GLIS 701: Comprehensive Examination

Performance Evaluation of Target Selection Tasks from Older Adults on Touch Screen Devices

Afroza Sultana

Introduction

Adults aged 65 years and older comprise one of the fastest growing populations, not only in Canada, but also in most of the developed countries in the world. Wey (2004) highlighted the importance of modern technology in the lives of older adults, and concluded that technology can help older adults to carry on their daily activities, independently and socially, both inside and outside of their houses, and especially, if they have age related declines in cognitive functionalities. With the emergence of wide range of useful applications for touch-based mobile phones and tablets (for example, applications for monitoring an individual's health related data, tracking an individual's location, scheduling, social networking, gaming, etc.), Wey's findings have become even more relevant today.

Statement of the Problem

Despite the potential for modern touch technologies, many of the older adults cannot take the full advantage of these technologies because of age related physical, sensory and cognitive declines (Walker, Philbin, & Fisk, 1997; Welford, 1981; Welford, 1988). Massimi, Baecker, and Wu (2007) indicated that mobile device manufacturers, in general, target younger adult users as their primary consumers. The device functionalities are primarily tuned to the speed and accuracy preferences of younger adult users. However, Walker et al. (1997) showed that older adult users have different speed-accuracy trade-off preferences than younger adult users. Furthermore, due to age related declines in motor functionalities, many older adult users face difficulties, especially, selecting a target (such as, a link or a button) on a touch screen interface (Motti, Vigouroux, & Gorce, 2013). Therefore, emphasizing only on the target selection behavior of younger adults may result in overlooking the accessibility issues for older users.

McCreadie and Tinker (2005) suggested that the willingness to accept any new technology depends on both need for, and accessibility of that technology. Therefore, it is clear that even though touch screen devices like mobile phones and tablets can provide useful services for older adults, if these devices are not accessible to them, adopting these devices will be difficult. To ensure the accessibility, there is a need to analyze the performance of older adult users when they interact with touch screen devices, particularly during a target selection task on the screen.

Context

General Frameworks of Performance Evaluation Studies

Performance evaluation studies for target selection tasks on computer screens have heavily relied on Fitts' model on channel capacity of human motor system for target acquisition tasks (Fitts, 1954; Fitts & Patterson, 1964). Fitts model suggests that bigger targets over smaller distances are easier to select (i.e., having lower index of difficulty) than smaller targets located far apart. In his empirical experiments, Fitts observed a strong correlation between task completion time and the difficulty of the selection tasks. This correlation is referred to as the speed-accuracy trade-offs for target selection tasks.

Fitts' model developed the fundamental framework for performance analysis studies for target selection tasks on computer screen. Card, English and Burr (1978) applied Fitts' model for the first time to evaluate the performance of target selection task with mouse and isometric joysticks. Since then, Fitts' model has been improved and extended several times, and have been applied successfully to quantify the performances of target selection tasks with both indirect (e.g., mouse, joystick, etc.) and direct (e.g., pen-based, and touch-based) pointing input devices (Bi, Li & Zhai, 2013; Grossman & Balakrishnan, 2005; MacKenzie, 1992; Soukoreff & MacKenzie, 2004).

MacKenzie, Kauppinen, and Silfverberg (2001) enhanced Fitts' model by looking into the smaller sub-movements that make up the trajectory of a target selection task. Adopting the human psychomotor model for target selection, which breaks motion into a series of consecutive sub-movements (Meyer, Abrams, Kornblum, Wright & Smith, 1988), MacKenzie et al. (2001) introduced several new sub-movement measures to evaluate target selection tasks with mouse (see Figure 1). In addition to identifying poor performance, this new model provides explanations for performance degradation, by for example, linking degraded performance to factors such as excessive movement direction changes. Some performance analysis studies with motor-impaired users applied sub-movement model and resulted in very precise understanding of target selection behavior of such users (Hwang, Keates, Langdon, & Clarkson, 2005).

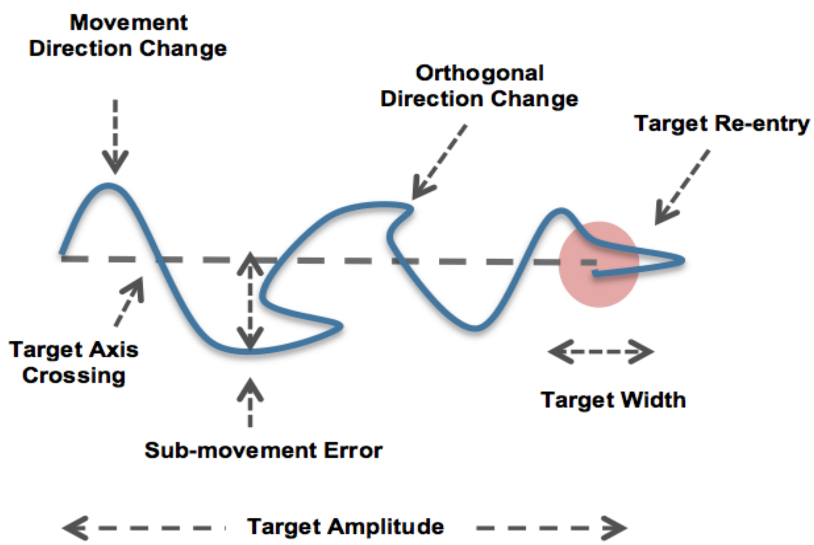

Figure 1. Sub-movement metrics to evaluate pointing performance (MacKenzie et al., 2001).

Performance Evaluation Studies for Direct Pointing Devices

Analysis of selection performance with direct pointing devices is underexplored relative to indirect devices, but similarly makes use of Fitts' model (Accot & Zhai, 2003; Grossman & Balakrishnan, 2005; Soukoreff & MacKenzie, 2004; Zhai, Kong & Ren, 2004). Sears and Shneiderman (1991) extended Fitts' model for touch screens by incorporating two touch-specific properties: (1) *homing time,* when a user initially places a finger on the screen, close to the target, before starting the target selection task, and (2) *target verification time*, the time taken by a user before physically selecting a target. Some studies concluded that the size, shape, angle, and position of the finger influence the target selection performance on touch screens, especially when the target size is very small (Cockburn, Ahlström & Gutwin, 2012; Holz & Baudisch, 2011; Zhai et al., 2004). Bi et al. (2013) improved Fitts' model for finger operated target selection tasks by augmenting a touch property in the model that is not bounded by the speed-accuracy trade-offs of Fitts' model, but rather depends on the properties of the finger.

Performance Evaluation Studies for Older Adults

Studies on target selection performance evaluation on older adult users reveal that they demonstrate significantly different target selection behavior than younger adult users, with both indirect and direct input devices. Firstly, older adult users face more difficulties than younger adults when selecting smaller targets (Hourcade & Berkel, 2008; Smith, Sharit, & Czaja, 1999). Secondly, they carry out a target selection task with a series of smaller sub-movements, having smaller pauses in between, rather than carrying out the task with a primary ballistic movement (Keates & Trewin, 2005; Ketcham & Stelmach, 2004; Walker et al., 1997). Thirdly, older adult users show high variability on target selection end points because of motor impairments (Paradise, Trewin & Keates, 2005; Rogers, Fisk, McLaughlin, & Pak, 2005; Walker et al., 1997). Fourthly, older adults demonstrated less precision in target selection, and require more corrective sub-movements after missing the targets (Smith et al. 1999). Finally, older adults emphasize on accuracy over speed, hence, requiring more time for target verification and selection than younger adult users (Findlater, Froehlich, Fattal, Wobbrock, & Dastyar, 2013; Keates & Trewin, 2005; Rogers et al., 2005; Salthouse, 1988; Smith et al., 1999). However, studies conducted on direct input devices suggest that pen-based and touch-based devices reduce the gaps between younger and older adult users target selection performance (Charness, Holley, Feddon, & Jastrzembski, 2004; Findlater et al., 2013; Rogers et al., 2005; Schneider, Wilkes, Grandt, & Schlick, 2008).

Methodology

Literature on pointing input device performance evaluation for target selection task shows that majority of such studies (both with older adults specifically, and users more generally) have relied on repetitive Fitts' tasks that are carried out in controlled laboratory experiments, often in a single session. Researchers have argued that, in real life, performance of the same user may vary significantly across different task sessions; hence data collected in one single lab session may capture only a smaller subset of the real word target selection behavior, especially with motor impaired users (Hurst, Mankoff, & Hudson, 2008), and such structured experiments may be unable to capture the contexts of the target selection tasks (Jansen, Findlater, & Wobbrock, 2011).

Unlike controlled lab experiments, in-situ data collection strategies (Brown, Reeves, & Sherwood, 2011) that are carried out in real life uncontrolled settings for a longer period of time, provide the flexibilities to capture both the variation and the context of target selection behavior. Recent performance evaluation studies, therefore, are slowly shifting towards collecting Fitts' tasks data from real life environment (Chapuis, Blanch & Beaudouin-Lafon, 2007; Evans & Wobbrock, 2012; Gajos, Reinecke & Herrmann, 2012; Hurst, Hudson, Mankoff & Trewin, 2013; Montague, Nicolau & Hanson; 2014; Weir, Rogers, Murray-Smith & Löchtefeld; 2012). However, most of these studies inherited the challenges of in-situ data, which include, inferring the user's intent for wrong target selection (i.e., whether the user aimed for the correct target but selected a wrong one, or mistakenly aimed for the wrong target), teasing out the environmental influences, and identifying the initiation and the termination points of a selection task.

Hurst, Hudson, and Mankoff (2007) argued that supervised machine-learning algorithm, such as the decision tree algorithm (Quinlan, 1986) can learn from pre-existing lab experiment data, and can classify lab quality data from the streams of in-situ data. Later, Gajos et al. (2012) applied a one-class classifier (Elkan & Noto, 2008) to isolate lab quality data from in-situ mouse data. Weir et al. (2012), and Montague et al. (2014) also gathered similar results applying Gaussian functions (Rasmussen & Williams, 2006) with direct input devices. One common trait of these studies is that they have applied very simple machine-learning algorithms for quality data prediction. Moreover, these in-situ studies relied upon assumptions and rough heuristics to infer the users' intent. For example, Evans & Wobbrock (2012) relied upon crowd-sourced workers to make visual assessments, and Montague et al. (2014) limited their study to a single online game (Sudoku) so that they could leverage characteristics of the game to identify errors. Such approaches only capture a subset of the user's errors, and thus, improved techniques are needed to fully understand the extent and nature of selection errors.

Significance of the research

Older adults now are more independent, educated, and tech-savvy than those of previous generations. Nonetheless, 40% of the older adult population cannot actively use modern technology in their daily lives due to a physical impairment (Pew Research Internet Project, 2014). A major reason behind this digital divide is that older adult needs have been overlooked in the design and evaluation of modern touch devices. Younger and older adults have significantly different target selection behavior on touch screen interfaces (Taveira & Choi, 2009). Therefore, assessing the performance of older adults from models developed for younger users may be insufficient. Specific performance evaluation models have been found to be helpful for understanding interacting pointing performance for motor impaired users (Hwang, Keates, Langdon, & Clarkson, 2005; Hwang, Langdon, Keates, Clarkson, & Robinson, 2002; Keates, Hwang, Langdon, & Clarkson, 2002; Keates & Trewin, 2005).

Furthermore, the existing literature of performance evaluation models for target selection tasks were developed for indirect input. These models, which are based on two-dimensional target selection activities, may not fully capture the three-dimensional movement behavior inherent to touch interaction. In addition, existing models do not address touch specific properties, such as angle and pressure of the finger during selection. Even though there are some extensions of Fitts' model for touch-based interactions, further exploration is required. Beyond basic selection, other gestures (e.g., swapping, dragging, and steering) should also be explored (Findlater et al., 2013; Wacharamanotham et al., 2011).

Last, but not the least, novel strategies for extracting quality data from in-situ touch interaction are needed. To date, only very simple algorithms have been used in combination with simple heuristics and assumptions to provide partial data. More research is needed to examine the feasibility of using more complex machine-learning algorithms to model multidimensional in-situ data. Considering the wide-range of machine-learning applications in other fields (Chandola, Banerjee & Kumar, 2009; Han, Kamber & Pei, 2006), we can anticipate that machine-learning algorithms can open new avenues of research in the field of performance evaluation of target selection tasks.

References:

Accot, J., & Zhai, S. (2003). Refining Fitts' law models for bivariate pointing. In *CHI 2003: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 193-200). New York, NY, USA: ACM.

Bi, X., Li, Y., & Zhai, S. (2013). FFitts law: modeling finger touch with fitts' law. In *CHI 2013: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 1363-1372). New York, NY, USA: ACM.

Brown, B., Reeves, S., & Sherwood, S. (2011). Into the wild: challenges and opportunities for field trial methods. In *CHI 2011: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 1657-1666). New York, NY, USA: ACM.

Card, S. K., English, W. K., & Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics, 21*(8), 601-613.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys (CSUR), 41*(3), 15.

Chapuis, O., Blanch, R., & Beaudouin-Lafon, M. (2007). Fitts' law in the wild: A field study of aimed movements. Laboratoire de Recherche en Informatique (LRI) Technical Report (1480), 1-11.

Charness, N., Holley, P., Feddon, J., & Jastrzembski, T. (2004). Light pen use and practice minimize age and hand performance differences in pointing tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(3), 373-384.

Cockburn, A., Ahlström, D., & Gutwin, C. (2012). Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *International Journal of Human-Computer Studies*, *70*(3), 218-233.

Elkan, C., & Noto, K. (2008). Learning classifiers from only positive and unlabeled data. In *KDD 2008*: *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (pp. 213-220). New York, NY, USA: ACM.

Evans, A., & Wobbrock, J. (2012). Taming wild behavior: the input observer for text entry and mouse pointing measures from everyday computer use. In *CHI 2012: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 1947-1956). New York, NY, USA: ACM.

Findlater, L., Froehlich, J. E., Fattal, K., Wobbrock, J. O., & Dastyar, T. (2013). Age-related differences in performance with touchscreens compared to traditional mouse input. In *CHI 2013: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 343-346). New York, NY, USA: ACM.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, *47*(6), 381-393.

Fitts, P. M., & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*(2), 103-112.

Gajos, K., Reinecke, K., & Herrmann, C. (2012). Accurate measurements of pointing performance from in situ observations. In *CHI 2012: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 3153-3166). New York, NY, USA: ACM.

Grossman, T., & Balakrishnan, R. (2005). A probabilistic approach to modeling two-dimensional pointing. *ACM Transactions on Computer-Human Interaction (TOCHI), 12*(3), 435-459.

Han, J., Kamber, M., & Pei, J. (2006). *Data mining: Concepts and techniques* (2nd ed.). San Francisco, CA, USA: Elsevier.

Holz, C., & Baudisch, P. (2011). Understanding touch. In *CHI 2011: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 2501-2510). New York, NY, USA: ACM.

Hourcade, J. P., & Berkel, T. R. (2008). Simple pen interaction performance of young and older adults using handheld computers. *Interacting with Computers*, *20*(1): 166–183.

Hurst, A., Hudson, S. E., & Mankoff, J. (2007). Dynamic detection of novice vs. skilled use without a task model. In *CHI 2007: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 271-280). New York, NY, USA: ACM.

Hurst, A., Hudson, S. E., Mankoff, J., & Trewin, S. (2013). Distinguishing Users By Pointing Performance in Laboratory and Real-World Tasks. *ACM Transactions on Accessible Computing (TACCESS), 5*(2), 5.

Hurst, A., Mankoff, J., & Hudson, S. E. (2008). Understanding pointing problems in real world computing environments. In *ASSETS '08: Proceedings of the SIGACCESS Conference on Computers and Accessibility*, (pp. 43-50). New York, NY, USA: ACM.

Hwang, F., Keates, S., Langdon, P., & Clarkson, J. (2005). A submovement analysis of cursor trajectories. *Behaviour & Information Technology*, *24*(3), 205-217.

Hwang, F., Langdon, P., Keates, S., Clarkson, P., & Robinson, P. (2002). Cursor characterisation and haptic interfaces for motion-impaired users. In S. Keates, P. Langdon, P. Clarkson, & P. Robinson, P. (Eds.), *Universal Access and Assistive Technology* (pp. 87-96), London, UK: Springer.

Jansen, A., Findlater, L., & Wobbrock, J. O. (2011). From the lab to the world: Lessons from extending a pointing technique for real-world use. In *CHI 2011: Extended Abstracts in SIGCHI Conference on Human Factors in Computing Systems*, (pp. 1867-1872). New York, NY, USA: ACM.

Keates, S. Hwang, F., Langdon, P., and Clarkson, P.J. (2002) Cursor measures for motion-impaired computer users. In *ASSETS '02: Proceedings of the SIGACCESS Conference on Computers and Accessibility*, (pp. 135-142). New York, NY, USA: ACM.

Keates, S., & Trewin, S. (2005). Effect of age and Parkinson's disease on cursor positioning using a mouse. In *ASSETS '05: Proceedings of the SIGACCESS Conference on Computers and Accessibility*, (pp. 68-75). New York, NY, USA: ACM.

Ketcham, C. J., & Stelmach, G. E. (2004). Movement control in the older adult. In R. W. Pew & S. B. V. Hemel (Eds.), *Technology for Adaptive Aging* (pp. 64-92). Washington, DC, USA: National Academies Press.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction, 7*(1), 91-139.

MacKenzie, I. S., Kauppinen, T., & Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. In *CHI 2001: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 9-16). New York, NY, USA: ACM.

Massimi, M., Baecker, R. M., & Wu, M. (2007). Using participatory activities with seniors to critique, build, and evaluate mobile phones. In *ASSETS '07: Proceedings of the SIGACCESS Conference on Computers and Accessibility*, (pp. 155-162). New York, NY, USA: ACM.

McCreadie, C., & Tinker, A. (2005). The acceptability of assistive technology to older people. *Ageing and Society, 25*(1), 91-110.

Meyer, D. E., Abrams, R. A., Kornblum, S., Wright, C. E., & Keith Smith, J. E. (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological review*, *95*(3), 340 – 370.

Montague, K., Nicolau, H., & Hanson, V. L. (2014). Motor-impaired touchscreen interactions in the wild. In *ASSETS 2014: Proceedings of the SIGACCESS Conference on Computers and Accessibility*, (pp. 123-130). New York, NY, USA: ACM.

Motti, L. G., Vigouroux, N., & Gorce, P. (2013). Interaction techniques for older adults using touchscreen devices: a literature review. In *IHM'13: Proceedings of the 25ème IEME Conférence Francophone sur l'Interaction Homme-Machine,* (pp. 125-134). New York, NY, USA: ACM.

Paradise, J., Trewin, S., & Keates, S. (2005). Using pointing devices: difficulties encountered and strategies employed. In *HCI International 2005: Proceedings of the 3rd International Conference on Universal Access in Human-Computer Interaction,* (pp. 22-27). Mahwah: Lawrence Erlbaum Association.

Pew Research Internet Project (2014). Older Adults and Technology Use. Retrieved from http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/.

Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81-106.

Rogers, W. A., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a screen or turn a knob: Choosing the best device for the job. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(2), 271-288.

Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. Cambridge: MIT Press.

Salthouse, T. A. (1988). Cognitive aspects of motor functioning. *Annals of the New York Academy of Sciences*, 515(1), 33-41.

Schneider, N., Wilkes, J., Grandt, M., & Schlick, C. M. (2008). Investigation of input devices for the age-differentiated design of human-computer interaction. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* (Vol. 52, No. 2, pp. 144-148). SAGE Publications.

Sears, A., & Shneiderman, B. (1991). High precision touchscreens: design strategies and comparisons with a mouse. *International Journal of Man-Machine Studies*, *34*(4), 593-613.

Smith, M. W., Sharit, J., & Czaja, S. J. (1999). Aging, motor control, and the performance of computer mouse tasks. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(3), 389-396.

Soukoreff, R. W., & MacKenzie, I. S. (2004). Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *International Journal of Human-Computer Studies, 61*(6), 751-789.

Taveira, A. D., & Choi, S. D. (2009). Review study of computer input devices and older users. *Intl. Journal of Human–Computer Interaction, 25*(5), 455-474.

Wacharamanotham, C., Hurtmanns, J., Mertens, A., Kronenbuerger, M., Schlick, C., & Borchers, J. (2011). Evaluating swabbing: a touchscreen input method for elderly users with tremor. In *CHI 2011: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (pp. 623-626). New York, NY, USA: ACM.

Walker, N., Philbin, D. A., & Fisk, A. D. (1997). Age-related differences in movement control: Adjusting submovement structure to optimize performance. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*(1), 40-52.

Weir, D., Rogers, S., Murray-Smith, R., & Löchtefeld, M. (2012). A user-specific machine learning approach for improving touch accuracy on mobile devices. In *UIST 2012: Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*, (pp. 465-476). New York, NY, USA: ACM.

Welford, A. T. (1981). Signal, noise, performance, and age. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *23*(1), 97–109.

Welford, A. T. (1988). Reaction time, speed of performance, and age. *Annals of the New York Academy of Sciences*, *515*(1), 1–17.

Wey, S. (2004). Use of Assistive Technology. In M. Marshall (Ed.), *Perspectives on Rehabilitation and Dementia* (pp. 202-208). London, UK: Jessica Kingsley Publisher.

Zhai, S., Kong, J., & Ren, X. (2004). Speed–accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies, 61*(6), 823-856.