Effects of age and disability on tracking tasks with a computer mouse: Accuracy and linearity

Cameron N. Riviere, PhD and Nitish V. Thakor, PhD
Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218; Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205

Abstract—Many individuals with movement disorders are unable to make efficient use of graphical computer interfaces commonly employed in personal computers. In this study, performance limitations in target tracking with a computer mouse were studied for eight young subjects aged 19 to 29 (M=23, SD=3), four old subjects aged 70 to 73 (M=72, SD=1), and five motor-disabled subjects aged 37 to 74 (M=65, SD=16). Subjects tracked simple one- and two-dimensional motions at various frequencies. Performance was measured using an accuracy index derived from root-mean-square error, and a linearity index based on coherence estimation. A maximum bandwidth of 2 Hz for accuracy of mouse use was found, which often decreased due to advanced age or motor disability. Tracking linearity of all groups decreased as frequency increased. A significant degree of nonlinearity existed in all results (p<0.05), with disabled subjects nearing complete nonlinearity in two-dimensional tracking. The data show that with advanced age and disability, mouse use becomes increasingly inaccurate and nonlinear. Assistive computer interfacing techniques, such as signal filtering, may improve mouse use.

Key words: human-machine systems, movement, myoclonus, tremor.

INTRODUCTION

Graphical user interfaces such as Microsoft Windows (Microsoft, Redmond, WA) are now the standard in personal computers (1), and successful use of a mouse or other pointing device is crucial to the rehabilitation and functional independence of many persons with movement disorders (2,3). Some severely disabled persons rely on pointing devices for personal communication (4). Computer mice and other pointing devices are used for drafting and other vocational rehabilitation applications at the Maryland Rehabilitation Center in Baltimore, MD, and at other centers nationwide (5). For persons with extremely poor manual control, commercial alternatives to mice include the Liaison (DU-IT, Shreve, OH), a joystick-like device for the chin; the HeadMaster Plus (Prentke-Romich, Wooster, OH), a sonar headset device; and the HeadMouse (Origin Instrument, Grand Prairie, TX), which uses a reflective disk mounted on the head. There has been much research in this area, including electroencephalogram (EEG)-based cursor controllers (6), video head motion detection systems (7), and electro-oculographic control (8). However, hand-operated interfaces are most natural for humans, and are preferable, provided reliable command information can be obtained from them.

In the workplace, productivity is essential if disabled workers are to be competitive. To obtain this needed productivity, a worker must be both accurate and fast. In any human task, there is a “tradeoff” between accuracy and speed: faster movement tends to be less accurate, and accurate motion tends to be slow (9). To measure the effectiveness of disabled persons in the use of pointing devices for vocational rehabilitation, it is therefore useful to consider the bandwidth, that is,
how fast a user can move while maintaining the needed accuracy. This is significant because the human body sets the information-processing limit in modern human-machine control (10).

Previous researchers (10–12) have thoroughly investigated the effective bandwidth for nondisabled human control. The purpose of this work is to determine how the bandwidth may change due to advanced age or movement disorders, and to infer from linearity results how the bandwidth might be increased. A mouse is used for these target tracking tasks in order to investigate the effectiveness of human subjects specifically using this most popular of commercially available computer interfaces (2). It is expected that young adult groups will exhibit the best performance, and are used as a baseline for comparison.

The human motor control system has been studied quantitatively by a great number of researchers, particularly through the use of manual target tracking tasks (9–21), including multiple-axis tests (17,21). The methods of control and information theory, in particular, have been widely used to develop linear and nonlinear models (9,17). In contrast, the body of research quantifying tracking performance of aged and disabled subjects is smaller (18,22–26).

Behbehani et al. (18) tested subjects by presenting them with step signals, or jumps, from one location on a video screen to another, and measured their ability to follow. They found that measures such as movement speed, reaction time, and a coordination index derived from phase plane analysis, showed significant differences between normal and Parkinson’s disease populations. Frith et al. (25) also found that patients with Parkinson’s disease performed significantly worse in target tracking tasks than healthy subjects, as did Flowers (22), who found no improvement with training in Parkinsonian subjects. Yamashita and Sameshima (26) tested tracking of a sine wave at the elbow by adults with cerebral palsy (CP). The subjects with CP had error levels similar to nondisabled subjects at frequencies below 0.56 Hz, but their error was significantly higher than the nondisabled subjects at higher frequencies. This suggests that the subjects with CP may have a decreased bandwidth, rather than simply experiencing an overall decrease in movement accuracy.

In human-machine control, the human is in general a nonlinear element (17). That is, in addition to the linear portion of the human response, there is a component of motion which is not a linear function of the target signal. Therefore, in addition to such established measurement techniques as root-mean-square (RMS) error (11,17,22,24), it is useful to have a measurement of tracking linearity, that is, the extent to which the human tracking responses signal can be represented as a linear function of the original target signal. Linearity of output, with respect to the input, is of interest because it greatly simplifies the design of control system components such as filters (27). Recent research has demonstrated that conditions such as essential tremor affect motion in ways other than simple additive oscillations, some of which are not fully understood (28,29). Furthermore, the nonlinear component of normal human tracking response generally contains frequencies within the target signal frequency band (13). Because of this, a measure of linearity aids in predicting the probable effectiveness of linear filtering to extract the target signal from the human output. We quantify mouse-based target tracking performance using two measures: an Accuracy Index (AI) derived from RMS error, and a novel Linearity Index (LI), based on adaptive coherence estimation.

**METHODS**

**Subjects**

Three groups of subjects participated in these experiments. The “young” group consisted of 8 healthy subjects (1 female, 7 male), aged 19 to 29 years (M=23 years, SD=3 years). The “old” group consisted of 4 healthy subjects (2 female, 2 male), aged 70 to 74 years (M=72 years, SD=1 year). The “disabled” group consisted of 5 subjects (2 female, 3 male), aged 37 to 74 years (M=65 years, SD=16 years). Table 1 summarizes the medical conditions of the disabled group. Only two members of the disabled group completed the two-dimensional experiment. All subjects gave their written consent to the experiments.

Some subjects had corrected vision, and others did not. With the old group, this study looks at the diffuse effects of aging on the human system. This work does not attempt to identify specific conditions of the subjects in the old group which may affect performance.

**Equipment**

An IBM-compatible personal computer, equipped with a Microsoft Serial Mouse and Microsoft Mouse Driver 9.0, was used for these tests. In the standard DOS environment, the sampling rate of the mouse was not constant. The Turbo C software for this work was
Table 1.
Disabled group subject information.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age</th>
<th>Gender</th>
<th>Diagnosis</th>
<th>Completed 2-D test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>f</td>
<td>anoxic myoclonus</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>74</td>
<td>f</td>
<td>rubral tremor</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>70</td>
<td>m</td>
<td>essential tremor</td>
<td>no</td>
</tr>
<tr>
<td>4</td>
<td>73</td>
<td>m</td>
<td>essential tremor</td>
<td>no</td>
</tr>
<tr>
<td>5</td>
<td>73</td>
<td>m</td>
<td>essential tremor</td>
<td>no</td>
</tr>
</tbody>
</table>

written so that constant sampling at 19 Hz was achieved. This was the highest constant sampling rate obtainable by the present software, and was judged acceptable because evaluation of the test data showed no frequency content above 9.5 Hz.

Procedure

Computer mice and other pointing devices sense motion in two spatial dimensions: horizontal (x) and vertical (y). To evaluate mouse-based performance by subjects, a target moving in two dimensions was therefore used. A battery of one-dimensional (sine) tracking tests was also carried out, to quantify the bandwidth at the simplest possible level.

One-dimensional (1-D) Experiment

Figure 1a shows a diagram of the computer screen for the 1-D experiment. A stationary vertical line segment was displayed on the computer screen. The length was approximately 3.9 cm on the computer screen (50 VGA pixels), corresponding to 2.6 cm of mouse displacement. A small square target then oscillated along the right side of the line segment in sinusoidal fashion. The subject tracked the target's motion with a small round mouse cursor, which moved along the left side of the line. The software was written such that the x component of the mouse signal was disregarded by the system, and only the y component was sampled and transmitted to the computer screen. The mouse cursor was thus constrained to move vertically, and was insensitive to sideways mouse movement. The young group and old group were tested at 8 frequencies (0.4 Hz to 4.6 Hz); the disabled group was tested at 5 frequencies (0.2 Hz to 2.0 Hz). Samples of raw data from the three groups are shown in Figure 2a.

Two-dimensional (2-D) Experiment

The computer screen for the 2-D tests was as shown in Figure 1b. A stationary circle was displayed on the screen, and the block-like target moved around the circle as the subject tracked it with the mouse cursor. The circle diameter was approximately 3.9 cm on the computer screen (50 VGA pixels), corresponding to 2.6 cm of mouse displacement. The x and y (horizontal and vertical) velocity components of the target were therefore both sinusoidal. Subjects were tested at 4 frequencies (0.4 Hz to 2.6 Hz). Samples of raw data from the three groups are shown in Figure 2b.

In a separate experiment, one young nondisabled subject and one elderly subject with essential tremor were each given 11 minutes of training on the 1-D task. After the first minute, no performance improvement was found. Each subject included here was therefore given one minute of training before each of the two experiments. For ease of learning, the lowest test frequency was used for training. The subjects in each group were relatively unskilled volunteers. It is likely that extended
training periods would yield performance improvements, as with Pew et al. (13), who trained subjects for 32 days. Each subject completed the 1-D experiment before the 2-D experiment. In all cases, data were recorded for 17 cycles of the target. To inform the subject beforehand of the test frequency, a sample cycle was displayed at the start of each test. The order of trials within each experiment was randomized. The distance from the screen to the subject’s face was approximately 65 cm.

In order to evaluate the bandwidth, it was necessary to obtain a frequency response (11), so sinusoidal target signals were used. One drawback of this approach is that mouse users are often not tracking a continuous signal, but attempting to acquire a discrete target. Furthermore, there is no explicit timing required in everyday mouse use. However, timing is implicitly introduced by the goal of productivity, thus requiring speed, which depends on bandwidth. The stationary line and circle displayed during the 1-D and 2-D tests, respectively, were provided in order to mitigate this drawback by providing advance information of the correct mouse path. In this way, the continuous tracking task is viewed, loosely, as a series of target acquisitions, separated by small time increments.

### Analytical Methods

Stark et al. (15) found that humans presented with a sinusoidal stimulus for tracking exhibited a transient response lasting between 0.5 s and 1 s, after which a steady state was typically reached. In all data analysis, the initial 1 s of data was therefore disregarded. Student $t$ tests were used to determine whether Accuracy Index results were significantly greater than zero, and to determine significance of differences between means among the three subject groups.

#### Accuracy Index ($AI$)

The position vector of the target on the computer screen was the input to the human sensorimotor system in these tests, referred to as the “target” and represented as $t$. The motion that the human subject made in response to the target was indicated by the mouse cursor location. This was considered the output, $o$, of the system. The error vector $e$ was obtained by the equation

$$ e = t - o. $$

For the 2-D tests, each of these vectors had two components. Thus,

$$ e = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}, \quad o = \begin{bmatrix} o_1 \\ o_2 \end{bmatrix}. $$

For the 1-D tests, these three are only one-dimensional vectors:

$$ e = [e_1], \quad t = [t_1], \quad o = [o_1]. $$

For each test, $e$ and $o$, the RMS values of the error vector $e$ and the target signal $t$, respectively, were calculated for each test. The subjects’ overall accuracy in tracking was represented by a measure called the Accuracy Index ($AI$), defined as

$$ AI = 1 - \sqrt{\frac{1}{N} \sum_{i=1}^{N} \frac{e_i^2}{o_i^2}}. $$

where $N$ was the vector dimension for the specific test. In the special case of the 1-D tests, there was no $x$-component, and the $AI$ could be simplified as

$$ AI = 1 - \frac{e_1}{o_1}. $$
For all tests, perfectly accurate motion generation resulted in zero error and, therefore, in an AI of unity. Leaving the cursor unmoved in the center of the screen resulted in a value of zero for AI.

The zero-crossing frequency for each group’s AI results was defined as the lowest frequency for which the AI was not significantly greater than zero (p<0.05). A negative AI value represented worse performance than would be attained by not tracking at all. Thus, in the absence of any signal filtering or other method to augment the human-machine interface, a zero AI would indicate that the signal the system was attempting to measure was no larger than the interference or noise. This is not to say that the tracking response contained no useful information; it was still possible that such a signal contained good information which could be extracted by standard methods such as filtering. In its raw state, however, the tracking response was no closer to the desired signal than would be the “null” signal resulting from no tracking at all (i.e., an untouched mouse). The zero-crossing frequency was therefore taken to indicate the bandwidth. A similar approach has been used by previous researchers (11).

**Linearity Index (LI)**

Coherence has been defined as follows:

\[ \gamma_{ij}(f) = \frac{G_{ij}(f)}{\sqrt{G_{ij}(f)G_{jj}(f)}} \]

where \(G_{ij}(f)\) is a cross-power spectral density of signals \(t_j\) and \(o_j\), and \(G_{jj}(f)\) and \(G_{ij}(f)\) are auto-power spectral densities. Coherence can be used as a measure of the linearity of the relation between two signals. A vector of coherence values for the two tracking dimensions is

\[ \gamma = [\gamma_1, \gamma_2] \]

using the convention \(\gamma_j = \gamma_{t, o_j}\), for \(j = 1, 2\). To evaluate tracking performance, a 2-D LI was developed, based on a 1-D version developed by Thakor et al. (30), that robustly indicated whether \(t\) and \(o\) were linearly related:

\[ LI = \frac{3}{2} \left( \frac{1}{1 + \sum_{j=1}^{M} 1 - \sqrt{\sum_{j=1}^{N} \gamma_j^2(f)}} - \frac{1}{3} \right), \]

where \(\hat{\gamma}\) was a numerical estimate of the magnitude of \(\gamma\), \(N\) is the dimension of the test (1 or 2), and \(M\) is the number of harmonics of \(t\) included (\(M=2\) here, due to the sampling rate). The definition given reflects the normalization of the LI to the range \([0,1]\). An LI value of 1 indicated a perfect linear system. LI values less than 1 showed the presence of a nonlinear component, with a proportional decrease in the linear component of the input-output relationship. An LI value of 0 indicated complete nonlinearity, that is, no linear component was present. LI results were computed for every four target cycles, and averaged over 12 target cycles in each test. The coherence magnitude was estimated using a least-squares technique. Impulse responses transforming \(t_j\) to \(o_j\), and vice versa, were estimated, and these estimates were improved using the least mean square (LMS) algorithm. The Fourier transforms of the impulse responses were then multiplied together to obtain \(\hat{\gamma}\). For more details on the technique, see Thakor et al. (30).

**RESULTS**

**Accuracy Index**

Figure 3 shows the AI results for the 1-D experiment. The young group was significantly more accurate (p<0.05) than the disabled group at all points (0.4-2.0 Hz). The old group performed significantly better (p<0.05) than the disabled group at 1 Hz, and significantly worse than the young group at 1.7 Hz and 2.0 Hz.

Figure 4 presents AI results for the 2-D tests. The young group exhibited the same decreasing trend as in Figure 3a. The young group performed significantly better (p<0.05) than the disabled group at all frequencies except the highest (2.6 Hz), and significantly better (p<0.05) than the old group at 0.4 Hz and 1.7 Hz. Differences between the old and disabled group 2-D results were not significant.

For the 1-D tests, the AI for the young group was significantly greater than zero (p<0.05) at frequencies up through 1.7 Hz. The old group AI was significantly greater than zero (p<0.05) at 0.4 Hz and 1.0 Hz, and the disabled group was significantly greater than zero (p<0.05) only at 0.2 Hz. In the 2-D experiment, the young group AI was significantly greater than zero (p<0.05) at 0.4 Hz and 0.8 Hz, and the old group and disabled group were not significantly positive at any point. Differences between 1-D and 2-D AI were significant (p<0.05) at 1.7 Hz for the young group and
Figure 3.
Accuracy Index for 1-D tests. Peak-to-peak target sine amplitude = 2.6 cm. (a) Young group. (b) Old group. (c) Disabled group. (Points plotted indicate mean ± standard deviation for subject populations of n=8, 4, and 5 for young, old, and disabled groups, respectively.)

0.4 Hz for the disabled group. Table 2 summarizes the zero-crossing frequency results, which were used to infer the bandwidth. The zero-crossing frequency for the 2-D tests was less than or equal to that for the 1-D tests for each group.

Figure 4.
Accuracy Index for 2-D tests. Target circle diameter = 2.6 cm. (a) Young group. (b) Old group. (c) Disabled group. (Points plotted indicate mean ± standard deviation for subject populations of n=8, 4, and 2 for young, old, and disabled groups, respectively.)

Linearity Index
The LI results for the 1-D experiment are presented in Figure 5. The LI of the young group results was significantly higher (p<0.05) than the old group at target
Table 2.
Zero-crossing frequencies for AI.

<table>
<thead>
<tr>
<th>Group</th>
<th>Zero-crossing frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-D</td>
</tr>
<tr>
<td>Young</td>
<td>2.0</td>
</tr>
<tr>
<td>Old</td>
<td>1.7</td>
</tr>
<tr>
<td>Disabled</td>
<td>0.4</td>
</tr>
</tbody>
</table>

frequencies up through 1.7 Hz, and significantly higher (p<0.05) than the disabled group at all frequencies through 2.0 Hz. The old group was significantly more linear (p<0.05) than the disabled group only at 1.7 Hz and 2.0 Hz.

Figure 6 presents the LI for the 2-D tests. The young group was significantly more linear (p<0.05) than the disabled group at all points except 1.7 Hz, but is significantly more linear than the old group only at 1.7 Hz. The LI for the old group was significantly higher (p<0.05) than the disabled group only at 1.7 Hz. The LI for the disabled group was within one standard deviation of 0.5 at almost all frequencies (i.e., the tracking response was almost completely nonlinear.)

DISCUSSION

Obtaining the frequency response allows us to determine the bandwidth for the three subject groups. Judging from the results presented here, it appears that 2 Hz is the maximum value of the bandwidth for tracking tasks using a computer mouse, which may decrease due to factors such as movement disorders and advanced age. Decreases in bandwidth with age and motor disability may be large (Table 2), and may limit the performance of rehabilitative interfaces. Figure 7 demonstrates how decreases in bandwidth may be manifest in individual subjects. For a male subject of the old group (aged 73), Figures 7a and 7b show the zero-crossing frequencies to be 1.7 Hz for both the 1-D and 2-D cases. The second subject, a female with rubral tremor (aged 75), has 1-D and 2-D zero-crossing frequencies of 1.0 Hz and 0.4 Hz respectively. Zero-crossing frequencies for a female subject with myoclonus (aged 37) are 1.7 Hz for the 1-D case and 0.4 Hz for the 2-D case. Based on these 2-D results, these two sample disabled subjects might be expected to have great difficulty in pointing device use.

Good productivity in computer mouse use requires a high bandwidth. Linearity of pointing device commands is desirable for simplicity of interface design. The data confirm the expectation that the young group
Figure 6. Linearity Index for 2-D tests. Diameter = 2.6 cm. (a) Young group. (b) Old group. (c) Disabled group. (Points plotted indicate mean ± standard deviation for subject populations of n=8, 4, and 2 for young, old, and disabled groups, respectively.)

As expected, the disabled group yields the worst results for both AI and LI. As Table 2 shows, the highest zero-crossing frequency obtained is 2.0 Hz, for the young group in the 1-D test. The young group 2-D zero-crossing frequency is lower, as are all zero-crossings for the other groups, in which the 2-D AI is never greater than zero.

We would expect the maximum bandwidth to be no less than 1 Hz, since Mann et al. (31) report that the average predominant frequency for 12 everyday tasks is 1 Hz. Elkind and Sprague (12), using target signals with rectangular spectra, report that the signal-to-noise ratio for healthy human pursuit tracking decreased sharply when bandwidths above 1 Hz were used, and information rates at 2 Hz were half those at 1 Hz. We would therefore expect to find a significant decrease in motion accuracy occurring in the vicinity of 1 to 2 Hz, which is in keeping with the findings for the young group.

The LI results shed further light on the nature of the human tracking responses. Stark et al. (15) used
linear system theory to analyze healthy human tracking of a sum of three sines, which appeared random, justifying it by comparing target and output amplitudes. However, when tracking a single sine, as was done here to quantify bandwidth, Sheridan and Ferrell (32) have pointed out that linearity cannot reasonably be assumed, primarily because of human anticipation of the signal. Anticipation is also present in practical teleoperation (such as pointing device use), and though sine tracking is an unlikely use for a rehabilitative interface, it provides a basis for quantitative study, and it is felt that a sinusoidal target takes anticipation into account more than a pseudorandom one. The fact remains, however, that general mouse use typically involves target acquisition rather than tracking. Further study involving different tasks and larger populations is therefore necessary before these results can be extended conclusively to general mouse use. It seems likely that further study with larger populations may also smooth out the nonmonotonic results sometimes visible at higher test frequencies.

The output data from these experiments are significantly nonlinear, even for the two nondisabled groups, as Figures 5 and 6 show. Yamashita and Sameshima (26) reported sine-tracking coherence results no lower than 0.8 for nondisabled subjects at frequencies up to 1.77 Hz, and no lower than 0.5 for patients with cerebral palsy at frequencies up to 1.33 Hz. These same frequencies of sine tracking are approximately where the 1-D LI results presented here begin to drop sharply (Figure 5) for the young and disabled groups respectively.

The degree of nonlinearity present is significant in all data (p<0.05), even at frequencies as low as 0.2 Hz. The effectiveness of linear filtering for extraction of human intentions from teleoperative interfaces would therefore be limited by this nonlinearity. Rehabilitative teleoperation, from control of simple pointing devices to assistive robotic systems, may be improved by the modeling and removal of this nonlinear component. The data suggest that nonlinear filtering may increase the bandwidth for human-machine control; thus, providing more precise and efficient rehabilitative teleoperation.

CONCLUSION

The maximum bandwidth for mouse-based target tracking by human subjects is approximately 2 Hz. Factors such as advanced age and movement disorders reduce this value in the subjects tested. Bandwidth limitations on human accuracy may limit the performance of assistive computer interfaces for persons with movement disorders. Significant nonlinearity is present in all recorded tracking by healthy and movement-disabled subjects, and increases with tracking frequency. Factors such as advanced age and movement disorders may cause more pronounced nonlinearity.

ACKNOWLEDGMENTS

The authors thank the test subjects; Dr. S.G. Reich and Dr. H.F. Rizer for their cooperation in the human subject studies; and K. Ng and M. Fung for data analysis.

REFERENCES