

METHOD AND SYSTEM FOR HANDWRITING RECOGNITION

Field of the Invention

The present invention relates to the field of handwriting recognition, and in particular to
5 a method and system for handwriting recognition.

Background of the Invention

Handwriting input is a method frequently adopted by computers or other electronic devices which have a touch screen or a touch panel (writing panel). With the extensive
10 application of notebook computers, mobile phones, Personal Digital Assistants (PDA) etc., fast, efficient and accurate handwriting recognition has been a developing trend.

However, from the user experience point of view, there are some problems in current methods and devices for handwriting recognition: The writing areas on mobile devices are usually small, making writing characters with lots of strokes quite difficult; Hand-hold writing is
15 ergonomically tiring, so users often hope that inputting a character takes short time during handwriting input; A character is usually composed of multiple components, and once users forget what exactly strokes of next component are written before writing the component, the recognition timeout expires and writing cannot be finished.

US Patent No.6970599 discloses a handwritten Chinese character input method, wherein
20 merely first several strokes of a Chinese character need to be written, and every time a stroke is entered, a list of Chinese characters that match the entered stroke is displayed, so that users can find a desired character without finishing all strokes of the character. However, this method still does not resolve the problem of how to recognize an incomplete handwritten character when users do not enter strokes of the character in stroke order or they forget some strokes of the
25 character.

Summary of the Invention

In view of the above problems, the present invention provides a handwriting recognition method and device using wildcard.

According to a first aspect of the present invention, provided is a handwriting recognition method, comprising the steps of: inputting a script to be recognized; recognizing the wildcard and element of the input script; matching the recognized element and wildcard information with characters pre-stored in a character library; and displaying matching results.

5 According to an embodiment of the present invention, the wildcard is customizable for users.

According to another embodiment of the present invention, the number of wildcard(s) is one or more.

10 According to another embodiment of the present invention, in the matching step, the recognized element rather than the wildcard is compared with the element of a character pre-stored in the character library.

According to another embodiment of the present invention, the location relationship of the wildcard and the element in the input script is recognized in the recognition step.

15 According to another embodiment of the present invention, in the matching step, the element whose location has been recognized, rather than the wildcard, is compared with the elements with the same location of a character pre-stored in the character library.

According to another embodiment of the present invention, matching result(s) is one or more characters pre-stored in the character library.

20 According to another embodiment of the present invention, further comprised is the step of sorting the matching results.

According to another embodiment of the present invention, further comprised is the step of selecting a character from the matching results.

25 According to a second aspect of the present invention, provided is a handwriting recognition device, comprising: a script receiver operative to receive a script to be recognized; a recognizer operative to recognize the wildcard and element of the script from the script receiver; a matcher operative to match the recognized element and wildcard information from the recognizer with characters pre-stored in a character library; and a display operative to display the matching results from the matcher.

According to another embodiment of the present invention, further comprised is a sorter

operative to sort the matching results from the matcher.

According to another embodiment of the present invention, further comprised is a selector operative to select a character from the matching results.

According to another embodiment of the present invention, handwriting recognition
5 device is one of personal computers, Personal Digital Assistants (PDA), mobile phones and other mobile electronic devices.

According to a third aspect of the present invention, provided is a computer program product for handwriting recognition, comprising computer readable storage medium on which computer readable program code is stored. The computer readable program code comprises:
10 computer readable program code of inputting a script to be recognized; computer readable program code of recognizing the wildcard and element in the input script; computer readable program code of matching the recognized element and the wildcard information with characters pre-stored in a character library; and computer readable program code of displaying the matching results.

15 Other features and advantages of the present invention will become apparent from the following description of the preferred embodiments, taken in conjunction with the accompanying drawings, which illustrate, by way of example, the principles of the invention.

Brief Description of the Drawings

20 Fig. 1 is the flowchart of the method of handwriting recognition according to one embodiment of the present invention;

Fig. 2 is the flowchart of the method of handwriting recognition according to another embodiment of the present invention;

Fig. 3 is the block diagram of the handwriting recognition device according to one
25 embodiment of the present invention; and

Fig. 4 is a schematic user interface view of handwriting recognition according to one embodiment of the present invention.

Description of the Invention

The embodiments of the present invention will be described in detail.

It is to be understood that there are various types of devices using the handwriting recognition method of the present invention, such as personal computers, Personal Digital Assistants (PDA), mobile phones etc., and that there are also many types of script receivers for
5 entering handwritten characters on said devices, such as touch screens, handwriting panels on which handwritten characters can be entered by handwriting pen, finger tapping or mouse clicking.

It should be noted that the handwritten character to be recognized (also called script here) may be a complete character or part of a complete character, and may be in any language, such
10 as Chinese Simplified, Japanese, Korean and Chinese Traditional. Some components of the handwritten character to be recognized may be various elements of the character, and other components may be replaced by one or more wildcards.

The term “element” mentioned here may mean some components of a character or some features of the components. For example, an element may be character form, stroke, the
15 combination of character form and stroke direction, the combination of stroke and stroke direction, etc.. The present invention may adopt an element-based handwriting recognition method, for example, to perform pattern recognition on character form or stroke recognition, and is not limited to any specific handwriting recognition method. Characters pre-stored in a character library in a device applying the method of the present invention are stored according to
20 said element, and the present invention does not relate to concrete implementations of the character library.

The term “wildcard” mentioned here is a mark used for replacing some elements of a character, for example, replacing some strokes, components or radicals. The wildcard may be customizable for users and may be any symbol or character, and preferably, is a simple mark
25 that does not appear in characters.

Fig. 1 is the flowchart of the handwriting recognition method according to one embodiment of the present invention, which will be illustrated according to recognition of strokes. In this embodiment, “×” is used as the wildcard, and the character the user desires to enter is “明”.

In step 101, the user enters on e.g. a touch screen the script to be recognized, that is, “” by handwriting pen, finger tapping or mouse clicking, etc..

In step 102, the wildcard and strokes of the script are recognized. In this embodiment, the entered script is recognized to include a wildcard and a stroke sequence including
5 “left-sloping, horizontal-vertical-hook, horizontal and horizontal”, and the value i of a counter for counting the matched characters is set to 0, wherein i represents the number of characters in the character library which match the entered script.

In step 103, the recognized stroke sequence and the wildcard information are compared with strokes of characters pre-stored in the character library. For example, the sequence “*,
10 left-sloping, horizontal-vertical-hook, horizontal and horizontal” is compared with strokes of characters pre-stored in the character library. Thus, all the characters that including “left-sloping, horizontal-vertical-hook, horizontal and horizontal” will be found. Further, a similarity can be assigned to every matched characters basing on the ratio between matched strokes and the total number of the strokes of the matched characters or the number of not matched strokes. For
15 example, the similarity of character “明” is 0.8.

Those skilled in the art may understand that there are already various methods for calculating similarity in the prior art, and the present invention is not limited to any certain method for calculating similarity. It should be noted that similarity mentioned here is not necessary to implementation of the present invention but merely serves as a basis for sorting a
20 plurality of matching results.

In step 104, it is judged whether similarity 0.8 reaches a certain threshold. If not, then the flow goes to step 106. In this embodiment, the threshold is set to 0.5. Since 0.8 is larger than 0.5, it means that the entered script matches the character in the character library, and the flow goes to step 105.

25 In step 105, the matching result “明” and the corresponding similarity $x_i = 0.8$ are stored, and $i = i + 1$.

In step 106, it is judged whether the stroke sequence has been matched with strokes of all characters in the character library. If not, then the flow returns to step 103. If yes, for example, matching results are “明” (similarity $x_1 = 0.8$), “服” (similarity $x_2 = 0.8$), “朋”

(similarity $x_3 = 0.9$), “萌” (similarity $x_4 = 0.6$), “背” (similarity $x_5 = 0.7$), “月” (similarity $x_6 = 0.5$) etc., then the flow goes to step 107.

In step 107, it is judged whether there is a plurality of matching results. If yes, then the flow goes to step 108; if not, then the flow goes to step 111.

5 In step 108, matching results are sorted according to their similarity x_i . And a sequence of matching results “朋”, “明”, “服”, “背”, “萌” and “月” is obtained according to similarity ranging from high to low and according to the matching order in case of equal similarity.

In step 109, said sequence of matching results is displayed for the user to select. In step 110, the user selects character “明” from the sequence of matching results, and the method ends.

10 In another example, the entered script is “”. In step 102, the script is recognized to include two wildcards and two stroke sequences, namely “dot, horizontal, vertical turning stroke, vertical, horizontal turning stroke, horizontal” and “vertical, horizontal turning stoke, left-sloping, dot”. In step 103, the two recognized stroke sequences are compared with strokes of characters pre-stored in the character library, to obtain similarity there between. After steps 15 104 to 106, there is only one matching result, namely character “赢”, and then the flow goes to step 111 after step 107. In step 111, it is judged whether the matching result is only one character, i.e. whether i is equal to 1. Since i equals to 1, the flow goes to step 109 in which the matching result merely including character “赢” is displayed. In step 110, the user selects this matching result and the method ends.

20 Fig. 2 is the flowchart of the handwriting recognition method according to another embodiment of the present invention, which will be illustrated according to pattern recognition of the character form. In this embodiment, “o” is used as the wildcard, and a plurality of wildcards at different locations are used. And the character which the user desires to enter is “湖”.

25 In step 201, the user enters on e.g. a touch screen the script to be recognized, that is, “” by handwriting pen, finger tapping or mouse clicking, etc..

In step 202, the entered script is recognized to include two wildcards and a character form including “月”, and the two wildcards are recognized to be on the left side of the character form “月”. Then, the value i of a counter for counting the matched characters is set to 0,

wherein i represent the number of characters in the character library which match the entered script.

In step 203, the recognized character form “月” is compared with character form of characters (e.g. “潮”) pre-stored in the character library, to obtain the similarity between the entered script and the character “潮”, e.g. 0.9, without comparing the wildcard. In this embodiment, the comparison of character form is based on the location relationship. Therefore, the entered script has relatively high similarity with the characters in the character library that have “月” as the right component of the character form, whereas it has quite low similarity with characters that have “月” as, for example, the upper, lower, left, or middle component of the character form. Additionally, there are two wildcards on the left side of “月” in this embodiment. Therefore, character forms in the character library that has two components on the left side of “月” better match the entered script and has much higher similarity therewith.

In step 204, it is judged whether similarity 0.9 reaches a certain threshold, e.g. 0.5. If yes, then the flow goes to step 205. If not, then the flow goes to step 206. In this embodiment, since 0.9 is larger than 0.5, the flow goes to step 205.

In step 205, the character and the corresponding similarity $x_i = 0.9$ are stored, and $i = i + 1$.

In step 206, it is judged whether the recognized character form has been matched with character forms in the character library. If not, then the flow returns to step 203. If yes, for example, matching results are “潮” (similarity $x_1 = 0.9$), “明” (similarity $x_2 = 0.6$), “棚” (similarity $x_2 = 0.9$), “湖” (similarity $x_3 = 0.9$), “月” (similarity $x_3 = 0.5$) etc., then the flow goes to step 207.

In step 207, it is judged whether there is a plurality of matching results. If not, then the flow goes to step 211; if yes, then the flow goes to step 208.

In step 208, matching results are sorted according to their similarity x_i . And a sequence of matching results “潮”, “棚”, “湖”, “明” and “月” is obtained according to similarity ranging from high to low and according to the matching order in case of equal similarity.

In step 209, said sequence of matching results is displayed for the user to select. In step 210, the user selects a character, namely the character the user desires to enter, from the

sequence of matching results, and the method ends.

The following example will further illustrate the impact on the final result of handwriting recognition of the location relationship between wildcard and character form. For example, “Δ” is used as the wildcard. When “” is entered, characters that have “口” at the outside have relatively high similarity, such as “国” and “围”. When “” is entered, characters that have “口” at the inside have relatively high similarity, such as “回” and “问”. When “” is entered, characters that have “口” at the lower part have relatively high similarity, such as “吝” and “言”. And when “” is entered, characters that have “口” on the left side have relatively high similarity, such as “吃” and “喝”.

Moreover, the combination of character form and stroke direction is used as recognition element helps to enhance the accuracy of the handwriting recognition method. For example, when the user desires to enter “芊” and the handwritten script is “”, it might be recognized as a wildcard and character form “千” in recognition of character form, so an undesired character “芊” might be obtained. However, through recognition of stroke direction along with recognition of character form, the first stroke of “千” is recognized as “left-sloping” instead of “horizontal”. In this manner, the accuracy and speed of handwriting recognition is improved greatly.

Fig. 3 is the block diagram of the handwriting recognition device according to one embodiment of the present invention. A handwriting recognition device 300 may be any device capable of handwriting input, such as a personal computer, a PDA, a mobile phone etc.. Handwriting recognition device 300 is provided with a script receiver 310, a recognizer 320, a matcher 330, and a display 340. Further, it may comprise a sorter 350 and a selector 360.

The script receiver 310 may be a touch screen, a handwriting panel and the like on which handwritten characters can be entered by handwriting pen, finger tapping or mouse clicking, for entering scripts to be recognized.

Upon receipt of the script entered through the script receiver 310, the recognizer 320 recognizes the wildcard and element in the script of said script receiver.

After the recognizer 320 completes the recognition of the wildcard and element, the matcher 330 matches the element and the wildcard information from said recognizer with

characters pre-stored in a character library, to obtain corresponding similarity without comparing the wildcard.

After matching with all characters pre-stored in the character library is completed, the sorter 350 sorts matching results according to the matching results obtained by the matcher 330 and the corresponding similarity.

The display 340 displays matching results directly from the matcher 330 or being sorted by the sorter 350.

The selector 360 selects a character from said matching results.

The present invention provides a method of recognizing handwritten characters using wildcards, reducing the time for handwriting characters. By introducing the concept of wildcard, users can write the simple wildcard to represent complex elements in a character. In particular, when complex characters are hand-written in a small writing area or when users forget what exactly a character to be input is written, the input efficiency is improved significantly.

Fig. 4 is a schematic user interface view of handwriting recognition according to one embodiment of the present invention. As shown in this figure, character “巛” has three components, namely “山”, “真” and “页”. The total number of strokes is 19, with 10 from the component “真”. In this case, users can just write a circle as wildcard to represent this element. With the handwriting recognition method and device according to the present invention, matching results of this script are “巛”, “巛” and “巛”. Then, users can select the desired character “巛”, thereby saving the time for inputting the 10 strokes of component “真” of character “巛”. Therefore, the present invention greatly reduces the number of strokes that need to be input during the process of inputting hand written characters, thereby improving the input efficiency. Users can input characters quicker and easier. Even when users forget the exact writing format of a character, they can still manage to input the character. The method of the present invention is easy to understand and easy to learn. Moreover, the method of the present invention is merely an addition to traditional handwriting methods and does not affect users if they do not use it.

While the foregoing has been with reference to specific embodiments of the invention, it will be appreciated by those skilled in the art that these are illustrations only and that changes in

these embodiments can be made without departing from the principles of the invention, the scope of which is defined by the appended claims.

What is claimed is:

1. A handwriting recognition method, comprising the steps of:
inputting a script to be recognized;
5 recognizing the wildcard and element of said input script;
matching said recognized element and the wildcard information with characters
pre-stored in a character library; and
displaying said matching results.
2. The handwriting recognition method according to claim 1, wherein said wildcard is
10 customizable for users.
3. The handwriting recognition method according to claim 1, wherein the number of said
wildcard(s) is one or more.
4. The handwriting recognition method according to claim 1, wherein in said matching
step, the recognized element rather than the wildcard is compared with the element of a
15 character pre-stored in the character library.
5. The handwriting recognition method according to claim 1, wherein the location
relationship of said wildcard and said element in the input script is recognized in said
recognition step.
6. The handwriting recognition method according to claim 1, wherein in said matching
20 step, said element whose location has been recognized, rather than the wildcard, is compared
with the elements with the same location of a character pre-stored in the character library.
7. The handwriting recognition method according to claim 1, wherein said matching
result is a plurality of characters.
8. The handwriting recognition method according to claim 6, further comprising the step
25 of sorting said matching results.
9. The handwriting recognition method according to claim 6, further comprising the step
of selecting a character from said matching results.
10. A handwriting recognition device, comprising:
a script receiver operative to receive a script to be recognized;

a recognizer operative to recognize the wildcard and element of the script from said script receiver;

a matcher operative to match the recognized element and the wildcard information from said recognizer with characters pre-stored in a character library; and

5 a display operative to display the matching results from said matcher.

11. The handwriting recognition device according to claim 10, wherein said wildcard is customizable for users.

12. The handwriting recognition device according to claim 10, wherein the number of said wildcard(s) is one or more.

10 13. The handwriting recognition device according to claim 10, wherein said matcher compares the recognized element rather than the wildcard with the element of a character pre-stored in the character library.

14. The handwriting recognition device according to claim 10, wherein said recognizer recognizes the location relationship of said wildcard and said element in the input script.

15 15. The handwriting recognition device according to claim 10, wherein said matcher compares said element whose location has been recognized, rather than the wildcard, with the elements with the same location of a character pre-stored in the character library.

16. The handwriting recognition device according to claim 10, wherein said matching result is a plurality of characters.

20 17. The handwriting recognition device according to claim 16, further comprising a sorter operative to sort said matching results from said matcher.

18. The handwriting recognition device according to claim 16, further comprising a selector operative to select a character from said matching results.

25 19. The handwriting recognition device according to claim 10, said device is one of personal computers, Personal Digital Assistants (PDA), mobile phones and other mobile electronic devices.

20. A computer program product for handwriting recognition, comprising computer readable storage medium on which computer readable program code is stored, the computer readable program code comprising:

- computer readable program code of inputting a script to be recognized;
- computer readable program code of recognizing the wildcard and element in the input script;
- computer readable program code of matching the recognized element and the wildcard
- 5 information with characters pre-stored in a character library; and
- computer readable program code of displaying the matching results.