

To Write not Select, a New Text Entry Method Using Joystick

Zhenyu Gu^(✉), Xinya Xu, Chen Chu, and Yuchen Zhang

Shanghai Jiao Tong University, Shanghai, China
{zygu, fireon7, cc060705}@sjtu.edu.cn, yczhang@gmail.com

Abstract. Existing joystick text entry methods for game and TV boxes are cursor-based selections on virtual keyboards. In this paper we present a new text entry method using joysticks as tangible devices to capture users' freehand writing gestures. The method has considerable accuracy to accomplish English text entry. On the prediction model, we introduced HMM algorithm so users can enter text assisted with automatic correcting. We conducted a pairwise usability test on the keyboard selection method and writing-with-joystick method. The result shows that both of them are very easier to learn and writing-with-joystick is faster than the keyboard selection method both on the prediction model or none-prediction model. Subjects also report that using the keyboard selection method to enter text can be boring when using handwriting is somehow natural. This result indicates that writing with joystick may be another text entry option for game console or Smart TV users.

Keywords: Text entry · Joystick writing · Online recognition

1 Introduction

It's very common for a user sitting in the couch, input several letters like a user name or movie title on an Xbox or Smart TV. There is a long demonstrated need for people to enter text on the game console or Smart TV with joysticks. New technologies like voice recognition has been an alternative text entry method but still can't replace physical interfaces such as keyboards or joysticks in many situations. Ubiquitous connections to the net increase the need for text entry in register, search, instant messaging (IM), and email and so on from TV or game consoles. An effective text entry method would greatly enhance all of these applications and it is a fundamental requirement for extended use of IM and email [1].

The most common text entry method with joysticks is using joystick to select characters from an onscreen keyboard. Entering lots of text this way can be very slow and tedious [2]. Onscreen keyboards occupy more screen real-estate, exacerbating the need for frequent window management, and impose a secondary focus of attention [3]. However, it's still very popular for everyday TV Box text entry because users can enter text immediately without learning. Some new key layouts [4, 5] were proposed to reduce selection time. The novices, however, have to visually search for characters and remember the location of them. Andrew D. Wilson (2006) presented a bimanual text entry technique designed for today's dual-joystick game controllers [1].

While this approach increases entry speed, it needs users to pay more attention resources and possess more motion control ability. Early joystick writing approaches are alphabetical text entry methods without an onscreen keyboard. Users could write with a joystick according to a gesture alphabet [3, 7], which is designed to be simple and easy to recognize. The idea is from touch-typing on PDAs with stylus dating back to 1990 s [6].

In this paper we present a new text entry method allow users to write with joystick free of the gesture alphabet. Instead of making users learn the gesture alphabet, the approach uses an online handwriting recognition system [8] to learn users' freehand writing gestures. Discriminant features are extracted from users' handwriting samples to train a SVM [9] model. Then the model will be used to recognize user's handwriting trajectory in runtime. Online learning enables improvement of the input performance, the accuracy will increase when users enter more letters. The usability test shows that, our system is fast to learn and increases the entry speed by 2.65 characters per minute over the selection keyboard.

2 Related Work

Joystick-based text entry methods still play important roles in cases to input a short text on the game consoles, smart TVs or in-car navigation systems. There's vast body of research work on this topic, which generally consists of two main branches: selection-based and gesture-based techniques.

Selection-based text entry techniques allow users to select characters from an onscreen keyboard. Alphabetical layout and Qwerty layout are the most popular keyboard layouts. Some other layouts [4, 5] modify the layout of keys, making frequently used keys easy to access. MacKenzie, I. S., Soukoreff, R. W., & Helga, J. (2011) also proposed a zone based text entry method for joystick called H4-Writer [10]. It splits the items of keyboard into 4 sections and uses a joystick to select until only one is left. With H4-Writer, users can enter 20 words per minute, using only 1 thumb and 4 buttons.

Gesture-based text entry techniques use joystick to write, usually referring to a gesture alphabet. As the joystick is physically constrained to "write" an accurate trajectory of character, the gesture alphabets usually simplify the characters to make them easy to write. Graffiti and Unistrokes are handwriting text entry methods with stylus introduced in 1990 s [7]. Each of them designed a single stroke alphabet, easy to write and well recognized. EdgeWrite [3] places a square frame around the joystick to assist people writing along the physical edges. The trajectory of the joystick can be simplified to a sequence of touched edges and corners, which is relatively easy to recognize. The Edgewrite alphabet is shown in Fig. 1. Compared to selection-based methods, gesture-based methods need less screen real estate. Users however have to learn the gesture alphabet, and therefore the input speed is slow at the beginning.

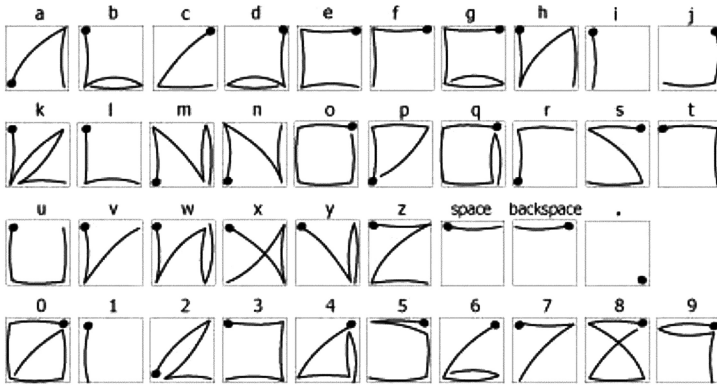


Fig. 1. The EdgeWrite gesture alphabet

3 Design

In this paper we present a new text entry method using joysticks as tangible devices to capture users' freehand writing gestures. First of all, the users' handwriting samples are collected to train a SVG (Scalable Vector Graphics) model, which will be used to recognize the users' handwriting trajectories. For sample providers, the system is considerable accurate even at the beginning. For new users other than the sample providers, we have found that the variety of samples have significant impact on the accuracy of the system. Besides samples covering more possible handwriting styles, online learning can improve the system performance as the users enter more text. Interactive feedback is also designed to guide the users to write more recognizable letters. We offered prediction input mode and non-prediction input mode as well. In the prediction mode, users entry text word by word while in the non-prediction mode, users entry text letter by letter.

3.1 Hardware & Interface

We test the prototype system on an Xbox game controller. A C ++ program was developed to deal with the signal in real time. We also designed an interactive interface including the input box and the prediction box (Fig. 2).

3.2 Online Handwriting Recognition

Handwriting recognition is the task of transforming a language represented in its spatial form of graphical marks into its symbolic representation [8]. Handwriting data can be converted to digital form either by scanning the writing on paper or by writing on an electronic surface. The two approaches are distinguished as off-line and on-line handwriting.

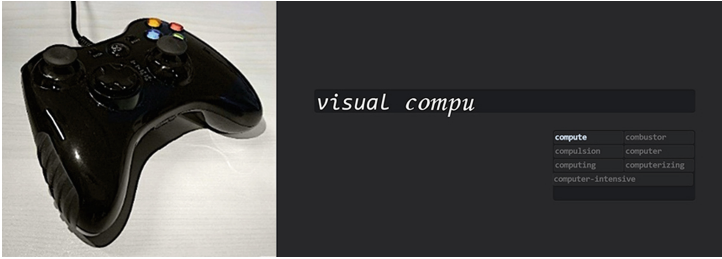


Fig. 2. The Xbox game controller and the interactive interface

The writing with joystick is an online handwriting recognition system referring a lot from that of touchpad. However, writing with joystick is quite different from the writing on touchpad. The trajectories of writing on touchpad spread on the plane and are often separated strokes while the trajectories of writing with joystick are continuous and most strokes are usually overlapped on the boundary, as swaging against the physical edge is natural and efficient for joystick writing. To segment the trajectories, we utilize state information of the stick on/off the boundary, bouncing back to the center or reversing its direction along the boundary. A character is generally segmented into on-boundary strokes and off-boundary strokes. And it also takes into account the sharp changes of directions (Fig. 3).



Fig. 3. Letter 'b' writing with joystick (left) and touchpad (right)

Feature extraction is one of the important cornerstone of any pattern classification system [11]. After a character is segmented into several strokes, each of the strokes will be transformed into a feature vector further. Seven kinds of features are extracted from their sequential and geometric information: distance, degree, absolute position, absolute degree, absolute distance and diff.

Feature vectors extracted from users' handwriting samples will be used to train a SVG (Scalable Vector Graphics) model. Then the model will be used to recognize user's handwriting trajectory in runtime.

Online learning enables improvement of the input performance. Though at the beginning extra selections are necessary to correct a few possible misrecognitions, online learning mechanism can increase the accuracy when users enter more letters. The mechanism is that when users confirm the entry result, the letters and trajectories

will be added to the model. Considering most of times joysticks are very personal devices, the system will finally be customized.

In the prediction model, we use the HMM (Hidden Markov Model) to help increase accuracy and efficiency according to the word corpus. Though each gesture may get some letters misrecognized, with this model users can entry word without interrupting to correct. The model assesses each letter's recognition result—a series of possible letters and their joint probabilities, and in conjunction with the weights of the words in the word corpus, to give a best guess. This will also help when there is a mistyping or misrecognition in the input word.

One challenge of writing with joystick is that the trajectories of some letters could be too similar to distinguish. Restricted by the moving range of joystick, for instances, the trajectories of h and b, r and n, a and d, are easy to miswrite and hard to recognize even by human being. An interactive feedback animation of real-time recognition results was designed to guide users to write more recognizable letters. For example, when users move the joystick down, turn it right to hit the edge and then keep move down along the round edge, it will show “i”, “r”, “h” in a sequence. If users move more distance along the round edge, it will show a “b” instead (Fig. 4).

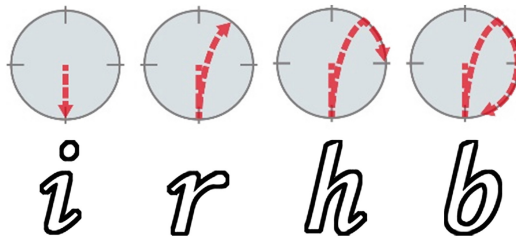


Fig. 4. The letter showed in the text box changes with the joystick's movement

4 Laboratory User Study

In order to evaluate the performance of the system, we have 15 subjects wrote each letter 10 times to get basic writing data. System training was controlled using a cross-validation procedure where 75 % of the training set was used for training and 25 % for validation. The model is not mature enough for more widely use but enough for a test.

We conducted a pairwise usability test on the keyboard selection method and writing-with-joystick method, both using an Xbox game controller and without prediction. Subjects were asked to enter text phrases as quickly as they could using both methods. It should be note that the system will keep capturing subjects' handwritings and prompting recognition results. After the test we retrained a model that was customized for the 20 subjects. The same subjects conducted another test on the writing-with-joystick method with prediction one day later. In this test, the original model and the customized model were both used (Fig. 5).

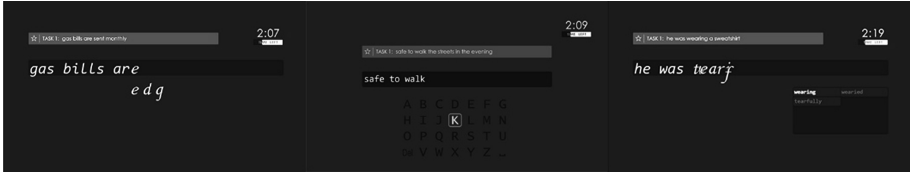


Fig. 5. The test interface of writing-with-joystick (left), keyboard selection (middle) and writing-with-joystick with prediction (right).

20 subjects were recruited for the test, aged from 20 to 24 years old. Each subject will take ten continuous sessions of tests using both two methods in an interlaced order. In each session, users needed to enter continuously for 3 min. To ensure the subjects not being disturbed, we also designed an automatic test system for both methods. Subjects could complete all ten sessions with themselves. Figure 6 shows the interfaces of the system. Polacek, O. and Sporcka, A. J. (2013) proposed that the relative position of the presented phrase and the transcribed text could also affect the test results [12]. So in the test, the position of the target phrases and the input box are all the same. The phrases are randomly selected from a collection of 500 phrases for evaluations of text entry methods published by MacKenzie and Soukoreff (2003) [13], which contain no numbers or punctuation symbols but only letters.

In the next test, subjects used the original model and the customized model to accomplish the text entry task respectively. The customized model is used to imitate the system after a long online learning process. We were interested in how the text entry performance improved with the adaptive system.

5 Results and Discussions

5.1 Writing-with-Joysitck and Keyboard Delection

Speed. Table 1 shows the average input speed across all subjects during ten sessions, measured with characters per minute (CPM). The average input speed across all sessions and subjects of writing-with-joystick is 22.73 CPM, and that of keyboard selection is 20.08 CPM (also seen in Table 1), which means writing-with-joystick is 13.2 % faster than keyboard selection method. The variance of writing-with-joystick is 1.85 when the variance of keyboard selection is 0.12, indicating the input speed of keyboard selection is more stable than that of writing-with-joystick. In fact, Fig. 6 shows that the input speed of writing-with-joystick is increasing when that of keyboard selection is stable.

Error. Soukoreff and MacKenzie (2003) divided the input error into two categories: corrected errors (errors committed but corrected) and uncorrected errors (errors left in the transcribed text) [14]. As Table 1 shows, the uncorrected error rate of both methods are very low, indicating subjects tend to correct the errors. The total error rate of writing-with-joystick is 5.94 % when that of keyboard selection is 3.49 %. We calculated the average corrected error rate of the first three sessions and the last three

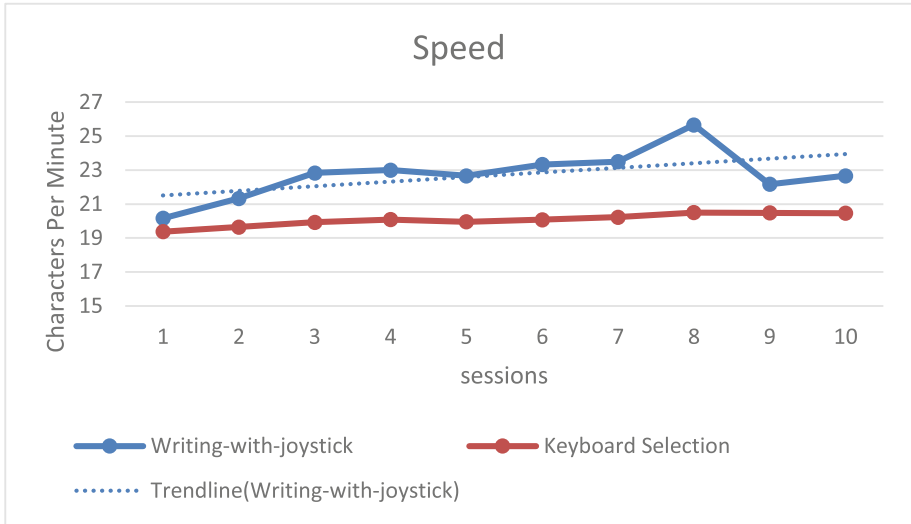


Fig. 6. Input speed of the two methods across ten sessions

sessions, and found that the session had a significant effect on the error rate of writing-with-joystick ($F_{1,38} = 5.325$, $p < 0.05$). In other words, the error rate has a significant decrease after several sessions. In fact, the total error rate of the first three sessions is 9.72 % when that of the last three sessions is 4.17 %. As online learning is not activated, it proves that interactive animations we designed play an important role in guiding subjects and making their handwritings more recognizable.

5.2 Writing-with-Joystick with Two Models

We compared the performance of the customized model and the original model. The average input speed using the customized model is 30.15 wpm (words per minute), higher than 28.76 wpm that using the original model. We found that using the retrained model had a significant effect on the input speed ($F_{1,38} = 5.724$, $p < 0.05$), which indicated that online learning did help increase the input speed.

Compared to input speed, improvement of error rate is more remarkable. By using the customized model, total error rate drops from 7.59 % to 3.8 %. F-test also shows

Table 1. Results of tests, W is writing-with-joystick and K is keyboard selection

Methods	Prediction	Model	Speed	Corrected error rate	Uncorrected error rate	Total error rate
W	×	original	22.73 cpm	5.65	0.29	5.94
K	×	original	20.08 cpm	3.28	0.21	3.49
W	√	original	28.76 wpm	4.67	2.92	7.59
W	√	customized	30.15 wpm	1.94	1.86	3.8

that the customized model has a significant effect on total error rate. Corrected error rate drops from 4.67 % to 1.94 % sharply, the reason may be that corrected errors are mostly produced by misrecognitions, which are significant fewer when using the customized model. Relatively uncorrected errors are mostly produced by personal errors, so have no big change.

6 Discussion

Text entry on game consoles, smart TVs or other platforms have two types: letters entry and words entry. We compared the performance of writing-with-joystick and keyboard selection when entered letter by letter, found that the input speed of writing-with-joystick was faster and gone up sharply. Keyboard selection is an easy-to-learn method which means there's little difference between novices and experts. This means that writing-with-joystick is more efficient than keyboard selection ever for novices or experts. The error rate of writing-with-joystick was higher at the first, but decreased a lot after several sessions. We found interactive animations played an important role in improve the performance when online learning was not activated.

Words entry is usual when fill a form or write an email. Using a retrained customized model, we found both input speed and error rate had a remarkable promotion, indicating that online learning was an effective way to improve the system. There are still much room for improvement though. In fact, when figured out the reasons for errors, we found that many errors were caused by misoperations such as an unmeant 'OK'. If we can cut down misoperations, the error rate will have a significant decrease.

7 Conclusion and Future Works

In this paper we have presented a new text entry method that allows users to write with joystick freely without a preset gesture alphabet. The approach uses an online handwriting recognition system to extract features from users' handwritings and train a SVG model. Then the model will be used to recognize user's handwriting in runtime. Interactive animations we designed help users figure out how it works and avoid miswriting. Online learning keeps collecting users' handwritings and confirmed recognition results and retraining new models, makes it an adaptive and customizable system.

Our prototype and user study demonstrate that writing-with-joystick is technologically practical and efficient in terms of usability. We have suggested a relatively simple way to extract features from the segmented joystick writing trajectories. The pairwise usability test shows that the writing-with-joystick system is more efficient than keyboard selection method as the base line even for novices or experts. With more samples of writing accumulated on line, the customized model of recognition has a significant promotion in both input speed and accuracy comparing to its initial unused state. That means online learning can improve the performance of the method further in a long run. In conclusion, writing-with-joystick is an efficient and promotable system that can be an alternative text entry method in platform like a game console or smart TV.

References

1. Wilson, A.D., Agrawala, M.: Text entry using a dual joystick game controller. In: Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM (2006)
2. Wobbrock, J.O., Myers, B.A., Aung, H.H.: Writing with a joystick: a comparison of date stamp, selection keyboard, and EdgeWrite. In: Proceedings of Graphics Interface 2004. Canadian Human-Computer Communications Society (2004)
3. Wobbrock, J.O., et al.: Integrated text entry from power wheelchairs. *Behav. Inf. Technol.* **24**(3), 187–203 (2005)
4. Rash, C.E.: Analysis and Design of Keyboards for the AH-64D Helicopter, DTIC Document (2005)
5. MacKenzie, I.S., Zhang, S.X.: The design and evaluation of a high-performance soft keyboard. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp. 25–31. ACM, Pittsburgh (1999)
6. Goldberg, D., Richardson, C.: Touch-typing with a stylus. In: Proceedings of the INTERACT 1993 and CHI 1993 Conference on Human Factors in Computing Systems, pp. 80–87. ACM Amsterdam, The Netherlands (1993)
7. Castellucci, S.J., MacKenzie, I.S.: Graffiti vs. unistrokes: an empirical comparison. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 305–308. ACM, Florence (2008)
8. Plamondon, R., Srihari, S.N.: Online and off-line handwriting recognition: a comprehensive survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(1), 63–84 (2000)
9. Bahlmann, C., Haasdonk, B., Burkhardt, H.: Online handwriting recognition with support vector machines—a kernel approach. In: Proceedings. Eighth International Workshop on Frontiers in Handwriting Recognition. IEEE (2002)
10. MacKenzie, I.S., Soukoreff, R.W., Helga, J.: 1 thumb, 4 buttons, 20 words per minute: Design and evaluation of H4-Writer. In: Proceedings of the 24th annual ACM symposium on User interface software and technology. ACM (2011)
11. Parizeau, M., Lemieux, A., Gagne, C.: Character recognition experiments using Unipen data. In: Proceedings. Sixth International Conference on Document Analysis and Recognition (2001)
12. Polacek, O., Sporka, A.J., Butler, B.: Improving the methodology of text entry experiments. In: IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom), pp. 155–160 (2013)
13. MacKenzie, I.S., Soukoreff, R.W.: Phrase sets for evaluating text entry techniques. In: CHI 2003 Extended Abstracts on Human Factors in Computing Systems, pp. 754–755. ACM, Fort Lauderdale, Florida (2003)
14. Soukoreff, R.W., MacKenzie, I.S.: Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM (2003)