and perform a dwell click. Eight different targets, each at a different orientation for each target size (W) and distance (D) are displayed to the participant.

Figure 3 is an illustrated example of the function of the Home location, the red circle represents the Home Location and the blue circles are examples of 2 different target locations. As it can be seen, one of the targets has a width (W) of 25 pixels and is at a distance (D) of 50 pixels from the center of the home to the center of the target location. The other target is of width (W) 50 pixels and at a distance (D) of 400 pixels from the home location. The measurements required to calculate the efficiency or performance of a device is described in [12].

Equation (1) is used to calculate the Index of Difficulty (ID) of a given task. In our case, the task is moving the

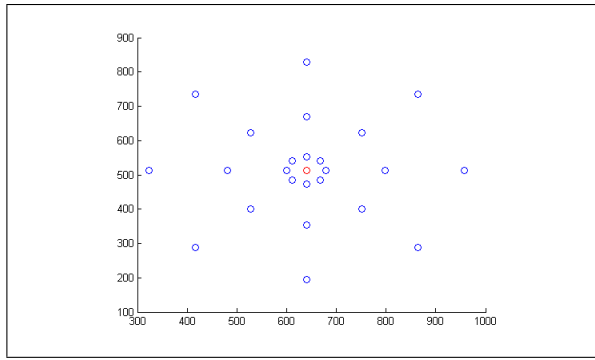


Fig. 2. Target locations

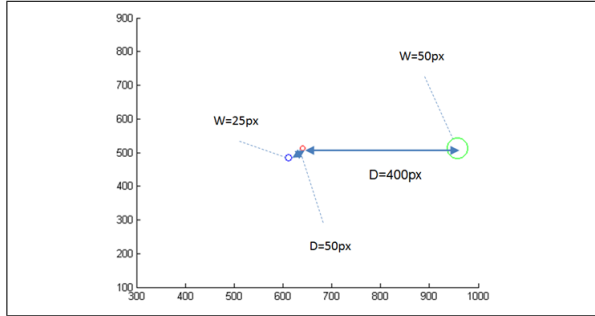


Fig. 3. Example target locations

cursor across the screen to complete the Fitts' test.

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

where  $D$  is the distance from the Home location to the target and  $W$  is the width of the target.

Equation (2) is used to calculate the effective Throughput ( $TP_e$ ) in bits/second.

$$TP_e = \frac{ID_e}{MT} \quad (2)$$

where  $MT$  is the mean movement time, in seconds, for all trials within the same condition, and

$$ID_e = \log_2\left(\frac{D}{W_e} + 1\right) \quad (3)$$

$ID_e$ , is the effective index of difficulty, in bits, and is calculated from the distance ( $D$ ) from the home to the target and  $W_e$ , the effective width of the target.  $W_e$ , is calculated from the observed distribution of selection coordinates in the test carried out.

$$W_e = 4.133 \times SD \quad (4)$$

where  $SD$  is the standard deviation of the selection coordinates.  $W_e$  is the effective width and using effective width incorporates the variability observed of human performance and includes both speed and accuracy [12].

#### IV. EXPERIMENTATION

##### A. Set up

For the experiment, each sensor for each of the devices being evaluated has to be mounted centrally on the monitor.

SmartNav has its own sensor but HeadTracker and CameraMouse, both use a standard web-camera. The participants are placed at a distance of 30-60 cm from the screen.

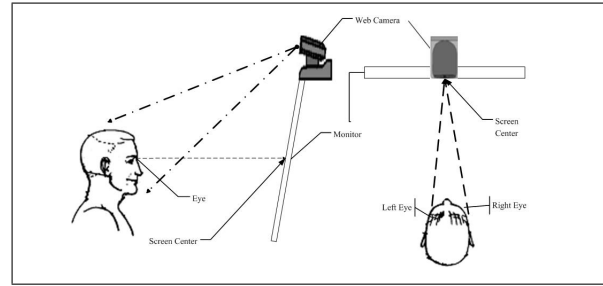


Fig. 4. Equipment setup

##### B. Fitts Test

Fitts Test is conducted in blocks of 5 sequences with 8 different orientations for each device as shown in Figure 4. Both the orientation and difficulty index of the target within each block are randomly selected. The approach is used to prevent the users from anticipating the location of the next target and thus ensure the independence between subsequent target selections. There is a rest period of 3 seconds (3000 milliseconds) between each block of test. The rest period is to reduce the fatigue a user might experience during a test and to prevent the current test from influencing the following one.

TABLE I  
PARAMETERS FOR FITTS' TEST

Width (W)	Distance (D)	Index of Difficulty(ID)
25	50	1.585
50	200	2.322
25	400	4.088
50	400	3.170
13	400	4.990

#### V. RESULTS

TABLE II  
MEAN MOVEMENT TIME

Index of Difficulty	1.585	2.322	3.170	4.088	4.990
Mouse	0.980	1.085	1.337	1.457	1.525
SmartNav	1.933	1.566	2.002	2.714	3.584
CameraMouse	1.517	1.677	2.238	2.699	4.271
HeadTracker	4.596	3.302	4.681	6.880	12.418

Figure 5, represents a coarse view of the performance of an individual using the different assistive devices. Using the coarse view, the devices with the lowest Movement Time (MT) are selected. The selected devices are CameraMouse and SmartNav as the regression lines obtained for these devices are consistently below the line obtained for the HeadTracker. The data from the standard mouse is also added to the fine grain view as a base line.

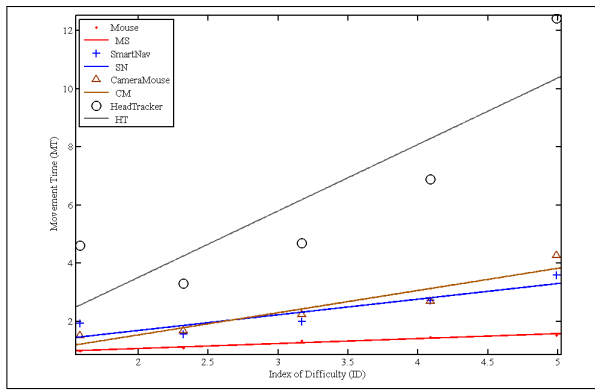


Fig. 5. Coarse Grain view of the Comparison between devices

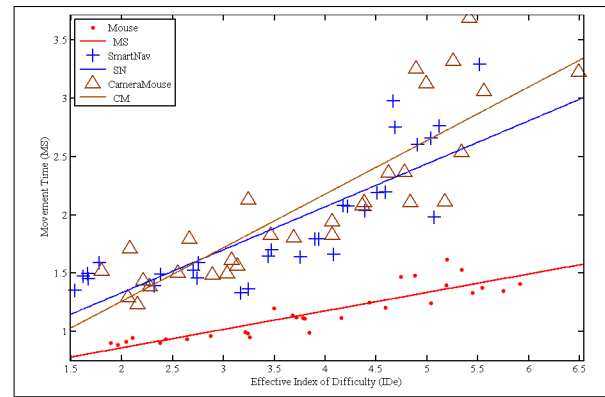


Fig. 6. Fine Grain view of the Comparison between devices

For example using SmartNav, the user would be able to set the resolution of their screen so as their icons/buttons are 25 pixels wide and spaced by 50 pixels would correspond to an ID of 1.585 bits and from Figure 5, it can be seen that the CameraMouse performs better than SmartNav. Similarly, for selecting text, having a width and height of 13 pixels from a document along a distance of 400 pixels would correspond to an ID of 4.990 bits and it can be seen from Figure 5 that SmartNav performs better than CameraMouse.

The fine view, in Figure 6 shows the comparison of CameraMouse and SmartNav using  $ID_e$  so as to get the effective performance of the user. It can be seen that SmartNav is better than CameraMouse for large Index of Difficulties. The Fitts' Test illustrate the performance of the user using each of the devices being evaluated. The best device for the user is the one which takes the least Movement Time to complete a task and from Figure 6, it can be seen that SmartNav is the most suitable device. this is so because although CameraMouse performances better than SmartNav for easy tasks, SmartNav is much better for the more difficult tasks. For example as it can be seen from Figure 6, for a task with an  $ID_e$  of 2 bits, CameraMouse would be the most appropriate device. However, for a task with  $ID_e$  6 bits as it can be seen in Figure 6, SmartNav would be the most suitable one. The test could also be carried out at regular intervals with the same devices to monitor the performance of the user. Any change in performance could be due to a change in the condition of the user.

## VI. CONCLUSIONS

Assistive devices need to fine tune to the residual abilities of disabled individuals. The modified Fitts' Test with the Home location and the randomness of the target selection increases the accuracy of the data collected. The Fitts Test may be used to determine the optimum device for an individual and the best settings for the screen of the individual. This would reduce the time to find a well suited device for an individual, increase usability of the individual and increase adherence by patients/carers. The test could also be used to monitor the progress of the user.

## REFERENCES

- [1] WHO, *World report on disability*. World Health Organization, Geneva : WHO, 2011.
- [2] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement. 1954.," *Journal of experimental psychology. General*, vol. 121, pp. 262–9, Sept. 1992.
- [3] D. Anson, G. Lawler, A. Kissinger, M. Timko, J. Tuminski, and B. Drew, "Efficacy of Three Head-Pointing Devices for a Mouse Emulation Task," *Assistive Technology*, vol. 14, no. 2, pp. 140–150, 2002.
- [4] R. Cloud, M. Betke, and J. Gips, "Experimentation with a Camera-Based Human-Computer Interface System," in *Proceedings of the 7th ERCIM Workshop "User Interfaces for All," UI4ALL 2002*, pp. 103–110, 2002.
- [5] M. Ashdown, K. Oka, and Y. Sato, "Combining head tracking and mouse input for a GUI on multiple monitors," in *CHI '05 extended abstracts on Human factors in computing systems*, vol. Portland, of *CHI EA '05*, (New York, NY, USA), pp. 1188–1191, ACM, 2005.
- [6] W. Akram, L. Tiberii, and M. Betke, "A Customizable Camera-Based Human Computer Interaction System Allowing People with Disabilities Autonomous Hands-Free Navigation of Multiple Computing Tasks," in *Proceedings of the 9th conference on User interfaces for all*, vol. 246 of *ERCIM'06*, (Berlin, Heidelberg), pp. 28–42, Springer-Verlag, 2006.
- [7] C. A. M. Pereira, R. Bolliger Neto, A. C. Reynaldo, M. C. D. M. Luzo, and R. P. Oliveira, "Development and evaluation of a head-controlled human-computer interface with mouse-like functions for physically disabled users.," *Clinics (São Paulo, Brazil)*, vol. 64, pp. 975–81, Jan. 2009.
- [8] C. Shan, R. Braspenning, and M. Betke, "Intelligent Interface to empower people with disabilities," *Intelligence*, pp. 409–432, 2010.
- [9] ISO TC 159/SC 4, *ISO 9241-9:2000, Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices*. International Organization for Standardization, 2002.
- [10] I. S. MacKenzie, "Fitts' law as a research and design tool in human-computer interaction," *Human-Computer Interaction*, vol. 7, pp. 91–139, Mar. 1992.
- [11] J. Accot, "Beyond Fitts' law: models for trajectory-based HCI tasks," *Proceedings of the SIGCHI conference on Human*, 1997.
- [12] S. A. Douglas, A. E. Kirkpatrick, and I. S. MacKenzie, "Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard," in *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, pp. 215–222, ACM, 1999.
- [13] M. R. Williams and R. F. Kirsch, "Evaluation of Head Orientation and Neck Muscle EMG Signals as Command Inputs to a Human-Computer Interface for Individuals With High Tetraplegia," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 16, no. 5, pp. 485–496, 2008.