# **MODELING HUMAN PERFORMANCE IN USING A SPATIAL SEGMENTATION CHINESE HANDWRITING RECOGNIZER**

Changxu Wu

Department of Industrial & Operations Engineering, University of Michigan Ann Arbor, Michigan Yongge Hu Intel China Research Center, Beijing, PR China

To predict and improve human performance in spatial segmentation Chinese handwriting recognizer, a mathematic performance model was developed based on Fitts' law and probability theory of discrete random variables. The model was verified by a behavioral experiment and it can predict subjects' task completion time well after the subjects experienced 2 stages of practice (R square  $> 0.9$ ). The mathematic model including its variants might be very helpful for designers of the user interface to select optimal parameters in layout of elements on the user interface and focus on relatively important and cost-effective factor(s) in the system to optimize the human performance. Further developments of the model in modeling human performance in using other input advices, and its value in developing proactive ergonomic design and analysis tools for input interface design are discussed.

# **INTRODUCTION**

 Nowadays pen computing and pen-based interface become one of major ways in human computer interaction, esp. in using personal digital assistance (PDA) and other mobile computing systems (Mackenzie & Chang, 1999; Davis et al., 1998; Frankish et al., 1995). Moreover, for more than 1.5 billion people in the world using graphic characters, e.g. Chinese, Japanese, and Korean, inputting graphical characters into computer remains a major impediment in human– computer interaction (Wu et al., 2003; Sarcher et al., 2001; Cheng, 1996). The development of handwriting recognition system (see Figure 1 as an example) has shed lights on the solution to this bottleneck (Sarcher et al., 2001) because they can save users' time in mentally mapping graphical characters onto keyboards, reduce mental workload, and offer direct manipulation of characters on user interface (Zhang, 1992; Wu et al., 2003).



 Figure 1. A mobile phone with a Chinese character handwriting recognizer. Adapted from Wu (2003).

 However, most of the previous human performance modeling studies were focused on text entry with standard keyboards and use English as inputting language (John, 1989; Wu et al., 2004 a, b). Very few human–computer interaction researches including human performance modeling, studied the human performance in using these graphical character handwriting recognizers.

 All of the handwriting recognizers can be categorized into two kinds: temporary and spatial segmentation depending on how the system separates different characters. A temporary segmentation handwriting recognizer only has one handwriting window. Hence, user has to write a character in the window, pause several hundred milliseconds waiting for the system to recognize the character, and then write the next character in the same window. Recently, Wu et al. (2003) published the first work in modeling human performance in Chinese temporary segmentation handwriting recognizer. By using mathematical modeling and structural equation modeling (SEM) methods, a mathematical human performance was built and validated by two behavioral experiments. The mathematical model is able to predict task performance time including handwriting and correcting time based on R (system recognition time per character), UD (user delay time), WT (handwriting time), and RA (system recognition accuracy).

 A spatial segmentation handwriting recognizer has multiple handwriting windows. For example, in using a double-window recognizer (see Figure 2), user can write first character in the left window and then continues to write the second one in the right window. When the user finishes writing the second character, he or she can begin to write the third character on the left window because the system may already recognize the first character on left window when user is writing on the right window. Therefore, spatial segmentation handwriting recognizer is more efficient in improving the human performance. Because of the multiple windows on the user interface, human performance modeling is focused on predicting task completion time (T) in terms of size and distance of different areas on the user interface (see Figure 2), handwriting time, and other variables of user and system. Once a mathematical performance model is developed, it is possible to propose optimal parameters setting in designing user interface and find relatively important and cost-effective factor(s) in the system to optimize human performance.



1: Handwriting area (one window per character)

 2: Alternative characters selection (ACS) area (when one of characters is selected in text editing area, e.g. "介" in Figure 2, several alternative recognized characters as system recognition result will appear in this ACS area to let user select a correctly recognized character, e.g. "个" in Figure 2)

 3: Text editing area (listing all of characters as recognition results) Figure 2. A user interface of a double-window spatial segmentation Chinese handwriting recognizer

#### **METHOD**

# **Mathematical Modeling**

 Total task completion time (T) is composed of two parts: copywriting time  $(D_1)$ —viewing text to be copied and writing characters onto the multiple windows; correcting time  $(D<sub>2</sub>)$  correcting wrongly recognized characters in the text editing area.

*Copywriting time*  $(D_1)$ *.* Copywriting time includes time to read the character to be input (UD, user read character one by one in unfamiliar text condition), handwriting time (WT) and movement time between the handwriting windows (MT).

 According to a review of human factors studies (Mackenzie, 1992), Fitts' law (Fitts, 1954) can be used to estimate finger or stylus movement time in pointing at two different targets on user interface (see equation 1).

$$
MT_{Fitts} = a_0 + kLog_2(2A/W)
$$
 (1)

MT<sub>Fitts</sub>: movement time; A: distance of movement from start to target center; W: target width;  $a_0$ , k: regression coefficients



Figure 3. Movements of stylus on the interface in writing characters (M1: Finishing the last stroke of the first character on the left and move to start writing first stroke of the second character on the right; M2: Finishing the last stroke of the second character on the right and move to start writing first stroke of the third character on the left; Side of each window =J; Average distance from area B to C=0.5J; Average distance from area D to A=1.5J)

 Based on Figure 3 and equation 1, the movement time of  $M1$  (MT<sub>1</sub>) and movement time of M2 (MT<sub>2</sub>) will be:

$$
MT_1 = a_0 + kLog_2[2 \times 0.5J/0.5J]
$$
 (2)

 $MT_2=a_0+ kLog_2[2\times1.5J/0.5J]$  (3) Suppose there are  $F$  multiple windows ( $F=2$  in equation 2 and 3,  $MT_1$  and  $MT_2$  can be generalized into equation 4 and 5:



$$
MT_2=a_0+kLog_2[2\times (FJ-0.5J)/0.5J]
$$
 (5)

If N characters are written  $(N>F)$ , the total time spend on movement between different windows  $(MT_F)$  will be:



MT. The copywriting time  $(D_1)$  of N character will be:

$$
\begin{cases}\nD_1 = [(F-1)MT_1 + MT_2] \mid N/F \mid + N \times WT + N \times UD & (8) \\
\text{(if N mod F=0)} & \text{(if N mod F=0)} \\
D_1 = [(F-1)MT_1 + MT_2] \mid N/F \mid + [(N \mod F)-1]MT_1 \\
\text{(if N mod F \neq 0)} & (9)\n\end{cases}
$$

*Correcting time*  $(D_2)$ *.* Estimation of correcting time is depending on handwriting recognition result of the system:

 1) Path A: after user finishes copywriting of all of N characters, first, user will visually search wrongly recognized characters in the text editing area (see Figure 2, takes  $D_{vef}$  ms) and point at one of wrongly character with stylus (takes  $D_{\text{mef}}$ ms). If the correctly recognized character corresponding to the pointed character appears in the alternative characters selection (ACS) area (takes the user  $D_{\text{vef}}$  ms to search the correctly recognized character in ACS area), user has to move stylus from the text editing area to the ACS area (takes  $D_{ec}$ ) ms) and select the correctly recognized character (takes  $D_{\text{mcf}}$ ms). Finally, after the character is corrected, the stylus is moved back to text editing area for the next character to be corrected (takes  $D_{ec}$  ms). Assume within the N characters, correcting process of  $n_A$  characters follows this path A.

 Based on other experiment studies in Chinese character reading, equations of expected  $D_{\text{vef}}$ ,  $D_{\text{vef}}$  and  $D_{\text{ec}}$  are developed (see appendix 1 for detailed deduction process):



$$
D_{\text{vef}} = [173 + 199 \text{Ln}(N)] + 439 = 612 + 199 \text{Ln}(I) \tag{11}
$$

$$
D_{ec} = a_0 + k \log_2[2 \text{Lec/E(Wec)}]
$$
 (12)

 If there is an a×b metrics of characters in the text editing area, based on probability theory, it is possible to estimate expected distance between two wrongly recognized characters. Then, this distance can be plugged into the Fitts's law to calculate the expected value of  $D_{\text{mef}}$  and  $D_{\text{mcf}}$  (see Appendix 2 for detailed deduction).

$$
D_{\text{mef}} = \text{ao+klog2} \{ 2 \frac{1}{ab} \sum_{m=1}^{a} \sum_{n=1}^{b} \left[ \sqrt{(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2} \right] \} \tag{13}
$$

$$
D_{\text{mcf}} = \text{ao+klog}_2\{2\frac{1}{cd}\sum_{m=1}^{c} \sum_{n=1}^{d} \left[\sqrt{(0.5c - m + 0.5)^2 + (0.5d - n + 0.5)^2}\right]\}\tag{14}
$$

Thus, total expected correcting time in path  $A(D_{2A})$  will be:  $D_{2A} = D_{\text{vef}} + D_{\text{mef}} + D_{\text{vef}} + 2D_{\text{ec}} + D_{\text{mef}}$  (15)

 2) Path B: After user finishes copywriting of all of N characters, visually search wrongly recognized characters in the text editing area and point at one of wrongly character with stylus, if the correct character does not appear in the ACS area, user has to move the stylus back to the handwriting area to rewrite the character (first-time rewriting). Suppose

within N character, correcting process of  $n<sub>B</sub>$  character follows this path B0. After B0, there are three branches in path B (B1-B3). Branch B1: if first-time rewritten character is not recognized by the system and the user does not find the correct character in the ACS area, user has to rewrite the character again. If the second-time rewritten character is recognized correctly, user will correct the next character otherwise the user stops correcting the character according to an instruction that user stops correcting character if the character is not recognized by the system two times. Suppose within  $n_B$  characters,  $n_{B1}$  characters' correcting process follows path B1. Branch B2: if first-time rewritten character is not recognized by the system but the user find the correct character in the ACS area, user will select the correct character in the ACS area and go on to correct the next character. Assume within  $n_B$  characters,  $n_{B2}$  characters' correcting process follows path B2. Branch B3: if first-time rewritten character is recognized by the system correctly, user will correct the next character immediately. The number of characters following this correcting process will be  $n_B-(n_B)+$  $n_{B2}$ ).

 Similar to the task analysis method in path B, the correcting time of B0-B3 are estimated (see Table 1).





Therefore, the total task completion time (T) will be:

 $T=D_1+n_AD_{2A}+n_BD_{2B0}+n_{B1}D_{2B1}+n_{B2}D_{2B2}$ 

$$
+(n_{B} - n_{B1} - n_{B2})D_{2B3}
$$
 (16)

# **Experiment Validation**

 A behavioral experiment was conducted to validate equation 16. In current experiment setting of the input interface, the value of UD followed normal distribution (mean: 439 ms; standard deviation: 48 ms) based on an experiment study of Chinese reading task (Gao & Zhong, 1995). The value of  $a_0$ was 12.8 and k was 94.7 (Fitts, 1954). I=8, R=550 ms, N=24. The expected value of T in the current setting will be:

T=3002n<sub>A</sub>+(3339+WT) $\times$ n<sub>B</sub>+(2757+WT) $\times$ n<sub>B1</sub>+439 $\times$ (  $n_{\text{B}}$ -  $n_{\text{B1}}$ -  $n_{\text{B2}}$ )+1983  $n_{\text{B2}}$ +24×(WT+622) (17)

In the experiment, the value of WT,  $n_A$  to  $n_{B2}$  will be recorded by self-developed software automatically. These parameters will be plugged into equation 17 and the estimated T will be compared with T value in experiment results.

**Participants.** 6 undergraduate students (3 male, 3 female), 20–22 years old, who never used same kind of handwriting system, participated in the experiment. They were not major in statistics or psychology. To exclude effects of handedness, subjects copied sentences with their dominant hands. Subjects were thanked and paid after the experiment.

 **Material.** Sentences to be copywritten in the experiment were selected by a pilot study (Wu & Zhang, 2000). Each sentence contained 24 Chinese characters (mean: totally 179

strokes per sentence, standard deviation: 11 strokes per sentence). The pilot study found no significant difference of copywriting time among these sentences by excluding individual handwriting time as a covariate variable. Moreover, the content of the sentences were academic terms in statistics and psychology, which all of the subjects were not familiar with.

 **Apparatus.** Hanwang (EM III 6045) spatial segmentation Chinese handwriting recognizer was used. The behavioral data was recorded by online video capture software. A selfdesigned software (IntelPen) was used to measure subjects' handwriting time (WT).

 **Experimental design and procedure.** One-factor within subject design was used in this experiment. The independent variable was the number of practice stages (4 stages totally). The dependent variable was the task completion time including copywriting and correcting time. The order of sentences presented in the experiment followed ABBA order paradigm (Yang, 1996).

 First, experimenters introduced how to use the recognizer to the subjects. Then, subjects were asked to copywrite the sentences on the screen as quick as possible with their regular writing style, and then correct wrongly recognized characters. Subjects were told to stop correcting character if the rewritten character was not recognized by the system two times. Each practice stage lasted around 15 minutes, including copywriting the sentences and correcting the wrongly recognized characters. After 4 practice stages, subjects' handwriting speed was measured.

 **Experimental Result.** The total task completion time (T) in the experiment and estimated T based equation 17 were plotted on Figure 4. Table 2 listed Pearson correlation coefficients between actual and estimated task completion time and R square of the model.





 Table 2. Pearson correlation coefficients between actual and estimated total task completion time and model's R square



\*\* p<.01 (2-tailed)

 It was found that both the Pearson correlation coefficients and R square of the model increased with number of practice increasing. There was significant difference between the estimated and actual total task completion time at the first and second practice stage (t=2.559 (df=44), p<.05; Wilcoxon W=379, Z=-2.723,  $p<01$ ). At the third and forth practice stage, there was no significant difference between the estimated and actual T (t=1.875 (df=39), p $> 0.05$ ; t=0.912 (df  $=46$ , p $> 0.05$ ).

# **DISCUSSION**

 A mathematic model of human performance (equation 16) in spatial segmentation Chinese handwriting recognizer is developed and validated by a behavioral experiment. The model is able to predict human performance after user experienced two stages of practice (R square>0.9). The model's underestimation of T at the early stage of practice may stem from several reasons: first, the subjects may spend additional time in getting familiar with the system and function of each area; second, handwriting time (WT) in equation 17 was measured at the end of practice. Subjects' handwriting time may be longer at the earlier stage of practice because they may spend additional time to adapt to the friction between the stylus and the interface which is lower than normal friction between regular pen and paper.

 This model including its variants might be very helpful for the designer of the user interface to select the optimal parameters in layout of elements on the user interface. Based on parameter optimization process (see Wu et al., 2003 for detailed optimization methods), it is possible to deduct: i) optimal size of text editing, ACS and handwriting area (a, b, c, d, J) as well as their optimal distance and width (Lec, Leh, Wec, Weh); ii) optimal number of alternative characters to be selected  $(I)$ ; iii) optimal system recognition time  $(R)$ . iv) optimal number of handwriting windows. For example, when number of handwriting window (F) increased from 2 to 3, 4, 5, and 6 (other parameters in the system are fixed), the task completion time per character (T/N) is reduced by 41 ms  $(F=3)$ ,  $204ms(F=4)$ ,  $357ms(F=5)$ , and  $488ms(F=6)$ . Therefore, the model predicts that user interface with 4 or more windows will improve the human performance efficiently.

 Moreover, the model may also help the designer of the system focus on the relatively important and cost-effective factor(s) in the system to improve the human performance. For example, if the designer hope to compare the importance of two factors—handwriting time (WT) and recognition time (R) in determining the human performance, in the current interface setting, equation 16 can be developed into:

T=199( $n_A+n_B+n_{B1}+n_{B2}$ )Ln(I)+( $n_B+n_{B1}+24$ )WT

 $+(n_B+n_{B1})R+2590n_A+2925n_B+2339n_{B1}+1569n_{B2}+439(n$  $B_B - n_{B1} - n_{B2}$  + 24 UD + 24  $\times$  183 (18)

 If other parameters are fixed, weight of WT will be higher than R. This means that increasing handwriting speed (e.g. by reducing the friction between stylus and screen) will improve human performance more efficiently than decreasing the system recognition time which might require more cost on hardware of the system than the cost of reducing the friction.

 Finally, by replacing the parameters in the model (e.g. handwriting time (WT) of Chinese to WT of other language or speaking time), it is possible to use the same model to predict human performance in using other input devices, e.g. English handwriting recognizers and even voice recognition systems. Since the current model does not consider parallel processing of different subtasks, it is promising to integrate the current model with queuing-network based cognitive architecture, e.g. QN-MHP (Liu et al., 2004; Wu et al, 2004a, b) which can simulate the parallel processing without drawing scheduling charts. Our comprehensive computational model of handwriting offers not only quantitative prediction of human performance, but also a step toward developing proactive ergonomic design and analysis tools for input interface design.

# **ACKNOWLEDGMENTS**

 This article is based upon work supported by Intel China Research Center. However, any opinions, findings and conclusions or recommendations expressed in this article do not necessarily reflect the views of that institute.

# **REFRENCE**

- Cheng, Y. F. (1996). Theory of Inputting Chinese Characters into Computer. In Y. F. Cheng (Ed.), *Theory & Techniques of Inputting Chinese Characters into Computer with Keyboard* (pp. 156-212). Beijing, China: TsingHua University Press.
- Davis, R. C., Lin, J., Brotherton, J. A., Landay, J. A., Price, M. N., & Schilit, B. N. (1998). *A Framework for Sharing Handwritten Notes Demonstrations.* Proceedings of the ACM Symposium on User Interface Software and Technology, San Francisco, CA, USA. 119- 120.
- Fitts, P. M. (1954). The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology, 47*(6), 381-391.
- Frankish, C., & Hull, R. (1995). *Recognition accuracy and user acceptance of pen interfaces papers: pen interfaces.* Proceedings of ACM CHI'95 Conference on Human Factors in Computing Systems, Denver, CO, USA. 503-510.
- Gao, D., & Zhong, Y. (1995). The Influence of the frequency of usage on the speed of recognition in the Chinese characters. *Chinese Journal of Psychological Science, 18*, 225-228.
- John, B. E., & Newell, A. (1989). *Cumulating the science of HCI: From S-R compatibility to transcription typing.* Paper presented at the CHI'89 Proceedings: Conference on Human Factors in Computing Systems, Austin, Texas, USA. 109-114.
- Liu, Y., Feyen, R., & Tsimhoni, O. (2004). *Queuing Network-Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance* (No. Tech Report 04-05). Department of Industrial & Operations Engineering, University of Michigan.
- Mackenzie, I. S. (1992). Fitts' Law as a Research and Design tool in Human Computer Interaction. *Human Computer Interaction, 7*, 91- 139.
- Mackenzie, I. S., & Chang, L. (1999). A performance comparison of two handwriting recognizers. *Interacting with Computers, 11*(3), 283-297.
- Murata, A. (1996). Empirical Evaluation of Performance Models of Pointing Accuracy and Speed with a PC Mouse. *International Journal of Human Computer Interaction, 8*(4), 457-469.
- Sarcher, H., Tng, T. H., & Loudon, G. (2001). Beyond translation: approaches to interactive products for Chinese consumers. *International Journal of Human Computer Interaction, 13*(1), 41- 51.
- Wu, Y., & Sun, A. (1995). Probability Theory. In Wu, Y., & Sun, A. (Eds.), *Probability Theory and Statistics* (pp. 180-181): South China Technology University Press.
- Wu, C., & Zhang, K. (2000). *A study of copywriting task with handwriting recognizers* (Technique Report No. 20001101). Beijing, P.R. China: Engineering Psychology and Human Factors Lab., Institute of Psychology, Chinese Academy of Sciences.
- Wu, C., Zhang, K., & Hu, Y. (2003). Human performance modeling in temporary segmentation Chinese character handwriting recognizers. *International Journal of Human Computer Studies, 58*, 483-508.
- Wu, C., & Liu, Y. (2004a). *Modeling Behavioral and Brain Imaging Phenomena in Transcription Typing with Queuing Networks and Reinforcement Learning Algorithms.* Proceedings of the 6th International Conference on Cognitive Modeling (ICCM-2004), Pittsburgh, PA, USA. 314-319.
- Wu, C., & Liu, Y. (2004b). *Modeling human transcription typing with queuing network-model human processor.* Paper presented at the Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society, New Orleans, Louisiana, USA. 381-385.
- Yang, Z. (1996). Reaction Time in Experimental Psychology. In Z. Yang & J. Yue (Eds.), *Experimental Psychology* (pp. 177): East Normal University Press.
- Zhang, K. (1992). A cognitive model of typing Chinese into computer. *Journal of Chinese Information Processing, 4*, 345-349.

#### **APPENDIX**

# Appendix 1. Estimation of D<sub>vef</sub>, D<sub>vcf</sub> and D<sub>ec</sub>

 Table 3 listed the expected reaction time to search a target Chinese character in a text composed of N characters (SD=48) calculated from the studies of Gao & Zhong (1995) and Yang (1996).





curve estimation function in SPSS 1.0 ( $R^2$ = .991, F(df=8)= 865.91, p<0.001).

 $D_{\text{vef}} = 612 + 199 \text{Ln}(N)$  (10)

Similarly, we can predict the time  $(D_{\text{vcf}})$  in searching a target Chinese character in the ACS area (suppose there are I characters in the area).

 $D_{\text{vcf}} = 612 + 199 \text{Ln}(I)$  (11)

 Based on Fitts law, if the average width of ACS area and text editing area is Wec and the distance between the centers of these two areas is Lec, the movement time between these areas will be:

$$
D_{ec} = a_0 + k \log_2[2 \text{Lec/E(Wec)}]
$$
  
Similarly, we can have:  $D_{eh} = a_0 + k \log_2[2 \text{Leh/E(Weh)}]$  (12)

# **Appendix 2. Estimation of Dmef and Dmef**



Figure 5. An a×b metrics of characters in the text editing area

 Assuming that the side of each small square (one character) in Figure 5 is 1 unit, and the size of the editing area is a×b units. Suppose the stylus starts from the center of editing area (OE, coordinate (0.5a, 0.5b), see Figure 5). The coordinate of the character  $(T_{mn})$  on n-th row m-th column is (m-0.5, n-0.5).

Thus, TO—the distance between T<sub>mn</sub> and OE will be:  
\n
$$
TO = \sqrt{[0.5a - (m - 0.5)]^2 + [0.5b - (n - 0.5)]^2}
$$
\n
$$
= \sqrt{(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2}
$$

Based on (Wu & Sun, 1995), the probability density function of  $T_{mn}$ will be:

$$
\phi(x, y) = \begin{cases} 1/ab & x \in [0, a]; y \in [0, b] \\ 0 & \text{Else} \end{cases}
$$

 According to the methods in calculating the expected value of discrete random variables (Wu & Sun, 1995), the expected value of TO, E(TO) will be:

$$
E(TO) = E[\sqrt{(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2}]
$$
  
= 
$$
\sum_{m=1}^{a} \sum_{n=1}^{b} \left[ \frac{1}{ab} \sqrt{(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2} \right]
$$
  
= 
$$
\frac{1}{ab} \sum_{m=1}^{b} \sum_{n=1}^{b} \left[ \sqrt{(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2} \right]
$$

For example, if a=8, b=3, then E(TO)=2.209. Because the movement angle can be ignored when using Fitts' laws (Austo, 1996),  $D_{\text{mef}}$  and  $D<sub>mcf</sub>$  in this setting will be:

$$
D_{\text{mef}} = \text{ao+klog}_2\{2\frac{1}{ab}\sum_{m=1}^{a}\sum_{n=1}^{b} \left[\sqrt{(0.5a-m+0.5)^2 + (0.5b-n+0.5)^2}\right]\}\tag{13}
$$

$$
D_{\text{mgf}} = \text{ao+klog}_2 \{ 2 \frac{1}{cd} \sum_{m=1}^{c} \sum_{n=1}^{d} \left[ \sqrt{(0.5c - m + 0.5)^2 + (0.5d - n + 0.5)^2} \right] \} \tag{14}
$$

# **Appendix 3. Nomenclature**

