

Typical Scenarios and Influencing Factors on Speech-handwriting

Licheng Xue^{1,2}

¹ State Key Laboratory of Brain and Cognition,
Institute of Psychology, Chinese Academy of Sciences
Beijing, P.R. China

Xianghong Sun¹

² Graduate University of Chinese Academy of Sciences
Beijing, P.R. China

Xia Wang³, Huanglingzi Liu³, Ying Liu³

³ Nokia Research Center
Beijing, P.R. China

Abstract—Combination of speech recognition and handwriting recognition as a text entry way has been widely used on mobile phone products. However, users seldom use speech recognition to input text. It probably because we didn't know what kind of situation was really suitable for using speech recognition to input text in daily life. The present study tried to answer what were typical scenarios and influencing factors for the combination of speech and handwriting in users' daily life. Three focus groups were conducted. Results suggested that environmental noise and social factors had an important influence on the usage of the speech recognition, while the attention demands and hands occupation limited the usage of the handwriting. Further, the speech-handwriting combination had special advantages in some scenarios, in terms of the entry speed, accuracy, and attention demands. And the combination way was less impacted by the environmental noises and social factors. These findings provided information for further field studies and system design.

Keywords—multimodal text entry; speech-handwriting; focus group; typical scenarios

I. INTRODUCTION

In recent years, multimodal text entry systems for handheld communication devices are paid more and more attention to, with the goal of supporting more transparent, flexible, efficient and power-fully expressive means [1] [2] [3]. The combination of speech and handwriting has some advantages such as natural effective expression, high channel compatibility as well as mutual complementation [4]. It has been widely applied for the multimodal interactive map system [5] [6] [7].

However, few studies are carried out in the text entry system, especially in a realistic scenario [8] [9] [10]. The present study aimed to investigate the influencing factors and typical scenarios on the usage of a speech-handwriting text entry method. We expected to get some information for the further studies. The main purposes were to:

1. Explore main criteria for evaluating the speech entry and handwriting entry methods.
2. Find the important factors influencing the usage of the speech and handwriting entry methods.

3. Explore the typical scenarios, in which the speech-handwriting entry method was especially applicable.

II. METHOD

A. Participants

16 participants were divided into three groups:

Group 1 (G1): consist of six subjects (including two designers, two experts in the field of speech recognition, one linguistics expert and one hardware designer) aged from 25 to 40 years old. Two of them were females.

Group 2 (G2): consist of five subjects (two user experience experts, two program developers, and one linguistics expert), two of them were females. The age ranged between the 25 and 40 years old.

Group3 (G3): consist of five subjects (four handwriting and keyboard users, and one key-board user), three of them were females. All of them were between 20 and 23 years old.

In addition, there were one moderator and one note taker. And the whole process was video recorded as backup and for further analysis.

B. Procedure

Each group took one hour, containing five stages:

Opening stage: The participants were asked to fill out the basic information. Equipments and the multimodal entry method were introduced. And the purpose of this discussion was declared (5mins).

Warm-up stage: Participants were required to freely talk about the experiences, characteristics as well as the influence factors when using the text entry in mobile phone (10mins).

The first discussion stage: Participants were instructed to discuss: Which scenarios the speech and handwriting entry method can be applied to? And further, some common scenarios were showed including street, dining hall, supermarket, bus, driving and office (15mins).

This study was supported by Nokia Research Center (NRC).

The second discussion stage: Participants were required to further discuss: Which factors may affect entry in the scenario? How important are they? (20mins)

The third discussion stage (ending stage): Participants were asked to discuss: in which kind of scenarios can the speech-handwriting entry be conveniently applied? Show the scene of taking a record when answering the phone as cue (10mins).

III. RESULTS

A. Main Criteria Used to Evaluate the Speech Entry and Handwriting Methods

TABLE I. EVALUATIONS OF HANDWRITING AND SPEECH ENTRY METHODS

Items	Entry Methods	
	Handwriting	Speech
Entry speed	Fast	Very fast
Accuracy	Middle	Low
Hands occupation	Hands	No
Word association	Not sure	Language model
Differences in pronunciation	No effects	Great impact
Words database	Big	Small
System speed	Slow	Fast
Experience requirements	Low	Low
Chinese, English, digital switch	Middle	Easy
Noise effects	No	Great impact
Motion effects	Great	No
Social effects	No	Great

Results shown in Table 1 indicated that each entry method had its own advantages and shortcomings. And the combination of them was expected to make up the deficiencies of each other. Generally speaking, the efficiency and accuracy of text entry were the main criterions for evaluating a text entry method. And these criteria could further provide important references for the design and evaluation of speech and handwriting entry methods in the future.

B. Typical Scenarios and Important Influencing Factors

As shown in Table 2, the handwriting and speech entry methods were used in typical scenarios such as the street, office, bus and so on. Furthermore, the handwriting entry was more applicable than the speech entry in many scenarios. However, the speech entry method had a special advantage when driving. Overall, the attention demands and hands occupation were important factors influencing the usage of the handwriting entry, while the speech entry method was restricted by the environmental noises and social factors.

TABLE II. TYPICAL SCENARIOS AND IMPORTANT INFLUENCING FACTORS

Scenarios	Entry Methods	
	Handwriting	Speech
Street	Attention demand	Noise
Dining hall	Hands occupation	Noise
Supermarket	Applicable	Noise
Bus	Applicable	Social effects Psychological feelings
Driving	Hands occupation Attention demand	Applicable
Office	Applicable	Social effects Psychological feelings

C. Scenarios Especially Applicable to the Combination of Handwriting and Speech Entry

During the third stage of focus groups participants were further asked to imagine in which situations would the combination of handwriting and speech entry method be especially applicable and has more advantages?

TABLE III. SCENARIOS APPLICABLE TO THE COMBINATION OF SPEECH AND HANDWRITING ENTRY METHOD

Scenario Name	Scenario Description
Note taking	In this scenario, this entry method could be a medium of information record. The speech part as the hidden input. For example, during the call, the speech from the call can be used as the input of text recorded; take notes in class and so on.
Alone (input fast)	Need to communication with many people by text message at the same time.
Alone (input at normal speed)	Most of the participants mentioned that they would use it as an assist. For example, alone at home or not so quiet office.
Hands occupied	When hands are busy on cooking or carrying something.
List input	Take a short note very quickly, like shopping list.
Public place	In open and quiet place, like park.
Show off	Show it to friends.
Switch fast	One can switch one entry method to another fast.
Special populations	One cannot perform fine motor such as the children, the elderly and the persons with disabilities.

From the description as shown in Table 3, the combination of speech and handwriting was especially applicable in special conditions, for example, record assist, show off and so on. These results indicated that two entries had complementary advantages: the speech entry had faster entry speed, low attention demand, while the handwriting entry had high accuracy and raises superior to environmental noises and social effects.

In fact, most of group members had mentioned that the combination of various entry methods would provide more

convenient. Users could make necessary adjustment of entry method to the properties of realistic environment and diverse situations.

IV. DISCUSSION

In the present study, we investigated the typical scenarios and important factors associated with the usage of speech and handwriting entry.

First of all, the efficiency and accuracy of text entry were the most important criterions for evaluating an entry method. Further, each method had both advantages and shortcomings. And the combined code of two entry methods would more convenient, in which users could make necessary adjustment of entry method to the current environment.

Furthermore, handwriting and speech entry methods were adaptive in typical scenarios (e.g. supermarket, bus, street). In addition, the handwriting was more widely used, while the speech entry method had a special advantage in specific situations (e.g. driving). Noises in environment and social factors had an important effect on the usage of the speech entry, while the attention demand and hands occupation limited the usage of the handwriting entry.

Finally, the speech-handwriting combined mode had more advantages in special scenarios, such as record assist, show off and so on. The combined mode had greatest advantages in terms of the entry speed, natural and attention demand. It had higher accuracy and was less impacted by the environmental noises and social factors.

V. CONCLUSION

Our study provided information for the question that what were the typical scenarios and which factors were associated with the usage of the speech and handwriting text entry. They were adopted in typical scenarios. In addition, the handwriting was more widely used, while the speech entry method had a special advantage in the specific situations. And further, environmental noises and social factors had an important effect on the usage of the speech entry, while the handwriting entry

was restricted by the attention demand and hands occupation using. The combined mode had greatest advantages, with high entry speed, low attention demand, natural, especially it was higher accurate and more applicable in a large crowd or noisy environment. These findings provided meaningful information not only for further filed studies, but also for evaluating and improving the speech-handwriting text entry system in the future.

REFERENCES

- [1] S. H. Dong, J. Wang, and G. Z. Dai, "Human-computer interaction and multimodal user interface," Science Press, August 1999, pp. 10-13.
- [2] "Spoken multimodal human-computer dialogue in mobile environments," Series: Text, speech and language technology, W. Minker, D. Bühler, and L. Dybkj, Eds, Kluwer Academic Pub., 2005, pp. 21-24.
- [3] S. Oviatt, "Multimodal interfaces," The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications, 2003, pp. 286-304.
- [4] M. Billingham, "Put that where? voice and gesture at the graphics interface," ACM SIGGRAPH Computer Graphics, vol. 32, 1998, pp. 60-63.
- [5] P. Cohen, D. McGee, and J. Clow, "The efficiency of multimodal interaction for a map-based task," Proceedings of the applied natural language processing conference, 2000, pp. 331-338.
- [6] M. Johnston, and S. Bangalore, "MATCH: An architecture for multimodal dialogue systems," Proceedings of 40th annual meeting of association of computational linguistics, 2002, pp. 376-383.
- [7] W. Wahlster, "Smart Kom: Foundations of Multimodal Dialogue Systems," Springer-Verlag, in New York, Inc. Secaucus, NJ, USA., 2006, pp. 3-6.
- [8] K. Garmer, J. Ylvén, and I. C. MariAnne Karlsson, "User participation in requirements elicitation comparing focus group interviews and usability tests for eliciting usability requirements for medical equipment: a case study," International Journal of Industrial Ergonomics, vol. 33, 2004, pp. 85-98.
- [9] P. W. Jordan, An introduction to usability: CRC Press, 1998, pp.8-11
- [10] J. Nielsen, and M. Bellcore, "The usability engineering life cycle," Computer, vol. 25, 1992, pp. 12-22.