Learning Chinese Characters with Gestures

[1]Jun Ji, [1]Han Yu, [2]Boyang Li, [1*]Zhiqi Shen, and [1]Chunyan Miao

[1]Nanyang Technological University, Singapore
[2]Georgia Institute of Technology, USA

{jijun, han.yu}@ntu.edu.sg, boyangli@gatech.edu, zqshen@ntu.edu.sg

Abstract

As China's economy continues to grow rapidly, there is increasing worldwide interest in learning the Chinese language. A significant common obstacle is learning to write the numerous Chinese characters, which bear little resemblance to Western alphabets. This research explores gesture-based interaction that allows the users to practice writing Chinese characters, each within a few minutes, and cultivates an in-depth and embodied understanding of the spatial organization and writing order of the strokes in the characters. We implement a prototype system based on Microsoft Kinect and report a study involving ten users. The user study demonstrates the success of our design in preventing unintentional commands with sporadic input. We identify design challenges of writing Chinese characters via 3-dimensional gestures without haptic feedback and discuss potential solutions.

Keyword: Natural User Interface, Gesture Recognition, Chinese Calligraphy, Microsoft Kinect.

I. Introduction

As China emerges as a global economic powerhouse, worldwide interest in learning the Chinese language is also on the rise. The Chinese language is commonly considered as difficult, and writing Chinese characters is arguably the most difficult part.

* Corresponding author

Two factors contribute to the difficulties of learning the Chinese characters. First, Chinese characters differ significantly from Western alphabets. A character represents a single unit of meaning, and a character (even in Simplified Chinese) can contain more than 20 strokes, which should be written following a specific order. In additions, there are many characters in the Chinese language. It is generally believed that one needs to know 2000~3000 characters to read simple newspaper articles [1].

The ability to read and write Chinese characters is an effective way to communicate with people speaking different Chinese dialects. As the characters express meaning rather than pronunciation, they represent the lowest common denominator of a large number of dialects, such as Cantonese and Hakkanese, whose pronunciation even native speakers of Mandarin have difficulty understanding.

We believe the best way of learning Chinese characters is through continued practice. In this paper, we present the Body Brush system – our novel approach for turning body movements into brush strokes to enable users to practice writing Chinese characters based on Microsoft Kinect [3]. To the best of our knowledge, our system is the first with the objective of helping people learn Chinese characters with a Kinect-based virtual interface.

Our system aims to provide people with the following benefits:

- Facilitating users to practice writing Chinese characters in their often fragmented free time;
- Providing an embodied experience of the spatial organizations and the order of the strokes of the characters, which we believe will reinforce the memory of the strokes.

II. Related Work

Microsoft Kinect provides packaged motion sensors and depth cameras, which are useful for the development of natural user interface (NUI) using movements in three dimensions. Recent research utilized Kinect for sign languages [2], dancing practice [3], and rehabilitation [4] and so on. Kinect-based interactive systems have been brought into the living room [5] and the kitchen [6]. Our system is different from other Kinect-based systems as users need to use gestures to control delicate drawing actions rather than making selections from a menu. In this study, we aim to investigate the design principles of Kinect-based interaction in this type of applications. Previous work also studied the simulation of the effect of traditional Chinese ink on paper [7-8]. These systems proposed often involve expensive styluses and sometimes even head-mounted-displays.

III. The Concept

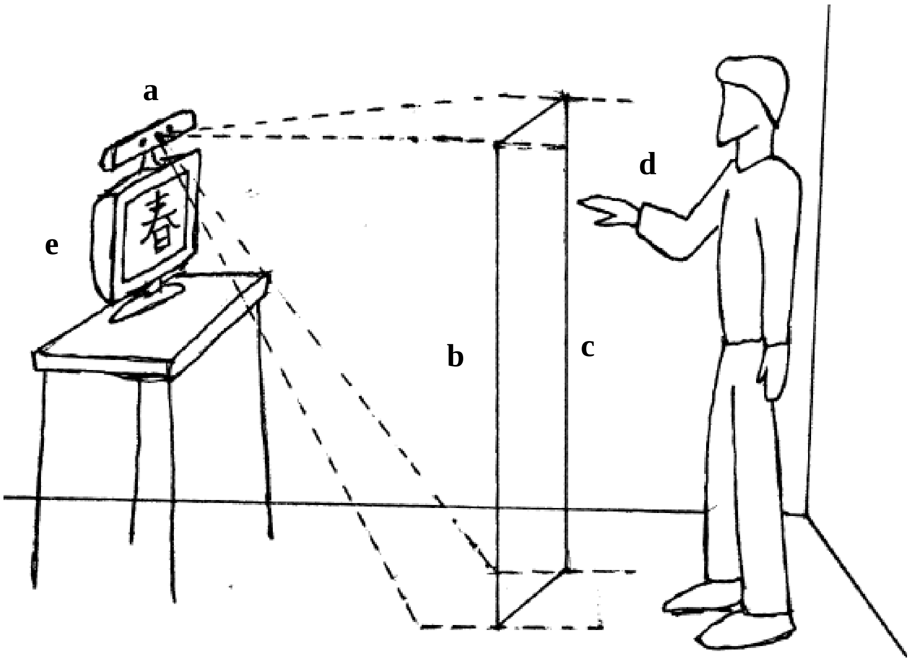

Figure 1. The design concept of the Body Brush system.

The Body Brush interface system consists of several components: human gestures, Kinect, a processor and a display as shown in Figure 1. Kinect is mounted on top of the display to be on the same level as the user's hand in order to sense the user's gestures (Figure 1-a). We create two virtual sensing volumes in the 3D space. The virtual sensing volume farther away from the user (Figure 1-b) is used to capture the user's writing actions, whereas the virtual sensing volume closer to the user (Figure 1-c) is used to capture the user's control actions. The human gesture inputs (Figure 1-d) are divided into two categories: 1) writing actions, and 2) control actions. By extending his hand into 1-b, the user signals that she is using the brush to touch the paper and is writing; by retracting her hand into 1-c, the brush is lifted up from the paper and the user can manipulate other interface artifacts to control the system.

The processor interprets the user's inputs and renders the outputs on the screen (Figure 1-e). It records the user's drawing actions and analyzes potential areas of improvements with regard to the user's writing behavior patterns. In addition, it controls network resources that enable the user to share his calligraphy artworks with his friends. The progress made by the user over different periods of time can also be displayed to let the user appreciate the results of his efforts.

IV. The Prototype

A prototype of Body Brush is created using Kinect for sensing the user's gestures (Figure 2). The graphical user interface for the Body Brush system is designed to contain the main features of a Chinese calligraphy training booklet (Figure 3). A list of characters can be selected from the scroll on the left. The scroll in Figure 3 shows a renowned poem from the Tang Dynasty, describing scenes of spring. The outline of the selected character is displayed on the main writing area in red. Ink is shown in black. A brief explanation of the character is displayed below in the brownish pages.

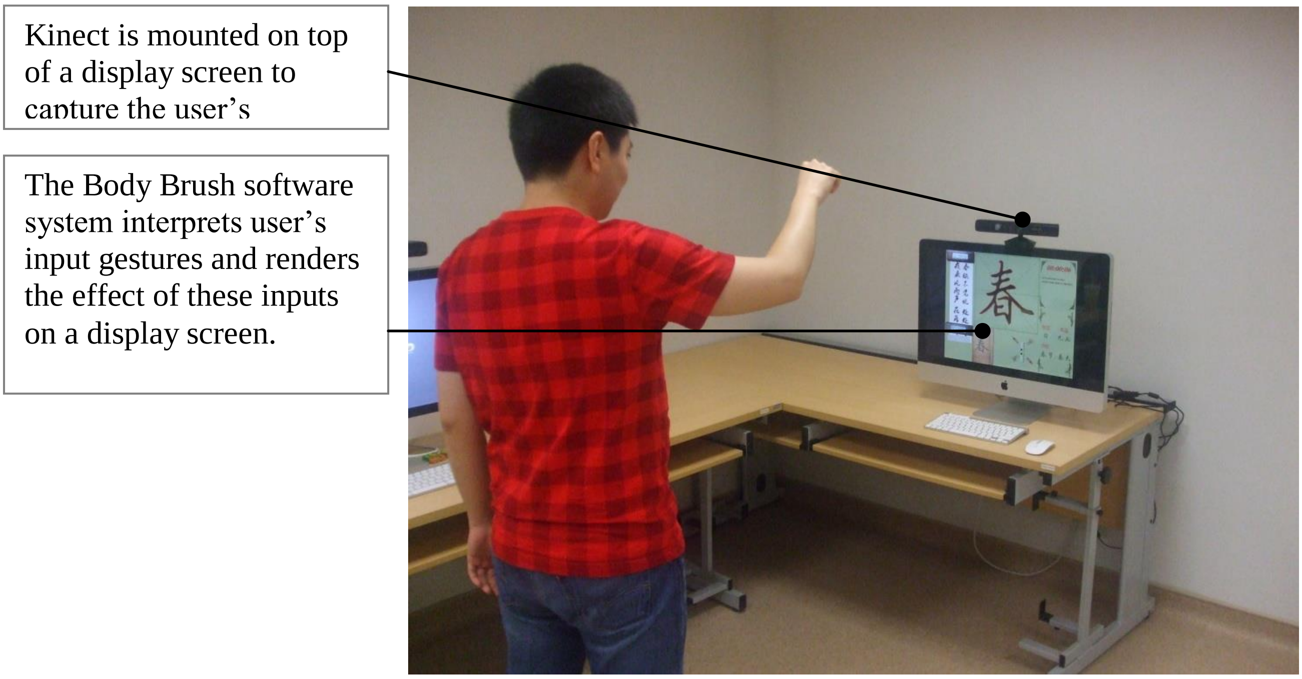

Figure 2: The prototype Body Brush system used in our usability test.

Our prototype system provides three training modes with increasing difficulty. Mode 1 provides the character outlines to be filled by users and guides the user through the correct sequence of strokes. As mentioned, Chinese calligraphy requires the strokes to be written in a specific order. In Mode 2, the system still provides the character outlines but does not enforce any order of writing. It tests if the user remembers the correct sequence. In the most difficult Mode 3, users can create their own freeform artwork for the given character without any guidance from the system.

Figure 4: Switching controls.

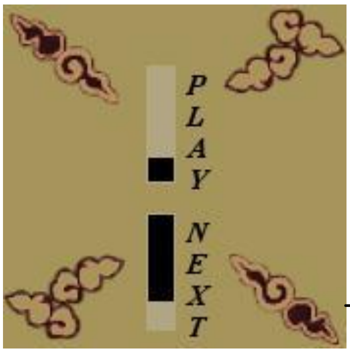

Figure 5: Filtering unintentional gestures.

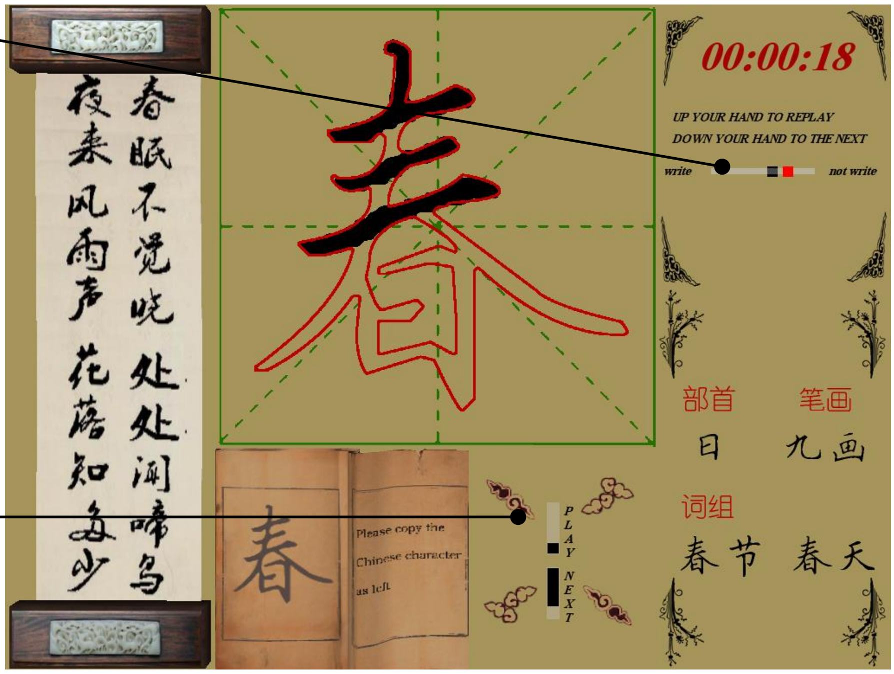

Figure 3: The display interface of the Body Brush prototype.

In the form of visual feedback, the interface artifact shown in Figure 4 indicates if the user's hand is currently in the writing volume or the control volume. The red square tracks the position of the user's hand, while the black square indicates the position of the virtual separation plane. If the red square appears on the left of the black square, the user's gestures will be interpreted as writing actions. Otherwise, the user's gestures will be interpreted as control actions.

A two-step confirmation mechanism as illustrated in Figure 5 is designed to prevent the user from inadvertently triggering control actions through unintentional gestures. The user points his hand up/down/left/right to select control actions. After that, to confirm his choice, she must put her hand to the left of her body for a period of time as indicated by the progression of the indicator bar, after which the selected action will be executed. Alternatively, the user can maintain her hand pointed to the right until the indicator bar is fully filled to cancel the selected action.

V. The User Study

A preliminary user study was conducted to improve the prototype system. Ten graduate students and research staff from the age group of 20~35 from Nanyang Technological University participated in our study. All of them speak Chinese. Each participant was asked to write ten Chinese characters taken from the poem and practice all three training modes. The time taken to complete a brush stroke from each of the three steps was automatically recorded. At the end of the trial, each user is asked to fill out a survey form designed based on the System Usability Scale (SUS) questionnaire [9].

Figure 6: The SUS Score from the survey.

38.75
out of 100 on *feeling that they would like to use the system*.

35
out of 100 on *feeling that the interactivities in the system simplifies practicing Chinese calligraphy*.

63.75
out of 100 on *feeling the various interfaces are well integrated*.

72.25
out of 100 on *feeling the features in the system are easy to learn.*

51.25
out of 100 on *feeling that they can confidently use the system*.

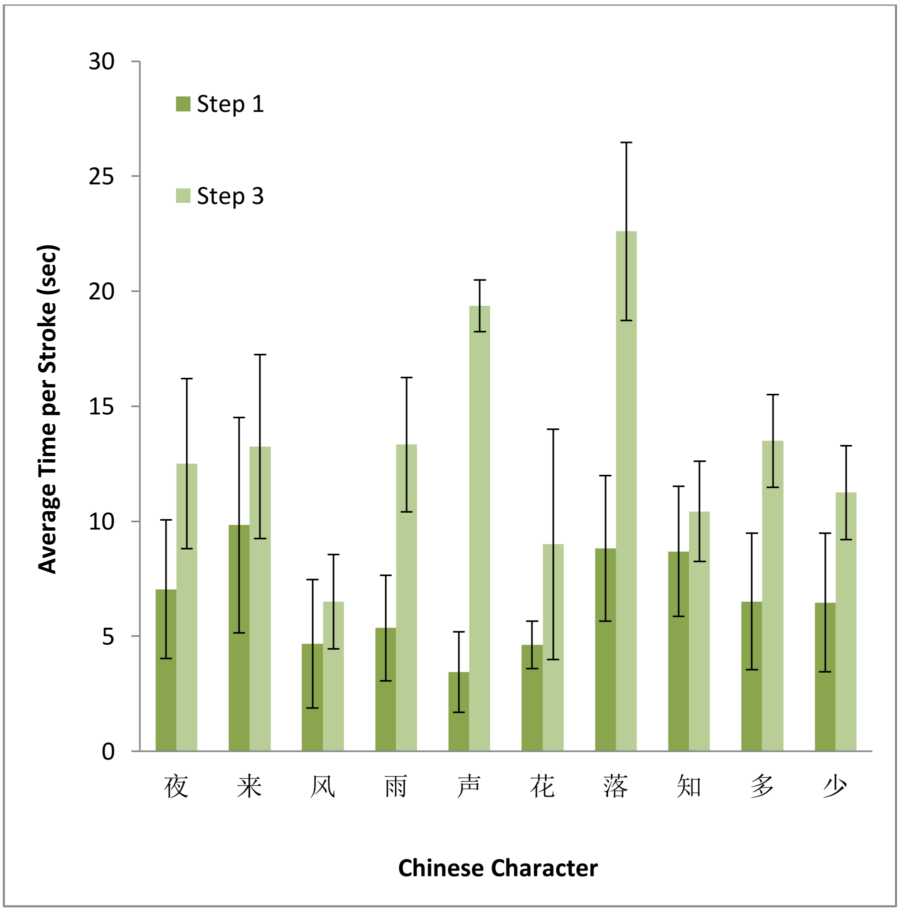

Figure 7: The averaged time taken (with standard deviation values) to paint a stroke of each Chinese character.

The SUS score derived from the survey results are shown in Figure 6. The participants found controlling the brush with the Kinect interface relatively easy to learn but hard to master. They could draw upon their existing knowledge of Chinese calligraphy without having to learn new control movements to write. However, to get used to writing in three dimensions was hard for the participants. It took some time for the test users to get familiar with this new experience. Overall, the survey questions concerning the ease of use aspect of the prototype scored an average of 45.83 out of 100 on the SUS; the questions concerning the ease of learning aspect of the prototype scored an average of 61.75 out of 100 on the SUS.

The average times taken to complete a brush stroke in Mode 1 and 3 for 10 different characters are illustrated in Figure 7. The fact that participants took longer to complete a stroke in Mode 3 (without any guidance) than Mode 1 (with guidance of character outlines and correct writing sequences) agrees with the intended difficulty level and suggests that character outlines provide useful guidance for writing Chinese characters. The difference between the modes is statistically significant based on a one-tailed Student's t-test ($t = 5.249 > 2.878$, degree of freedom = 18, $\alpha=0.01$).

Participants indicate that the main difficulty of Mode 3 was controlling the width of the stroke, which increases as the users stretches their hands further forward, even though they found the indication of position of the virtual brush relative to the virtual paper (as shown in Figure 4) to be helpful. The participants attributed the difficulty to the lack of haptic feedback of how forcefully the brush touched the paper. Thus, most participants had to erase stray brush strokes several times before a character can be finished with the desired effect in Step 3. This problem was less of a concern in Step 1 since the system ignored stray brush strokes made outside the outlines of the character.

Accidental button pushes did not happen throughout the user study. Due to the two-stage control action triggering mechanism, even when the user's gesture inadvertently selected a control action (e.g., next, replay, etc.), without the confirming gesture, the selected control action would not be executed. Most participants in the user study agree that practicing writing Chinese characters with a Kinect-based interface is novel and can potentially bring multiple benefits to the users.

VI. Limitations and Future Work

Real brushes and paper provide strong haptic feedback when writing. The lack of such feedback when writing with the Kinect interface creates a number of control challenges. This section discusses some limitations of our system and their potential solutions.

The lack of haptic friction increases small unconscious hand movements. As a result, it becomes difficult to hold the virtual brush steady or to keep lines straight. This suggests that we should not expect the users to replicate the strokes precisely. Instead, we should recognize the strokes based on the general directions of hand movement. Parameterized gesture recognition (e.g. [10]), which recognizes gestures together with its parameters such as angle and length, is ideal for this application.

Although the idea of separating the 3D space into two volumes makes intuitive sense, the association between stroke width and position in space appears difficult for the participants to master. Each participant has different perception of the association. Haptic feedback can correct for individual differences, but without haptic feedback, it becomes difficult to synchronize individual perception with the effects of the virtual brush. This problem suggests we should use less ambiguous gestures, such as finger pointing, to control the stroke width.

VII. Conclusions

Our study suggests character outline is a useful type of guidance. Some challenges in the interaction design are also unveiled and will be addressed in future work. Overall, we think Kinect-based interaction can provide a fun interface for practicing writing and an effective learning method for difficult Chinese characters.

Acknowledgements

This research is supported in part by Interactive and Digital Media Programme Office, National Research Foundation hosted at Media Development Authority of Singapore (Grant No.: MDA/IDM/2012/8/8-2 VOL 01).

References

[1] “Chinese language” (n.d.) In Wikipedia. Retrieved Jan 9, 2013 from http://en.wikipedia.org/wiki/Chinese_language

[2] Kinect. http://www.xbox.com/kinect/.

[3] Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H. and Presti, P. American Sign Language Recognition with the Kinect. In Proc. ICMI’11, ACM Press (2011), 279-286.

[4] Alexiadis, D. S., Kelly, P., Daras, P., O'Connor, N. E., Boubekeur, T. and Moussa, M. B. Evaluating a Dancer's Performance using Kinect-based Skeleton Tracking. In Proc. MM’11, ACM Press (2011), 659-662.

[5] Chang, Y.-J., Chen, S.-F. and Huang, J.-D. A Kinect-based System for Physical Rehabilitation: A Pilot Study for Young Adults with Motor Disabilities. Research in Developmental Disabilities, Volume 32, Issue 6 (2011), 2566-2570.

[6] Oh, J., Jung, Y., Cho, Y., Hahm, C., Sin, H. and Lee, J. Hands-Up: Motion Recognition using Kinect and a Ceiling to Improve the Convenience of Human Life. Ext. Abstracts CHI’12, ACM Press (2012), 1655-1660.

[7] Panger, G. Kinect in the Kitchen: Testing Depth Camera Interactions in Practical Home Environments. Ext. Abstracts CHI’12, ACM Press (2012), 1985-1990.

[8] Chu, N. S.-H. and Tai, C.-L. MoXi: Real-Time Ink Dispersion in Absorbent Paper. In Proc. ACM SIGGRAPH, ACM Press (2005), 504-511.

[9] Otsuki, M., Sugihara, K., Kimura, A., Shibata, F. and Tamura, H. MAI Painting Brush: An Interactive Device that Realizes the Feeling of Real Painting. In Proc. UIST’10, ACM Press (2010) 97-100.

[10] Bangor, A., Kortum, P. T. and Miller, J. T. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, Volume 24, Issue 6 (2008), 574-594.

[11] Wilson, A.D., and Bobick, A.F. Parametric Hidden Markov Models for Gesture Recognition. IEEE Trans. Pattern Analysis and Machine Intelligence. Volume 21, Issue 9, (1999), 884-900.

Jun Ji is currently a research associate at the School of Computer Engineering, Nanyang Technological University, Singapore. He received his M.Sc. from Shanghai Jiaotong University, China in 2007. His research focuses on digital media technologies involving natural gesture-based human computer interactions.

Han Yu is currently a Ph.D. candidate at the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. He is a Singapore Millennium Foundation (SMF) Ph.D. scholar. He obtained his B.Eng. in Computer Engineering with 1st Class Honours from NTU in 2007. He worked as a systems engineer in Hewlett-Packard Singapore Pte Ltd. from 2007 to 2008.

His research interests include trust management in multi-agent systems and intelligent agent augmented interactive digital media in education. His works have been published in leading international journals and conferences including *Proceedings of the IEEE*, *AAMAS*, *IJCAI*, and *IUI*. In 2011, he has received the Best Paper Award in the 13th IEEE International Conference on Communication Technologies (ICCT'13).

Boyang Li is currently a Ph.D. candidate at School of Interactive Computing, Georgia Institute of Technology. He obtained his B.Eng. in Computer Engineering with 1st Class Honours from Nanyang Technological, Singapore in 2008. His research interest lies in building AI that can create, tell, manage, and understand various forms of narratives, and that can support humans to maximize their creativity. He is particularly interested in learning rich knowledge structures in support of those tasks.

Zhiqi Shen is currently with the School of Computer Engineering, Nanyang Technological University, Singapore. He obtained B.Sc. in Computer Science and Technology from Peking University, M.Eng. in Computer Engieering in Beijing University of Technology, and Ph.D. in Nanyang Technological University respectively.

His research interests include Artificial Intelligence, Software Agents, Multi-agent Systems (MAS); Goal Oriented Modeling, Agent Oriented Software Engineering; Semantic Web/Grid, e-Learning, Bio-informatics and Bio-manufacturing; Agent Augmented Interactive Media, Game Design, and Interactive Storytelling.

Chunyan Miao is an Associate Professor in the School of Computer Engineering, Nanyang Technological University (NTU), Singapore. Her current research is focused on human-centered computational/ artificial intelligence and interactive media for the young and the elderly.