

09 MAR 2000

**An apparatus of generating Chinese short message
by using numeric keys**

Technical Field

5 The present invention relates to a Chinese short message generation apparatus of using numeric keys to generate a Chinese short message.

Prior Art

10 With the development of communication technology, the line telephone and the radiotelephone have been the necessities of modern life. People use the 12 keys on telephone (0~9, *, #) to input the telephone numbers of a called party for putting through the communication line. Recently, with the advance of electrical display technology, the telephone further combined with, for example, the liquid crystal display, the user can check whether the input
15 of telephone numbers is correct or not. Further, because the telephone can accommodate even a larger storing device, such as a random access memory or a read only memory, this renders that the telephone system with a data base or short message serve function can be implemented. Therefore, How to use numeric keys on the numeric keyboards to input characters efficiently has
20 been an important topic.

 As telephone short message serve, written message is sometimes clearer than spoken message, especially, while a user needs to send special phrases, such as an airplane schedule, English or the codes that may lead to mistake, etc. Therefore, at present, many telecommunication companies, such as China
25 Telecommunication, Taiwan Mobile Phone, Eastcom, etc., have their short message serve centers to provide such service. Their method is that the user uses the character input method and short message function provided on the digit mobile phone to send and receive short message. Especially, message still can be received when the user's phone is busy, which makes the phone be

used in both ways. Taking Taiwan Eastcom Company's short message serve for example, only need to set a short message serve center number "+886931413131" on mobile telephone, then enable short message serve performed. To send short message characters, users first select the "Edit short
5 message" function in the "Mail" or "Message" of mobile telephone's function menu, then use Chinese character input method to input short message, and further input the telephone numbers of the opposite party and press a "send" transmit key. Finally, the short message is sent out. On the other hand, to receive short message, after the indication of received short message is
10 displayed on the mobile telephone, and the "Read message" function in "Mail" or "Message" is selected, then the short message received can be read and automatically stored in a SIM card.

There are well-known Character input methods of using numeric keys, for example, "Numeric Keys Chinese Character Input Method" of R.O.C
15 publication No. 73889 and "Numeric Keys (Chinese Characters' spelling symbols) Input Method" of R.O.C publication No. 73890. In above two patents, Chinese characters are inputted by designating the Chinese pronunciation symbols, i.e. spelling symbols and phonetic symbols, to eight keys 1~8 in an add character key on the right side of the keyboard, and each
20 Chinese pronunciation symbol is given two codes, e.g. 11 means phonetic symbols (—) or spelling symbols (i). Additionally, the "Yao's character shape input method" of R.O.C publication No. 119811 is a Chinese character input method of using 16 character radicals to input Chinese characters. Additionally, the well-known telephone character input method and
25 apparatus, such as Japanese Examined Patent Publication No. 4-31615 said, designate pronunciation symbols to numeric keys 0~9, according to its use order. When the user desires to input characters, he must press a certain key repeatedly, until desired pronunciation symbols appear.

By taking the input method of the CD 928 plus type Chinese mobile

telephone of the Motorola Company commercial available in Taiwan in 1999 as a example, the well-known technology of using numeric keys to input Chinese characters is further described. The distributing graph of pronunciation symbols on the numeric keyboard is shown in Fig. 8. If the user desires to input "有", the pronunciation symbols of "有" are "一又 ~", and the pronunciation symbols "一" and "又" are designated to the key "2" and the key "8", respectively. So the user must firstly press key "2", and a display portion will display possible pronunciation symbols(《, ㄩ, ㄣ, 一); then press the key "8", and the display portion will display possible pronunciation symbol combinations (《 ㄣ, 《 又, ㄩ ㄣ, 一 ㄣ, 一 又, . . .) of the key "2" and the key "8". At that time , the user uses the shift right direction key "->" to shift for 4 times, and presses the confirming key "OK", then the display portion will display the homophone sequence of the pronunciation symbols"一又" (有, . . .). For the first candidate character just is the desired character, the user presses the confirming key "OK" right now, then can get the character"有".

For the same reason, if the user want to input the short message"有事请速回家", the requested input sequence of each character is as follows:

(->: Shift right selecting key , OK: Confirming key).

20

有(一又 ~): 2, 8, ->, ->, ->, ->, ok, ok (8 keys)

事(尸 `): 9, ok, ->, ->, ->, ->, ok (7 keys)

请(ㄍ一ㄥ ~): 5, 2, 9, ->, ->, ok (6 keys)

速(ㄌㄨㄝ `): 9, 5, ->, ->, ->, ok, ->, ok (8 keys)

25

回(ㄏㄨㄟ ~): 8, 5, 5, ->, ok (5 keys)

家(ㄐㄧㄚ ~): 2, 2, 1, ok (4 keys)

Therefore, when using said Motorola mobile phone as a well-known example of Chinese characters input method to input said short message,

8+7+6+8+5+4=38 keys is needed in total. In addition, for increasing Chinese characters input efficiency of numeric keys, it is common to store the frequent used phrase in memory beforehand, such as short message serve (less than 10 characters) of mobile phone. While input, it is simply need to input the code of the short message. On the aspect of the caller, it uses so-called "love code " (take characters and numbers' analogy sound) to input message. For example, if input "53019", that means"我深情依旧".

Said numeric keys short message input method of the telephone Chinese characters input apparatus and the caller of the well-known example have following problems:

1. Because the number of Chinese phonetic symbols is 37, if plus 5 tone symbols, the total is 42 symbols. The number of numeric keys is 10, so each numeric key is designated 4~5 phonetic symbols, averagely. Therefore, when inputting phonetic symbols, the user can't input correct phonetic symbols without selecting.

2. The pronunciation symbol structure of Chinese characters is often comprised of consonants, transition vowels, vowels and tones. If averagely user need to press 3 keys to input a pronunciation symbol, then averagely must press 3+3+3+3=12 keys to input a character. That makes a lot of trouble to the user of utilizing the numeric keys.

3. Although the short message serve on the mobile phone needs less pressed keys and is very convenient, the not recorded or minor modified short message sentences can't be inputted. On the other hand, if too many short message sentences were recorded, more searching time is needed and the input efficiency is reduced..

4. Although the input method of a caller is one key corresponding to one character, due to one key corresponding to a plurality of Chinese characters, the sender often can't confirm whether his/her inputting is correct or not, the receiver sometimes find it difficult to get the correct message from the

numbers.

Summary of The Invention

To solve said problem, the present Chinese short message generating
5 apparatus of using numeric keys comprising:

Key-characters corresponding dictionary for storing the numeric key
codes as search keywords and the corresponding characters using in the
Chinese short message;

Chinese dictionary for storing the numeric key code sequence as the
10 search keywords and the corresponding candidate phrase;

Character combine frequency dictionary for storing a plurality of
continue characters as the search keywords and the using frequency of the
continue characters;

Control portion is used for determining whether the key signal inputted
15 by the user is the character converting, editing or ending control signal etc. If
it is not the control signal, then refer to said key-character corresponding
dictionary, and acquire the corresponding candidate characters; if it is the
control signal but not the ending signal, then activate the character converting
portion or the editing portion; if it is the ending signal, then output the
20 converted characters;

After said control portion inputs the key signal sequence, its
corresponding candidate characters and the sequence length, all the candidate
phrases are extracted according to said Chinese dictionary, using the starting
or the ending position of each candidate character or phrase as reference
25 feature. After the candidate phrases are connected to a candidate phrase
direction graph, the phrase length weighting and the frequency weighting
acquired according to said Character combine frequency dictionary are used
as an evaluation function, to get the optimum path and sent the converted
Chinese character sequence back to the character converting portion of said

control portion;

Said control portion inputs the editing signal, the key signal sequence, the conversion character sequence and the candidate characters or phrases, and performs corresponding editing process based on the type of the editing
5 signal (modifying, deleting or inserting). Then, the editing result is sent back to the editing portion of said control portion.

According to the structure of the present invention's Chinese short message generating apparatus of using numeric keys, the control portion uses the keyboard portion to input a key signal, and determines whether it is a
10 control signal; if not, by using the key signal as the search keywords and referring to the key-character corresponding dictionary, extract its candidate characters, display them on a display portion and store them in a buffer area; if the key signal is a control signal as well as a converting signal, after the character converting portion extracts the candidate phrases of the inputted
15 key signal sequence in sequence by referring to the Chinese dictionary, and forms a direction graph for reference feature according to the starting and ending positions of the candidate characters or phrases, with the phrase length weighting and the frequency weighting resulted from the character combination frequency dictionary, the optimum path having the highest
20 evaluation function value is derived by the dynamic programming method, and the characters or phases are displayed on the display portion and stored in the buffer area; if the control signal is the editing signal, after the editing portion performs the editing operations, such as candidate phrase selecting, signal sequence deleting or inserting etc., for the inputted key signal
25 sequence, based on the type of the editing signal, the editing result is displayed on the display portion and stored in the buffer area; if the control signal is an ending signal, then the conversion character sequence stored in the buffer area is output to the output portion.

Brief Description of Drawings

Fig. 1 is a system block diagram according to the embodiment of the present invention.

Fig. 2 is a processing flow chart of the control portion according to the embodiment of the present invention.

Fig. 3 is a structural schematic diagram of the key-character corresponding dictionary according to the embodiment of the present invention.

Fig. 4 is a processing flow chart of the character converting portion according to the embodiment of the present invention.

Fig. 5 is a structural schematic diagram of the Chinese dictionary according to the embodiment of the present invention.

Fig. 6 is a structural schematic diagram of the character combination frequency dictionary according to the embodiment of the present invention.

Fig. 7 is a processing flow chart of the editing portion according to the embodiment of the present invention.

Fig. 8 is a schematic diagram of the keyboard configuration according to the prior art.

Detailed Description of The Preferred Embodiment

Fig. 1 is the system block diagram of the Chinese short message input apparatus using numeric keys according to the embodiment of the present invention. The reference numeral 100 designates an input portion comprised of the input means, such as a keyboard, etc. The reference numeral 300 designates a display portion comprised of output means, such as a liquid crystal display, etc. The reference numeral 400 designates an output portion having electronic signal output function for signal output. The reference numeral 800 is a buffer area for storing the processing intermediate result temporally. The reference numeral 600 designates a key-character

corresponding dictionary comprised of key codes as search keywords and their corresponding candidate characters. The reference numeral 710 designates a Chinese dictionary comprised of key code sequences as search keywords and their corresponding Chinese candidate phrase. The reference numeral 720 designates a Character combination frequency dictionary comprised of three Chinese characters as search keywords and their corresponding occurrence frequency. The reference numeral 500 designates a control portion for performing the processes of character converting, editing, and result outputting etc., and storing the intermediate processing results in buffer area 800, according to the key signal or the control signal input by the user through said keyboard portion 100. The reference numeral 700 designates a character converting portion for converting the key signal sequence of the intermediate result processed by said control portion 500 to a character sequence according to the Chinese dictionary 710 and the Character combination frequency 720. The reference numeral 200 designates a editing portion for performing the relative editing process for the key signal sequence or the conversion character sequence, according to the control signal determined by the control portion 500.

The present embodiment firstly collects about 300 common used short messages and love codes as an expression material base by using a variety of mediums, such as network or operating instruction, and then statistically gets about 110 characters in total after collating. Finally establishes said key-character corresponding dictionary 600, according to the corresponding relationship between the characters and the numeric keys, its structural schematic diagram shown in Fig. 3. In the meantime, collates about 90 Chinese characters as well, and then establishes said key-character corresponding dictionary 600, according to the corresponding relationship between the characters and the numeric keys, its structural schematic diagram shown in Fig. 5. Then, the occurrence frequency of three continue characters

in said expressions material base is computed by using statistical means, and stored in the Character combination frequency dictionary 720, its structural schematic diagram shown in Fig. 6.

The processing flow chart of the control portion 500 is shown in Fig. 2.

- 5 Step s200 sets the initial value of stream length n to 0. Then, after inputting the key signal by the keyboard portion 100 in step s201, judge whether the signal is the control signal or not in step s205. If no, perform step s204, selecting the corresponding candidate characters according to the key-character corresponding dictionary 600; and perform step s245, increasing
- 10 variable n by 1. Next perform step s250, storing the key signal and the candidate characters into the buffer area 800 and displaying the key signal on the display portion 300, and then perform said step s201. If the judging result is yes in said step s205, perform step s210, judging whether the control signal is the converting signal. If yes, perform step s215, extracting the key signal
- 15 sequence and the candidate characters from the buffer area 800, and performing the character converting process in the character converting portion 700; then do step s260, storing the output data of the character converting portion 700 in the buffer area 800 and displaying the conversion character sequence on the display portion 300; and next perform said step
- 20 s201. Otherwise, if the judgment of said step s210 is not the converting signal, perform step s220, judging whether the control signal is the editing signal. if yes, firstly perform step s225, extracting the key signal sequence, the control signal and the candidate characters or phrase, the converted character sequence in buffer area 800, and then performing editing process in the
- 25 editing portion 200, and then perform said step s260. If the judgement of said step s220 is not the editing signal, perform step s230, judging whether the input signal is the ending signal or not. If it is not the ending signal, return to said step s201; otherwise, perform step s235, outputting the conversion character sequence in the buffer area 800 to the output portion 400, and then

terminating the process.

The processing flow chart of the character converting portion 700 is shown in Fig. 4. Having input key signal sequence, its corresponding candidate characters and sequence length n by the control portion 500 in step s401, set the variable i to 1 in step s405; then after setting the variable j to 1 in step s410, extract the i th to j th signal sequence from the inputted key signal sequence as search keywords; after extracting the corresponding candidate phrase according to the Chinese dictionary, increase the variable j by 1 in step s420; and then determine whether the variable j is greater than n in step s421. If no, return to said step s410; if yes, after increasing the variable i by 1 in step s425, determine whether the variable i is greater than n . If no, return to said step s405. If yes, after using the starting position and the ending position of each candidate character or phrase as the reference feature, and combining them to a direction graph in step s435, and then extracting the corresponding Character combination frequency from left to right according to the Character combination frequency dictionary and computing the phrase length weighting in step s440, use the dynamic programming method to compute the evaluate function value of each node; then, in step s445, having extracted the characters of the best path as the conversion character sequence, and having sent the conversion character sequence and the candidate characters or phrases back to the control portion 500, end the process.

Evaluation function = Phrase length weighting + Frequency weighting of three sequential characters

Phrase length weighting = Character number of phrase - 1

The processing flow chart of the editing portion 200 is shown in Fig. 7. Having input the editing signal, the key signal sequence, the conversion character sequence and the candidate characters or phrases by the control

portion 500 in step s701, perform step s710 for determining whether the control signal is a candidate phrase selection. If yes, perform the selection of the candidate phrase, and modify the conversion character sequence by a user in step s715; then, after sending the editing signal, the key signal sequence, the conversion character sequence and the candidate phrase to the control portion 500 in step s740, end the process. If the determining result in said step s710 is not the candidate phrase selection, perform step s720 to determine whether the control signal is a deleting signal. If yes, after deleting the corresponding signal and the candidate characters or phrases in the key signal sequence and the conversion character sequence in step s725, return to said step s740. If the determining result in said step s720 is not the deleting signal, perform step s730 to determine whether the control signal is an inserting signal. If yes, after inserting the key signal in a designated position of the key signal sequence according to the control signal in step s735, return to said step s740. If the determining result in step s730 is not the inserting signal, then return to said step s740.

Taking the user's desiring to input the Chinese short message "有事请速回家" as an example, the operating condition of the present invention embodiment will be further described. Firstly, after the user inputs the key signal 9, 4, 7, 4, 5, 9 in sequence by using the inputting portion 100, the control portion 500 determines that each of those input signals is not the control signal in step s205. Then, in step s240, the starting and ending positions of the corresponding candidate character and the corresponding inputting key signal, shown as follows, are got respectively, according to the key-character corresponding dictionary 600, and stored in the buffer area 800. For example, the starting and ending positions of all the candidate characters of the inputting key 7 are 3 and 3.

9 -> 九, 走, 早, ..., 我, 有, 到, 家

- 4->四, 是, ..., 速, 子, 事, 上
 7->七, 痴, 去, 亲, 请, ..., 情, 妻, 见
 4->四, 是, ..., 速, 子, 事, 上
 5->五, 我, 无, 不, 回, ..., 玩, 晚
 5 9->九, 走, 早, ..., 我, 有, 到, 家

Next, after the user's inputting the converting control signal, the control portion 500 inputs the key signal sequence (947459) and said candidate characters in buffer area 800 into the character converting portion 700. Then, the following candidate phrase and the starting and ending positions of its corresponding input key signal as 5 and 6 are extracted sequentially in step s415, according to the Chinese dictionary 710 and said key signal sequence.

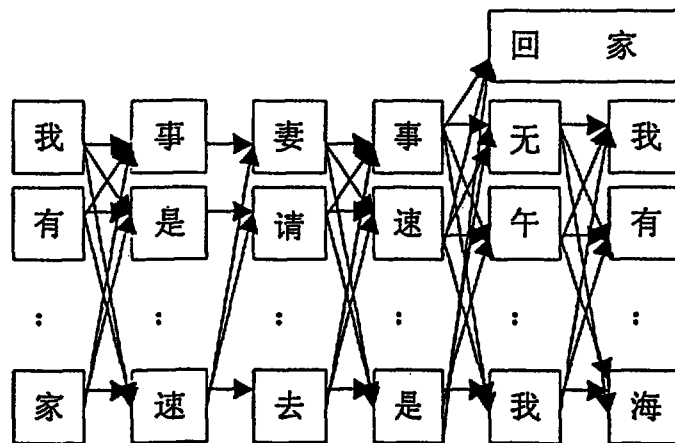
59->回家

15

Then, the following direction graph is combined by using the starting and the ending position of said candidate characters or phrases as reference feature, in step s435.

20 (Key signal sequence) 9 4 7 4 5 9

25



Next, phrase length weighting and frequency weighting of three

sequential characters of each candidate character or phrase are computed by using above-mentioned evaluation function in step s440, based on the Character combine frequency dictionary 720. For example, the phrase weighting of the candidate phrase"回家" is $(2-1)=1$, and the three sequential
5 characters frequency weighting"事回家" is 0.09,"速回家" is 0.11,"是回家" is 0, etc. Then,"有事请速回家", the characters on the best path of the embodiment, is got by using the dynamic programming method. Therefore, it is possible to get the desired short message only by inputting 6 keys, which greatly reduce the inputted key number than the 38 keys of the conventional
10 example.

The present invention is not limited to the embodiment, and can be implemented if only its essence does not change, for example, increasing or decreasing the key number on the numeric keyboard. On the other hand, the key number can be more or less than 10 keys. For example, the Character
15 combination frequency dictionary can use the message of a two Character combine frequency (bigram) as the basis of selecting characters and phrases. Additionally, the character converting portion can use the method of preferentially converting the long phrase and the characters of high using frequency to take the place of the dynamic programming method to select
20 candidate characters or phrases. This also is a modification of the embodiment of the present invention, only that the Character combine frequency dictionary portion in Fig.1 can be omitted, and the candidate characters of the key-character corresponding dictionary are arranged in accordance to their using frequency.

25

Effect of the Present Invention

The above-mentioned invention solves the problem of the conventional example. In other words, its effects are:

(1) The number of input keys of using numeric keys to input the short

message can be greatly reduced, and the input efficiency of using the numeric keys is increased. For example, in the process example of the embodiment, when the inputted key signal is "947451", the character sequence is converted to "有事请速回院", one character need one key only.

5 (2) The Chinese short messages can be input rapidly, without storing the codes of short messages.

(3) A variety of short messages can be derived automatically without storing a lot of the sentential forms, which makes it possible to save a lot of memory space.

10 (4) Simple learning function, such as extending the number of characters of the key-character corresponding dictionary, can be used, so that the number of characters and the number of sentential forms of the short message can be extended effectively.

(5) Because characters can be used to replace number codes to input and
15 output message, the message input and output efficiency of callers can be increased greatly.

In summery, the present invention has high practicability.

Claims

1. A Chinese short message generating apparatus using numeric keys, comprising of:

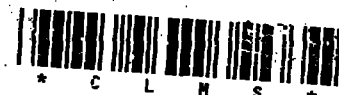
5 a key-character corresponding dictionary for storing numeric key codes as search keywords and its characters used in the corresponding Chinese short message;

a Chinese dictionary for storing numeric key code sequences as search keywords and its corresponding candidate phrase;

10 a character combination frequency dictionary for storing a sequence of characters as search keywords and the using frequency of the sequence of characters;

a control portion for determining whether the key signal inputted by a user is a character converting, editing or ending control signal etc. or not; If it is not a control signal, extracting the corresponding candidate characters according to said key-character corresponding dictionary; if it is a control signal and not an ending signal, then activating the character converting portion or editing portion operation according to the control signal; if it is an ending signal, then converting characters for output;

20 a character converting portion for extracting all the candidate phrases according to said Chinese dictionary, after a key signal sequence, its corresponding candidate characters and sequence length being inputted by said control portion; and then having used the starting or ending position of each candidate character or phrase as reference feature to combine a candidate phrase direction graph, by using a phrase length weighting and a frequency weighting got according to said Character combine frequency dictionary as an evaluation function, computing the best path, and transferring the converted Chinese character sequence back to said control portion;



an editing portion for inputting an editing signal, a key signal sequence, a conversion character sequence and candidate characters or phrases by said control portion, and performing the corresponding editing process according to the type of editing signals(modifying, deleting or inserting), and then
5 returning the editing result to said control portion.

2. A Chinese short message generating apparatus according to Claim 1, wherein the Character combination frequency dictionary using three sequential characters as search keywords.

3. A Chinese short message generating apparatus according to Claim 1,
10 wherein the Character combination frequency dictionary using two sequential characters as search keywords.

4. A Chinese short message generating apparatus using numeric keys, comprising:

a key-character corresponding dictionary for storing numeric key codes
15 as search keywords and their corresponding characters used in the Chinese short message arranged according to its using frequency;

a Chinese dictionary for storing numeric key code sequences as search keywords and their corresponding candidate phrases;

a control portion for determining whether the key signal inputted by a
20 user is a character converting, editing or ending control signal etc. or not, if it is not a control signal, then extracting the corresponding candidate characters according to said key-character corresponding dictionary; if it is a control signal and not an ending signal, then activating the character converting portion or editing portion operation according to the control signal; if it is an
25 ending signal, then converting characters for output;

a character converting portion, after inputting a key signal sequence by said control portion, its corresponding candidate and sequence length, converting the key signal sequence to a Chinese character sequence by using the method of preferentially converting the long phrase and the characters of



high using frequency, according to said Chinese dictionary, and then returning it to said control portion;

an editing portion for inputted an editing signal, a key signal sequence, a conversion character sequence and candidate characters or phrases by said
5 control portion, and after performing corresponding editing process according to the type of editing signals(modifying, deleting or inserting), returning the editing result to said control portion.

