Yan, Qifeng

The Birth of Mobile Chinese Keypad & Hybrid Input Methods

DOI: 10.1145/1520340.1520370

Published: 05/04/2009

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:
The Birth of Mobile Chinese Keypad & Hybrid Input Methods

Qifeng Yan
Nokia Research Center
Keilalahdentie 2-4
02150 ESPOO, Finland
Qifeng.yan@nokia.com

Media Lab, University of Art and Design Helsinki
Hämeentie 135 C, 00560, Helsinki, Finland
Qifeng.yan@taik.fi

Abstract
Almost all the language input devices were designed based on western linguistic and psychological model. They are just localized by changing the printings without any key layout modification for eastern countries. In this paper, the design process and user study of a Chinese style keypad and a hybrid input method are introduced.

Keywords
Chinese keypad, Chinese input, Hybrid input, Handwriting recognition, UI Localization.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
For a long run, Asian language input has to build on top of western key layout. Many input methods are invented based on existing keypad but only few of them survive because almost all of Chinese input methods need to code more than 6000 Chinese characters in GB2312 standard with the key layout designed for Latin characters. This needs extra efforts to remember the way how the local language elements are mapped to the keyboard. Figure 1 shows how Chinese radicals are mapped to QWERTY keyboard in the popular Wubi input
method. Normally a few months are needed to remember the positions of the radicals and roots.

![Wubi QWERTY Keyboard for simplified Chinese input](image)

Figure 1. Wubi QWERTY Keyboard for simplified Chinese input

Many Asian countries publish standards to spell local characters with Latin letters. E.g. China’s Pinyin system. However, many old people do not know latin letters and Pinyin.

Touch screen technology is flying and the usability of virtual keypad is becoming better and better. This provides opportunities to Asian languages because they can easily have their own keypad layouts. In this context, we start reconsidering the way how Chinese language input is adapted to western keypad, and try to develop a more intuitive keypad and a completely new Chinese input experience based on the characteristics of Chinese characters.

**Design process**

*Design Ideation*

For following reasons we decided to start from mobile device Chinese input: 1. Touch screen is going to prevail in mobile industry; 2. Short messaging is the key use case in mobile device input.

In 1988, National Language Committee of China did a throughout Chinese character frequency research on almost all available literatures in 1926-1988. The Top 1000 characters can cover 91.9% cases. [1] We think the statistics is a bit old and the results might be different in mobile short message context. Therefore, we analyzed more than 20,000 real life mobile short messages in China and found an important fact that the characters used frequently in short messages are much more concentrated than general statistics. Totally only 4329 different characters are ever used. Top 400 can cover 86.8% and Top 600 can cover 92.9%. Based on above studies, we focus on improving the input easiness for the Top 600 characters in short message context.

Currently, there are 2 main categories of Chinese input methods. The 1st category is key press based. It included Pinyin & Stroke which are popular in mainland China, Cangjie which is popular in Hongkong, and ZuYin which is popular in Taiwan. [2] This kind of input needs extra mental efforts to remember the keypad or keyboard. Another category is handwriting recognition based and the logic is simple. Users just need to draw a character on the screen and the engine tries to recognize the picture and come up a candidate list for users to select from.

Chinese characters consist of radicals, and radicals consist of strokes. When doing research on Stroke input method which uses 5 basic strokes to code Chinese characters, we found that recoding of characters needs a lot of mental efforts because strokes can not construct the characters directly. So the most intuitive way is to use radicals to build up characters. However, there are 185 different radicals and we can not give each of them...
a dedicated key. So strokes are still needed to input rare characters. Combining above thoughts, an idea of a new input method dedicated for Chinese character input comes up as shown in Figure 4: Stroke-Radical hybrid input on finger touch virtual keypad. The stroke and radical graphics design was optimized according to previous research [3].

The keypad layout is designed according to Chinese characters' square shape. This has not been considered in previous Chinese input method. Considering the similarity of some radicals, we grouped similar radicals together to share one key. According to the frequency of the radicals existing in the top characters and those characters' rank in the top list, we put 30 keys to the keypad considering users memory workload. The most frequently used radicals are in the middle and the less frequently used ones are arranged to the edges.

Radical prediction function is also implemented. Namely, when a radical key is pressed, a candidate list will pop up showing 20 characters containing this radical in frequency order. Users only need to continue inputting if the character is not yet in the list.

User research
Totally 50 native Chinese speakers were recruited for 3 runs of user research.

Some problems were found during the 1st run research: 1. all users can not find any rule for remembering the positions of the radicals on the keypad. The layout design according to frequency does not make any sense to them. The keypad need to have some hints to remember. 2. From the eye tracking data, we found that users always tend to look for the radical first in the keypad square from the position of the radical in the character square. E.g. "亻" normally exists in the left side of the character so the users try to look for "亻" from the left side of the keypad. Actually most radicals stay in the same position in different characters, if you imagine each character as a "square".

Therefore, the HID1 was updated to HID2 accordingly. We rearrange the position of radicals according to their regular position in characters. And the radicals are regrouped to make the keypad meaningful. E.g. the radicals 金(Metal), 木(Wood), 水(Water), 火(Fire), and 土(Earth) representing the 5 elements in traditional Chinese culture are arranged in a single line to help users to remember. 女 and 人 are put together in the middle to make a Chinese word: "woman".

Figure 2. Keypad layout of HID1 and HID2.

To evaluate the design of the new keypad layout, we did a simple memory test. Each participant was given 2 minutes to memorize the key layout (both HID1 and HID2), and 5 radical positions were then asked in each layout afterwards. The results show that 92% of the
radical positions from HID2 layout were remembered correctly and only 26% in HID1. The easiness of remembering HID2 keypad is quite good. So we decided to enlarge the keypad to 6x7 radicals in HID3 to cover more Chinese characters. In this version, radical input can cover 89% of the characters and Stroke input is just for inputting rare characters.

And we did 3rd run user study on 24 users to check the easiness of remembering the HID3 keypad. This time, users were given only 30 seconds but the results are still quite encouraging as shown on figure 3.

![figure 3. Memory map of the Chinese radical keypad.](image)

The 3rd run user study results are quite positive. In HID3, the number of key presses is 50% less, the error rate is 50% lower and the input speed is 20%-30% faster than users’ primary input method. 65% users would like to use it as the primary input method.

**Conclusion**

Hybrid input is proved to be a promising concept and it is worth to invest more resources to go deeper.

Especially in the age of capacitive finger touch, handwriting recognition input is facing a new challenge. Probably this hybrid input method is a solution for Asian languages and can even work better than Latin languages in finger touch input. In the next step, we will try to introduce hybrid input to Korean and Hindi languages because their native speakers are also suffering using Latin keypad to input their languages.

**Example citations**

