

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
29 January 2004 (29.01.2004)

PCT

(10) International Publication Number
WO 2004/010273 A2

(51) International Patent Classification⁷: G06F 3/00

(21) International Application Number:
PCT/CA2003/001073

(22) International Filing Date: 17 July 2003 (17.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/396,576 18 July 2002 (18.07.2002) US

(71) Applicant and

(72) Inventor: TANG, Hongzhao [CA/CA]; 32-110 Stewart Street, Ottawa, Ontario K1N 6J6 (CA).

(74) Agents: KINSMAN, L., Anne et al.; Borden Ladner Gervais LLP, World Exchange Plaza, 100 Queen Street, Suite 1100, Ottawa, Ontario K1P 1J9 (CA).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

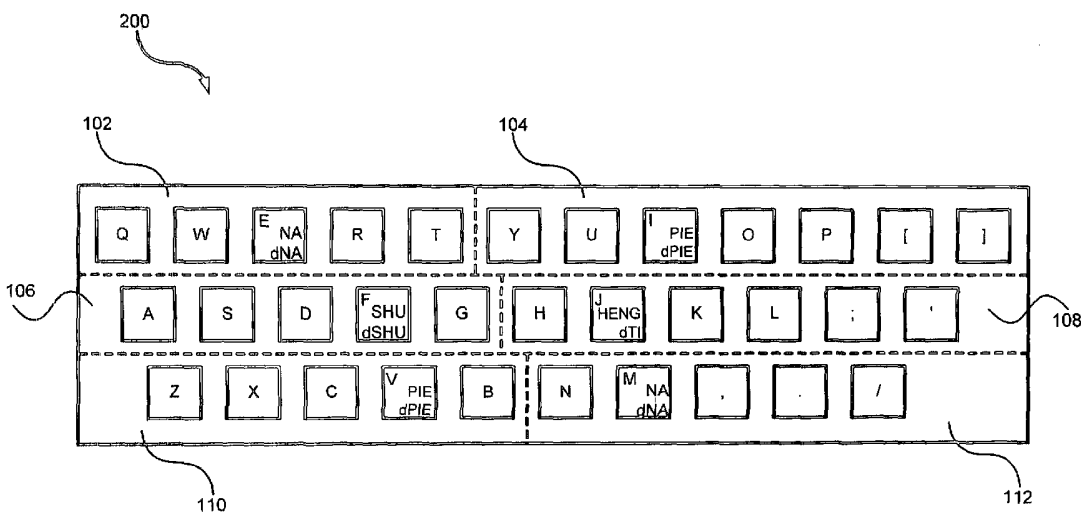
(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: RELATIVE STROKE IDEOGRAPHIC CHARACTER INPUT KEYBOARD



(57) Abstract: The invention provides a character input method (CIM) for ideographic languages to input textual data on the basis of relative stroke positions within the characters to produce electronic files. By identifying key-strokes and assigning each of the key-strokes to appropriate multiple keyboard positions, this character input method simulates the relative positions of these key-strokes within a particular character. The user can select the appropriate keyboard position for a particular essential stroke depending on that essential stroke's position relative to other strokes within the layout of a particular character or character-root. The keyboard positions for other non-key strokes are also decided by their relative stroke positions within the character.

WO 2004/010273 A2

Relative Stroke Ideographic Character Input Keyboard

FIELD OF THE INVENTION

The present invention generally relates to a system and method for the input of ideographic characters, such as the characters used in Chinese, Korean and Japanese languages. More particularly, the present invention relates to a relative stroke input method, and keyboard, for electronic processing of strokes to produce characters in a language based on characters formed from strokes.

BACKGROUND OF THE INVENTION

Modern computers typically provide a keyboard as an input tool to users. Keyboards are commonly designed around a Roman alphabet. The origins of the keyboard lie in typewriters which grew to prominence in countries that used Romance languages and a script derived from Latin. Other keyboards are available for alternate languages such as Greek and Cyrillic languages that have different scripts. These keyboards all share a common structure in that they are designed for the input of characters in languages where each word is created from a fixed set of characters. Thus, from a fixed character set, these keyboards permit the entry of both any word in the language and any word that can be added to the language.

However, there exist languages that do not rely upon phonetic characters to create words, and instead have their script formed by a set of ideographic characters. Many languages of this type exist in Asia, and include various Chinese languages and dialects, Korean and Japanese. Due to the nature of ideographic languages, it is difficult to use the standard computer keyboard to input characters directly. To overcome the inherent problems associated with ideographic character input a number of solutions have been developed to allow ideographic input.

These approaches can be grouped into five input methodologies. The first input methodology uses the Roman letters to phonetically spell words in a language, and upon indicating the end of a word a software routine replaces the phonetic spelling with an ideographic character. The second input methodology is based on the creation of ideographic characters from component roots, with each component root assigned a key or key combination. The third methodology is based on the creation of characters from the component strokes used to build a character. The fourth methodology is based on assigning specific codes to characters based on other characteristics, such as telegraph

codes are assigned for different ideographic characters. A fifth methodology is used to group the input methods designed as hybrids of the previously described methods.

The first methodology, the input methods based on pronunciation, is a commonly applied technique. Each ideographic character is transliterated by the typist into a string of Latin characters and then entered. This methodology provides a simple method for typists that understand how different letters in the Roman script are pronounced, but also relies upon a wide variety of typists to use a standard pronunciation for words. As is well understood by linguists, for a variety of reasons the populations of different geographic regions develop different pronunciations for the same character. As a result, people in different regions will prefer different strings of Roman characters to represent the same ideographic character. To compensate for this a strict pronunciation guide can be enforced to a single pronunciation, but this disadvantages different regions, as they are no longer entering characters using what they think of as a native pronunciation. Additionally, many characters are pronounced in identical or sufficiently similar manners to each other so that a plurality of characters are represented by identical strings of Roman characters. To compensate for this problem, upon the completion of a string of Roman characters, the typist is typically presented with a list of the possible characters to choose from. While this allows the resolution of homonyms to characters, it interrupts a keyboard driven process and slows down the typist. The heart of these problems is the mapping of ideographic characters to letters used in a script based on a different character organisation method. Those skilled in linguistics will readily appreciate that an additional problems is that there are phonemes in on language that cannot be properly represented using letters from another language. These phonemes create difficulty for transliteration based input. One skilled in the art will also appreciate that the discord between writing ideographic characters of one language using a text entry method based on the pronunciation of the ideographic characters using the letters of a different language creates a disconnect that results in an information gap between the character graphics and the keyboard typing. Despite the fact that this character input methodology is among the most popular input method, largely due to the simplicity of implementation, it presents several drawbacks and is inaccessible to those without strong language skills.

The second methodology, the one based on character roots, relies upon the fact that most ideographic languages build characters using other characters as roots. Complex or compound characters are usually composed of a collection of simpler character-roots. Each of these character-roots can be devolved into component strokes.

One example of such character roots is the set of radicals that typically appear on the left side or atop complex Chinese characters. Using these character roots, which are often characters themselves, many text entry methods have been designed. Skilled users have been able to become very efficient in text entry, but as a result of the large set of character roots, it is difficult for non-experts to learn and use these methods effectively. As a result, only a small number of typists can use these methods skilfully, and then only after considerable training and practice.

The third methodology builds characters from constituent strokes. In the actual writing of ideographic characters, a set of strokes is used to form the character. The set of strokes is both finite, and relatively compact when allowance for slight variances is made. Thus, to enter a character, the strokes used to create the character are entered based on a predefined structure. This is an analog of how the characters are written. However, input methods in this grouping have been consistently viewed as inefficient, as it is difficult to become proficient at the entry of text due to the complex relationship between the strokes. Due to the inter-relationship of the strokes used in characters, it is difficult to arrange a mapping of strokes to keys that allows for rapid and convenient text entry. As a result of the historical awkwardness of this input methodology it is commonly viewed as an inefficient means for character entry.

The fourth methodology is based on the assignment of special codes to different characters. This methodology has grown from the assignment of telegram codes to characters, or the development of codes that are based on the shape used at each corner of the character. This methodology is overly complex and has not been adopted outside of very small niche markets.

The fifth methodology is a hybridization of previous methodologies. One such hybrid combines the features of pronunciation and character roots to provide a mixed method of character entry.

Evaluation of the efficacy of character input methods can be based on the following three criteria: ease in learning; convenience; and efficiency in use. In introducing an input method to a new typist, the methods that are easiest to learn are most appreciated, while for experienced typists, the efficiency of text entry is valued. For non-professional users the convenience of character input is paramount, as it is possible to learn moderately difficult systems, but systems that are not easy to use tend to be rejected despite their efficiency. Current character input methods all fall short in one or

more of these criteria, and thus effective character/text input is still a major problem for ideographic languages.

There is, therefore, a need for a new method for inputting ideographic characters using standard keyboards that presents a simple and efficient character input method.

5 SUMMARY OF THE INVENTION

It is an object of the present invention to obviate or mitigate at least one disadvantage of previous ideographic character input methods. It is a further object of the present invention to provide a method and system for stroke based character input that utilises multiple placements of a single stroke to improve the convenience and efficiency
10 of the character input.

In a first aspect, the present invention provides a keyboard, for the entry of ideographic characters based on component strokes. The keyboard comprises a plurality of stroke keys and a non-stroke key. Each of the plurality of stroke keys represents a component stroke, at least two of the plurality of stroke keys representing the same
15 component stroke, each key provides a signal representative of the stroke associated with the key. The non-stroke key, or optionally a plurality of non-stroke keys, provides a signal representative of the end of the entry of a character.

In an embodiment of the first aspect of the present invention, the plurality of stroke keys includes two sets of two stroke keys, keys of the first of the two sets representing a
20 first component stroke, and the keys of second of the two sets representing a second component stroke. In another embodiment of the present invention, the keyboard is arranged to form two contiguous regions, horizontally adjacent to each other, the regions being a right and left sided region. Preferably wherein the keys of the first and second set are positioned so that one of the keys of each set is in the right sided region, and the
25 other key of each set is in the left sided region.

In another embodiment of the present invention, the ideographic characters are Chinese characters, and the at least two of the plurality of stroke keys representing the same component stroke represent a stroke selected from a list including stroke PIE, stroke dPIE, stroke NA and stroke dNA. In a further embodiment of the present invention,
30 the keyboard has the layout of a QWERTY keyboard and one of the at least two of the plurality of stroke keys representing the same component stroke resides in the row above the home row, and preferably the other of the at least two stroke keys representing the same component stroke resides in the row below the home row.

In a further embodiment of the present invention, the keyboard includes a stroke interpreter, operatively connected to the plurality of stroke keys and to the non-stroke key. The stroke interpreter receives and buffers the provided signals, and selects an ideographic character from a database based on the buffered signals representative of strokes when a signal representative of the end of the entry of a character is received. The stroke interpreter preferably includes means for both narrowing the number of characters for selection with each stroke; and selecting a character from the narrowed number of characters on the basis of a signal representative of one of the plurality of non-stroke keys. Additionally, the keyboard preferably includes a conflict resolution module, operatively connected to the database and the stroke interpreter. The conflict resolution module selects ideographic characters from the database when the buffered signals are not uniquely associated with one character, the selection based on a signal representative of one of plurality of non-stroke keys received in response to a presentation of the characters in the database associated with the buffered signals.

In another embodiment, the database contains a plurality of characters indexed according to the component stroke values in a sequence associated with each of the plurality of characters.

Other aspects and features of the present invention will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments of the invention in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the present invention will now be described, by way of example only, with reference to the attached Figures, wherein:

Fig. 1A, Fig. 1B and Fig. 1C provide tables showing 34 common strokes, stroke titles for identification, sample characters containing the strokes, and a stroke classification system;

Fig. 2 is a table showing a collection of stroke pairs with common graphic origin but different representations;

Fig. 3 is a table of four exemplary primary strokes, their point versions, and stroke titles;

Fig. 4 is a table of two exemplary essential strokes, their point versions, and stroke titles;

Fig. 5 illustrates a keyboard of the prior art

Fig. 6 is a schematic diagram showing partitioned sections of standard QWERTY keyboard positions;

Fig. 7 illustrates the numeric keypad of a hand-held calculator;

Fig. 8 illustrates the numeric keypad of a telephone set;

5 Fig. 9 provides a tabular assessment of the distribution of common characters, using the 6763 characters in the GB2312-80 character internal code table as example, over different groups by the number of strokes per character.

10 Fig. 10a and Fig. 10b provides a tabular assessment of distribution and frequency of strokes, using the 6763 characters in the GB2312-80 character internal code table as example and under stroke recognition of this invention. Stroke LING is not included;

Fig. 11 is a schematic representation of an exemplary QWERTY keyboard arrangement for primary strokes, according to an embodiment of the invention;

15 Fig. 12 is a schematic representation of the arrangement of all strokes on the standard QWERTY keyboard, according to an embodiment of the invention;

Fig. 13A and Fig 13B provide a summary table of the standard QWERTY keyboard arrangement for all the 34 strokes and sample characters which contain relative stroke positions concerning the particular stroke;

20 Fig. 14 is a schematic representation of the arrangement of all strokes on the version of QWERTY keyboard used by handheld computers and Personal Digital Assistants (PDAs), according to embodiment of the invention;

Fig. 15 illustrates an exemplary version of an arrangement of different stroke-units on a numeric keypad, according to embodiment of the invention;

25 Fig. 16 is a schematic representation of a simplified version of arrangement of all strokes and stroke units on the QWERTY keyboard used by handheld computers and Personal Digital Assistants (PDAs), according to embodiment of the invention;

Fig. 17 provides a reference list of technical terms used in the application in both English and Chinese languages.

30 DETAILED DESCRIPTION

Generally, the present invention provides a method or a system for inputting ideographic characters using a keyboard based upon the stroke sequence of a character, which takes into consideration the relative stroke positions within a character.

Ideographic characters are typically two-dimensional arrangements of strokes, as opposed to the linear nature of indo-european descendent languages. Each character is created using a unique arrangement of strokes, so that different strokes are either overlapping in some fashion, or are either horizontally or vertically offset from each other, or are in other positions relative to each other. The present invention provides a stroke based keyboard mapping, and a method for creating ideographic characters that devolves each character into a set of strokes. The strokes are mapped to a conventional keyboard which is used to enter the characters.

While previous methods of stroke based input have been viewed as awkward to use, the present invention provides an input mechanism derived from an analysis of the strokes used to create common characters, and the relationship between the strokes used. In a presently preferred embodiment, characters are produced by entering the component strokes used to create the character, in the sequential order used to commonly write the character, in which the characters are usually hand-written.

For the purposes of the following discussion, the ideographic writing system used in Chinese characters is used as the basis for a new keyboard design. One skilled in the art will appreciate that there may be variances between the presented design and embodiments designed for different ideographic character sets, but that the fundamental advance of the present invention is applicable to other ideographic languages. Additionally, one grouping of the identified strokes are designated as essential strokes. The term essential stroke is used in reference to the nature of the strokes in relation to each other and to other strokes used the language. In different languages, other stroke combinations may be deemed essential, and may be substituted for the stokes described herein as essential without departing from the present invention.

The stroke table provided in Figures 1A to 1C (referred to herein as Figure 1) presents a collection of, but is not limited to, the strokes currently in use, stroke titles for identification in this application, sample character(s) for each stroke, and a classification system for all strokes. There are strokes that are not present in the table of Figure 1, but they can easily be treated as variations of the strokes in the list. Additionally, a number of stroke names have the letter "d" preceding them, which refers to "Dian" or "point/dot". The letters "H", "S", "P", or "N" in the titles of compound strokes represent the class name "HENG", "SHU", "PIE" and "NA" respectively.

A review of the types of strokes is useful in understanding the nature of the invention and the advantages it imparts. The stroke LING is recognized as a special

stroke. Ling is unique for both its round shape and for the fact that it serves as a single stroke character. The stroke HENG and stroke SHU are the two fundamental strokes, from which strokes NA and PIE can be derived respectively. All strokes except stroke LING can be classified into four classes, Class HENG, Class SHU, Class PIE and Class NA. Each class can be divided into a Primary Category and a Compound Category. Some compound categories can be further divided into stroke groups, such as Group Left and Group Right, based on certain common graphic features of strokes within the group.

Class HENG includes stroke HENG, stroke dTI, and strokes with titles from H1 to H11 in the stroke table of Figure 1. Primary Category of Class HENG includes stroke HENG and stroke dTI. Compound Category of Class HENG includes strokes with titles from H1 to H11 as shown in Figure 1. All strokes in the Compound Category of Class HENG can be further divided into two stroke groups, Group Left and Group Right, on the basis of how a compound stroke in Class HENG is drawn graphically. Group Left of Class HENG includes strokes with titles from H1 to H7b in Figure 1 and Group Right of Class HENG includes strokes with titles from H8 to H11 in Figure 1. Different versions of Group Left and Group Right of Class HENG are possible.

Class SHU includes stroke SHU, stroke dSHU, and strokes with titles from S1 to S6 in Figure 1. Primary Category of Class SHU includes stroke SHU and stroke dSHU. Compound Category of Class SHU includes strokes with titles from S1 to S6 in Figure 1. All strokes in the Compound Category of Class SHU can be further divided into two stroke groups, Group Left and Group Right, on the basis of how a compound stroke in Class SHU is drawn graphically. Group Left of Class SHU includes strokes with titles from S1 to S4b in Figure 1 and Group Right of Class HENG includes strokes with titles from S5a to S6 in Figure 1. Different versions of Group Left and Group Right of Class SHU are possible.

Class PIE includes stroke PIE, stroke dPIE, stroke P1 and stroke P2 in the Figure 1. Primary Category of Class PIE includes stroke PIE and stroke dPIE while Compound Category of Class PIE includes stroke P1 and stroke P2.

Class NA includes stroke NA, stroke dNA, and stroke N1 where stroke NA and dNA are in the Primary Category and stroke N1 is in the Compound Category of this class. Stroke N1 can also be treated as a variation, or different representation, of stroke NA.

As a result, an "alphabetic" stroke table can be created and a ranking of all strokes can be arranged on the basis of stroke classes, stroke categories within each

class, stroke groups within the compound categories, and the graphic features of individual strokes. Figure 1 also presents one version of such stroke ranking from stroke LING to stroke N1.

5 In this description, the phrase "stroke unit" is used to describe a collection of strokes that typically shares similar graphic characteristics. A single stroke by itself can be a stroke unit. A stroke unit can also be composed of a stroke and its variations. Variations of certain basic strokes are all provided with differential titles in the stroke table in Figure 1.

10 Certain pairs of strokes have essentially the same historic origin. The present day differences in the appearance of these strokes is commonly traced to the relation between the stroke and the adjacent strokes in a character. A sampling of these paired strokes appears in Figure 2. for the purposes of generating stroke ranking table, these strokes can be treated as a single stroke.

15 Among the collection of all strokes, several strokes are more frequently used to form common characters set and therefore, are identified as primary strokes. One example of such common characters set includes those 6763 Chinese characters in the GB2312-80 character internal code table created by the People's Republic of China (PRC) in the early 1980s. These primary strokes include stroke HENG, stroke SHU, stroke PIE, and stroke NA in the stroke table of Figure 1. Each of these primary strokes
20 has a variation or point version (stroke dTI, stroke dSHU, stroke dPIE and stroke dNA respectively). For the purpose of character input in this invention, each primary stroke and its point version are recognized into the same stroke unit and are always arranged in the same keyboard position(s). These primary strokes are shown in Figure 3.

25 Among the primary strokes, stroke PIE and stroke NA, as well as their point versions, stroke dPIE and stroke dNA, are defined as essential strokes. Their keyboard arrangements are critical in designing a well functioning and user friendly stroke based character input method. These essential strokes are shown in Figure 4. The essential stroke unit PIE is composed of stroke PIE and stroke dPIE. Essential stroke unit NA is composed of stroke NA and stroke dNA. Stroke unit HENG is composed of stroke HENG
30 and stroke dTI while stroke unit SHU is composed of stroke SHU and stroke dSHU. All these stroke units of primary strokes and their point versions can also simply be referred to as stroke HENG, stroke SHU, stroke PIE, and stroke NA respectively as each primary stroke is treated as the same as its point version for keyboard arrangements in this invention. These strokes are considered essential as they are not only commonly used

strokes, but often occur in close proximity to each other, and are very often found adjacent to each other. A statistical analysis of a character set indicates that the placement of these strokes is often the stumbling block that results in the awkward use of previous stroke based input systems.

5 The general graphic layout of a particular character can be summarized as:

(1) Each character is composed of at least one single stroke and occupies a rectangular area;

(2) Strokes within a particular character-root or character, can be in different positions horizontally or vertically relative to each other, crossing over
10 each other, surrounded by each other, or be in other positions to each other; and

(3) When a character is composed of both character-root(s) and independent stroke(s), these character-root(s) and independent stroke(s) can be in different positions horizontally or vertically relative to each other, be crossing over each other, or be surrounded, either partially or fully, by each other, or in
15 other positions to each other.

While the way to write a character has been significantly influenced by many factors including the traditional writing tools and writing/printing system, the general rules of character writing can be summarized as "from left to right", "from top to bottom", "from outside to inside", and "from center to symmetrical (left and right) sides". These general
20 rules apply to characters of only a few component strokes as well as to those characters with complex structures. The present invention utilises these conventions in dictating the order in which strokes should be input, so that stroke input is performed in a consistent manner.

In the analysis of the characters and their component strokes, while the
25 relationships among various strokes are examined, those involving two particular strokes, herein referred to as stroke PIE and stroke NA, and their closely related variations, stroke dPIE and stroke dNA respectively, is preferably exploited in order to increase the convenience and efficiency of character input. These two strokes occur next to each other, or in very close proximity to each other, frequently, though the orders of their stroke
30 sequences and relative stroke positions among them vary. Because these strokes occur close together so often, it is advantageous to ensure that they appear on different sides of the keyboard and to avoid one hand being required to move around for consecutive stroke typing when it's not necessary. Furthermore, taking into consideration the different stroke orders and the graphic layouts, i.e. different relative positions, when stroke PIE

and stroke NA do appear next to each other or close to each other, the present invention seeks additional special treatment for these strokes. An optimal placement of the stroke PIE and stroke NA will try to match the keyboard typing of these strokes with the different stroke orders and different relative stroke positions of these strokes within the characters.

5 Because of the existence of different combinations of stroke PIE and stroke NA with different orders and different relative stroke positions, the present invention provides two locations on keyboard for each of the PIE and NA strokes and their respective variations. By providing a first appropriate stroke PIE location for accessibility by the right hand, and a second appropriate stroke PIE location for accessibility by the left hand, and similar first

10 and second locations for the NA stroke, it can be guaranteed not only that the PIE and NA stroke combinations can be entered using alternating hands when it is necessary, but also that different orders of these stroke combinations can be accommodated. Moreover, the placement of each of stroke PIE and stroke NA in two locations on the keyboard allows for convenient input of the strokes when they are not graphically adjacent to each

15 other, but end one line of a character and start the next. As described above, these strokes are considered essential to the writing system due to their frequency with relation to each other. In other ideographic languages the essential characters may vary, and are dealt with in a similar fashion. A detailed analysis of the rest of the keyboard position arrangement, and further information about how the characters are devolved into strokes

20 is described below.

The stroke based character input method of the instant invention is designed to accommodate the relative stroke positions within the characters or character-roots through use of the dual or multiple keyboard positions assigned to essential strokes and through the appropriate placement of other strokes. Since, in combination within a particular

25 character, stroke PIE/dPIE and stroke NA/dNA can have different relative stroke positions to each other with reversing stroke orders and different graphic layouts, and also considering their relative positions to other strokes, these strokes are each assigned to multiple (more than one) keyboard positions in order to simulate the various relative stroke positions within a character or character-root layout.

30 Thus, in the present invention, it is recognized that the relationship between PIE/dPIE and NA/dNA is crucial in the development of an ideographic keyboard for Chinese style characters. Because these strokes often occur next to each other it is advantageous to place the strokes on opposing sides of the keyboard so that they can be accessed by different hands, reducing the amount of unnecessary finger travel required

to input the pairing. Because these strokes and their pairings also occur both before and after a wide assortment of other strokes it is not possible to achieve the objective of input convenience and efficiency with traditional "one stroke over one key" arrangement. To overcome this, each of these strokes is provided in two different and appropriate locations on the keyboard. Thus, if the keyboard is segmented into right and left handed/sided segments, each segment will have one of either PIE or dPIE, and one of either NA or dNA. Thus, any practical combination of a first stroke, followed by a pairing of PIE/dPIE and NA/dNA, can be accomplished using alternating hands when it is necessary. This allows for a reduction in the character input time and reduces both the complexity of the character entry and the learning curve associated with the character entry. The location of each of the strokes, in relation to both each other and the other strokes on the keyboard can be determined using either an analysis of the frequency of the use of the strokes, a division of the keyboard into zones used for different classes of strokes, or other methods that will be known to those skilled in the art.

Figure 5 illustrates a standard QWERTY keyboard **100** which serves as the basis for a presently preferred embodiment of the present invention. One skilled in the art will appreciate that other keyboard layouts, can be just as easily used to practice the present invention, though the key mappings may differ from those indicated in this description. Conventional keyboards are used so that the cost of implementation is reduced, as a simple keyboard interpreter can be provided to convert the standard keyboard signals into the required strokes.

As illustrated in Figure 6, the standard keyboard **100** can be divided into six segments: upper left section **102** (with letters "Q", "W", "E", "R", and "T"), upper right section **104** (with letters "Y", "U", "I", "O" and "P"), middle left section **106** (with letters "A", "S", "D", "F" and "G"), middle right section **108** (with letters "H", "J", "K" and "L"), lower left section **110** (with letters "Z", "X", "C", "V" and "B") and lower right section **110** (with letters "N" and "M"). By convention, the middle left and middle right sections, **106** and **108** respectively, constitute the home row of keyboard positions. The upper left and upper right sections, **102** and **104** respectively, are in the upper row of keyboard positions. The lower left and the lower right sections, **110** and **112** respectively, are in the lower row of keyboard positions. Similar partition of keyboard positions can also be applied to other versions of keyboard and to numeric keypads.

Figure 7 presents an example of the numeric keypad of a hand held calculator and Figure 8 shows an example of the numeric keypad of a telephone set. On both of

these devices, at least three horizontal key groupings can be defined (upper, middle and lower). Each of the horizontal key groupings can be divided into vertical groups as well (right, centre and left).

In this invention, the character input is related primarily to the writing system of characters rather than to the pronunciation of them. Though the number of strokes used to create a frequently used character varies, the average number of strokes used to create an ideographic character using the stroke set defined in Figure 1 lies between 10 and 11.

Figure 9 provides a distribution table of the 6763 characters from the GB2312-80 internal code table over character groups by the number of strokes per character. This shows that it is feasible to use the stroke sequence as the basis to design a character input method. As most of the complex characters can be broken down into several character-roots, writing such a complex character is essentially an exercise of writing several character-roots sequentially. Each root character is also written using sequential component strokes.

Even though 34 different strokes are identified in Figure 1, some strokes are used more frequently than others. Among all strokes, the primary strokes make up about 80% of all strokes used to produce the 6763 characters in the GB2312-80 character set as shown in Figure 10. Therefore, the keyboard arrangement for primary strokes should be given priority over other strokes. Once the keyboard arrangements for primary strokes are determined, other strokes can be assigned to appropriate keyboard positions based on their relationships with the primary strokes and those among themselves.

Among the primary strokes, the treatment of essential strokes is critical because of their diagonal nature. In the common set of Chinese characters, over 90% of the characters can be identified as having at least one of the essential strokes. Also, approximately 50% of the characters contain one or more combination of essential strokes. The ideal objective in designing a stroke based character input method is to map the strokes over keyboard positions in such a way that it will make the typing/input easy, convenient and efficient. One major difficulty associated with arranging the essential strokes to achieve this objective is the existence of different essential stroke combinations, which represent different relative positions of essential strokes within the characters.

There are at least four groups of essential stroke combinations. Group One essential stroke combinations are essential stroke(s) NA (or dNA) horizontally followed

by essential stroke(s) PIE (or dPIE). These Group One combinations usually appear at the top of a character. Group Two essential stroke combinations are essential stroke(s) PIE (or dPIE) horizontally followed by essential stroke(s) NA (or dNA). Group Three essential stroke combinations are essential stroke(s) PIE (or dPIE) vertically followed by essential stroke(s) NA (or dNA). A Group Three combination usually appears on the right side of a character. Group Four essential stroke combination includes those that essential stroke PIE (dPIE) is crossed over by essential stroke NA (dNA).

The different stroke orders and different relative stroke positions of individual essential strokes like these among different groups of essential stroke combinations are the primary reasons why a simple stroke sequence based character input method is inconvenient and difficult to use in absence of appropriate keyboard treatment. An essential stroke combination in Group One starts with stroke NA (or dNA) and horizontally finishes with stroke PIE (or dPIE) while a essential stroke combination in Group Two starts with stroke PIE (or dPIE) and horizontally finishes with stroke NA (or dNA). The orders and graphic layout of strokes in Group Three and Group Four essential stroke combinations are also unique and different from those of essential stroke combinations of other groups. All these unique and different stroke orders and graphic layouts of strokes represent different relative positions among essential strokes. Without giving adequate and appropriate consideration to these stroke relationships, it is difficult to design a stroke sequence based character input method that can conveniently handle these strokes of different combinations, not to mention the fact that there exists a list of other non- essential strokes whose relative positions within a character also affect the convenience and efficiency of the input process.

These different essential stroke combinations can be treated as various fixed character roots and their graphic appearances and relative stroke positions are preferably used as the basis to determine the keyboard positions of individual essential strokes. Traditional simple stroke sequence based character input methods assign each stroke to one single keyboard position. Under such an arrangement, not all the different stroke orders and relative stroke positions of various essential stroke combinations are accommodated for. If the stroke orders of Group One essential stroke combinations are accommodated, the stroke orders of Group Two essential stroke combinations cannot be. When the vertical relative stroke positions of Group Three essential stroke combinations are matched by keyboard positions of essential strokes, it is impossible to accommodate the horizontal relative stroke positions of other groups of essential stroke

combinations. As a result, it is inconvenient and often awkward to use the same keyboard position to input an essential stroke when that essential stroke is associated with different essential stroke combinations of reversing stroke orders and of different relative stroke positions. Even though a character input method can be based solely and simply on the stroke sequences with strokes each being assigned to one single keyboard position, such methods are often inconvenient and inefficient to use when compared to pronunciation based and character root based input methods. Therefore, the issue of input convenience and efficiency can be solved by accounting for relative stroke positions in all groups.

The existence of different stroke orders and relative stroke positions among essential stroke combinations calls for unique keyboard arrangement for these essential strokes in order to create a connection between the keyboard arrangement of strokes and the relative stroke positions within characters. The present invention solves this issue, as described above, by a dual position arrangement for essential stroke strokes. On a keyboard, four particular positions, referred to as Position One, Position Two, Position Three and Position Four, are identified and reserved for these essential strokes. Position One and Position Two are preferably on the same horizontal line of the keyboard, with Position One on the left side of the keyboard and Position Two on the right side of the keyboard. Position Three and Position Four are preferably on the same row of the keyboard with Position Three on the left side of the keyboard, and Position Four on the right side of the keyboard. The keyboard row of Position Three and Position Four is preferably below the line of Position One and Position Two.

Corresponding to the fact that essential stroke combinations can be in reversed orders, and in order to make it possible to match the keyboard arrangements of individual strokes with the various relative stroke positions of these essential stroke combinations, each of the essential stroke units is assigned to two keyboard positions. Particularly, the essential stroke unit NA is preferably assigned to both Position One and Position Four, while essential stroke unit PIE is preferably assigned to both Position Two and Position Three. As such, each individual essential stroke is associated with two different keyboard positions. Different combinations of these positions can then accommodate the different stroke orders and different relative positions among these essential strokes. With this unique arrangement, the user can choose the appropriate keyboard position of a particular essential stroke to use depending on the nature of the target essential stroke combinations, i.e., the order and relative positions among essential strokes. If a Group One essential stroke combination is the target combination, Position One and Position

Two can be used to input the necessary essential stroke sequences. To input an essential stroke sequence of Group Two and Group Four, Position Three and Position Four can be used. If the Group Three essential stroke combination is the target, the stroke sequence can be processed using Position Two and Position Four. Position Three and Position Four can also be used to input Group Three essential stroke combinations if only the orders of stroke sequences are considered. Even though either of the two keyboard positions reserved for a particular essential stroke could be used to input that essential stroke regardless of its relative positions within the character, the selection of an appropriate keyboard position corresponding to that essential stroke's current relative position to other strokes within the character will make the input process more natural and more efficient.

This principle of matching the keyboard arrangement of strokes with their relative positions within a character can also be applied in situations where sequenced strokes with specific orders or relative positions in a essential stroke combination are recognized as one stroke unit, or when the character root is the basis of a character input method. In so doing, these stroke units or character roots are assigned to appropriate keyboard positions based on the relative stroke positions they represent. In one embodiment of such arrangement, stroke units or character roots of Group One essential stroke combinations can be assigned to either Position One or Position Two on the keyboard. And stroke units or character roots of Group Two and Group Four essential stroke combinations can be assigned to either Position Three or Position Four. The stroke units or character roots of Group Three essential stroke combinations can either be assigned to Position Two or Position Four, or be treated the same as those stroke units of Group Two and Group Four essential stroke combinations.

When applied to the standard QWERTY keyboard where six different keyboard sections are identified as illustrated in Figure 6, this dual position keyboard arrangement results in the following types of combinations of essential stroke positions and keyboard sections. Type One: Position One for essential stroke NA/dNA is in upper left section **102**, Position Two for essential stroke PIE/dPIE in upper right section **104**, Position Three for essential stroke PIE/dPIE in middle left section **106** and Position Four for essential stroke NA/dNA in middle right section **108**. Type Two: Position One for essential stroke NA/dNA is in upper left section **102**, Position Two for essential stroke PIE/dPIE in upper right section **104**, Position Three for essential stroke PIE/dPIE in lower left section **110** and Position Four for essential stroke NA/dNA in lower right section **112**. Type Three: Position

One for essential stroke NA/dNA is in middle left section **106**, Position Two for essential stroke PIE/dPIE in middle right section **108**, Position Three for essential stroke PIE/dPIE in lower left section **110** and Position Four for essential stroke NA/dNA in lower right section **112**. The selection of a particular one of these three types of combinations for implementation will be decided by the relative positions of essential strokes to other non-essential strokes, especially to other primary strokes.

Because of the horizontal nature of the stroke HENG, it preferably appears either right beneath or above the essential stroke combinations in the graphic layout of a character. Also, when a single essential stroke precedes or follows the stroke HENG, this individual essential stroke could be above, beneath, crossing over the stroke HENG. All these situations suggest a keyboard position for stroke unit HENG on a horizontal line of keyboard positions that is in between the two horizontal lines of keyboard positions assigned for essential strokes or essential stroke units.

In the case of relative positions between essential stroke units and stroke SHU, the stroke SHU is preferably next to essential stroke combinations of either Group One or Group Two in the stroke sequence of a character. This suggests a keyboard position for stroke unit SHU on a horizontal line of keyboard positions that is in between the two lines of the keyboard positions assigned for essential strokes or essential stroke units.

As shown in Figure 10, stroke HENG/dHENG and stroke SHU/dSHU are the most frequently used strokes to form common characters set. Their graphic appearances and their various relative positions to each other and to other strokes, especially to essential strokes, within characters suggest that the keyboard position for stroke SHU be on the left side of the keyboard position for stroke HENG.

Taking into account the relative stroke positions among all primary strokes, the above analysis results in a Type Two combinations of essential stroke positions and keyboard sections such that Position One for essential stroke unit NA is preferably in the upper left section **102** of the keyboard, Position Two for essential stroke unit PIE is in the upper right section **104** of the keyboard, Position Three for essential stroke unit PIE is in the lower left section **110** of the keyboard and Position Four for essential stroke unit NA is in the lower right section **112** of the keyboard. In the meantime, Stroke unit HENG is assigned in the middle right section of the keyboard and stroke unit SHU is assigned to the middle left section of the keyboard.

One implementation this dual position arrangement of primary strokes laid over the standard QWERTY keyboard is shown in Figure 11. Keyboard **200** provides

essential stroke unit PIE located in the positions typically occupied by the letter "I" in the upper right section **104** and "V" in the lower left section **110**. Essential stroke unit NA is located in the positions typically occupied by the letter "E" in the upper left section **102** and the letter "M" in the lower right section **112**. Stroke unit HENG occupies the position of letter "J" in the middle right section **108** while stroke unit SHU is assigned to the position of letter "F" in the middle left section **110**.

Once the keyboard positions for primary strokes, which are the most frequently used strokes to form characters, have been decided, the arrangements of other non-primary strokes on the keyboard are then preferably determined both by their relative positions to primary strokes, the relative positions among themselves, and the number of keyboard positions available. On reduced format keyboards where there are not enough appropriate keyboard positions for each of all the non-primary strokes, more non-primary strokes can be grouped together to form stroke units or be grouped into the stroke units of primary strokes to reduce the number of keyboard positions required. This also makes it possible to apply the present invention to other simplified versions of keyboard that may have fewer key positions than the standard QWERTY keyboard, or to numeric keypads of various electronic appliances.

Figure 12 presents one version of the arrangement for all strokes on keyboard **200**, which takes into consideration the relative positions among the various strokes. Figure 13 is a tabular summary of such a keyboard arrangement for all strokes and sample characters that contain the relative stroke relationships concerning each of the particular strokes. Under this version of keyboard arrangement, there exist some keyboard positions not assigned with any strokes while some other keyboard positions are each assigned with multiple strokes. This is due to the fact that the relative stroke position within characters is the most important fact in determining the keyboard positions of these strokes.

As a direct result of this unique keyboard arrangement of strokes, the overall level of convenience and efficiency of keyboard typing is improved significantly as indicated by the reduced distance that each hand has to travel over the keyboard, the low frequency that one particular hand is required to type continuously over different keyboard positions, the low frequency that one particular finger is required to type continuously over different keyboard positions, the high frequency that the home row of the keyboard is used relative to the top row and the bottom row of the keyboard, the overall continuity of stroke typing, and high finger coordination.

Figure 14 also presents an embodiment of the invention on a different configuration keyboard **202** which is used by many handheld computers and Personal Digital Assistants (PDAs). While the overall horizontal line-up of strokes are not changed, the assignment of strokes to keyboard positions on the upper row and the lower row of the keyboard is modified in response to a different positioning of keys.

The present invention can also be used to apply stroke based character input to Numeric Keypads. In applying the present invention to reduced key keyboard, or numeric keypads, more individual strokes have to be grouped into the same stroke units in order to accommodate the fact that now there are significantly less key positions. In one such embodiment, all strokes in Class PIE are grouped into one stroke unit and identified as Class unit PIE. All strokes in Class NA are grouped into one stroke unit and identified as Class unit NA. All strokes in the Compound Category of Class HENG are grouped into one stroke unit and identified as Compound unit HENG. All strokes in the Compound Category of Class SHU are grouped into one stroke unit and identified as Compound unit SHU. Primary unit HENG is composed of stroke HENG and stroke dTI while Primary unit SHU is composed of stroke SHU and stroke dSHU. Since primary strokes make up about 80% of all strokes used to produce common characters, this arrangement with reduced number of stroke units will not significantly increase the number of character groups with duplicating input codes.

As there are now only six stroke units in this arrangement, they can be assigned to the numeric keypads of any electronic appliances with the application of the dual position arrangement of the invention. Two positions will be reserved for each of the Class unit PIE and Class unit NA so the user can select the appropriate position to use for a particular class unit depending on that class unit's relative position to other strokes or stroke units within a character or character root.

One exemplary numeric keypad arrangement is shown in Figure 15. On keypad **204** Class unit NA is placed in both the upper left **102** and upper right **104** sections of keyboard **204** and is preferably assigned to the upper left and the lower right keys. Class unit PIE is placed in both upper right **104** and lower left **110** sections of the keyboard **204** is preferably assigned to the upper center and lower left keys. Primary unit HENG is placed in the middle right section **108** and is preferably assigned to the middle center position while Primary unit SHU placed in the middle left section **106** and is assigned to the middle left position. The middle right position is preferably reserved for Compound unit HENG and the lower center position is preferably reserved for Compound unit SHU.

The upper right position is reserved for stroke LING or for other special use. As LING is considered to be a special character, one skilled in the art will appreciate that to indicate the end of the entry of a character a user may be provided with either an end of character key, or may use the LING stroke to indicate the end of the character, thus LING can be thought of as a special case of a non-stroke key. As LING does not appear as a part of a character, an interpreter can determine whether the use of LING indicates a character, or represents the end of character signal.

This simplified version of the present invention can also be applied to the QWERTY keyboard such that only a portion of the keyboard will be required. In one exemplary arrangement on keyboard **206** is depicted in Fig. 16, Class unit PIE is assigned to keyboard positions of "Y" and "V", Class unit NA is assigned to keyboard positions of "T" and "N", Primary unit HENG is assigned to the keyboard position "H", Compound unit HENG is assigned to keyboard position "J", Primary unit SHU is assigned to the keyboard position "G", and Compound unit SHU is assigned to keyboard position "B". Generally, on a keyboard, four particular positions on two different horizontal lines of keyboard positions, the upper line and the lower line, are identified with two positions, the left position and the right position, specified on each line. The left position on the upper line and the right position on the lower line are reserved for essential stroke unit NA (stroke NA and stroke dNA). The right position on the upper line and the left position on the lower line are reserved for essential stroke unit PIE (stroke PIE and stroke dPIE).

Two essential strokes, such as PIE and dPIE, or NA and dNA, can share the same multiple keyboard positions. In cases where a essential stroke is recognized in the same stroke unit with a non-essential stroke, this stroke unit containing the essential stroke or essential strokes can be assigned to multiple keyboard positions. In so doing, the user can select the appropriate keyboard position for a particular essential stroke depending on that essential stroke's relative position to other strokes within a particular character or character root.

In one embodiment of the present invention four particular positions are provided on a keyboard for stroke PIE/dPIE and stroke NA/dNA units, two positions being located on a different horizontal line of keys than the remaining two positions. In a particular embodiment, one of the two horizontally higher positions is designated to the stroke PIE/dPIE and the other to stroke NA/dNA. And also one of the two horizontally lower positions is designated to the stroke PIE/dPIE and the other to stroke NA/dNA.

In embodiments where stroke PIE and stroke dPIE are recognized as two distinct strokes, stroke PIE can be assigned to multiple different keyboard positions while stroke dPIE can also be assigned multiple different keyboard positions. The multiple keyboard positions reserved for stroke PIE can overlap the multiple keyboard positions reserved for stroke dPIE.

In embodiments where stroke NA and stroke dNA are recognized as two distinct strokes, stroke NA can be assigned to multiple different keyboard positions while stroke dNA can also be assigned multiple different keyboard positions. The multiple keyboard positions reserved for stroke NA can overlap the multiple keyboard positions reserved for stroke dNA.

The principle of this dual/multiple keyboard arrangement under the present invention is based upon stroke recognition/classification systems or stroke tables. This allows the relationship between strokes to be accounted for in the placement of the strokes on the keyboard. Such relative stroke positions exist universally among characters and are independent of the way strokes are classified. In one conventional stroke classification system, stroke dSHU, stroke dPIE and stroke dNA in Figure 1 are recognized under one composite stroke "Dian". Therefore, the stroke classification in Figure 1 is essentially an extension of such conventional stroke recognition in the sense that stroke dSHU, stroke dPIE, stroke dNA are recognized differentially in Figure 1. However, even under the conventional stroke recognition system where a composite stroke "Dian" is recognized, the present invention can be implemented to achieve the objective of ease of learning, convenience and efficiency to use.

An ordinal ranking of strokes can be arranged at different levels of stroke classes, stroke categories, stroke groups, or at the individual stroke level. These arrangements may be used either alone or in combination. The examples below are based on the placement of strokes in an alphabetic stroke table, such as that shown in Figure 1.

In an ordinal alphabetic stroke table, Stroke LING is ranked ahead of all the other strokes and stroke classes. Class HENG is ranked ahead of Class SHU, Class SHU in turn is ranked ahead of Class PIE, and Class PIE in turn is ranked ahead of Class NA. In an alphabetic stroke table, within each stroke class where all strokes are divided into a Primary Category of strokes and a Compound Category of strokes, the Primary Category of strokes is ranked ahead of the Compound Category of strokes. In an alphabetic stroke table, within a Compound Category of strokes where all strokes are divided into a stroke Group Left and a stroke Group Right, the stroke Group Left is ranked ahead of the stroke

Group Right. In an alphabetic stroke table, all strokes in the stroke Group Left in Class HENG are ranked by the order of stroke H1, stroke H2, stroke H3, stroke H4, stroke H5, stroke H6a and stroke H6b, stroke H7a and stroke H7b. All strokes from the stroke Group Right in Class HENG are ranked by the order of stroke H8, stroke H9, stroke H10a and stroke H10b, and stroke H11. In an alphabetic stroke table, all strokes from the stroke Group Left in Class SHU are ranked by the order of stroke S1, stroke S2, stroke S3, stroke S4a and stroke S4b. All strokes from the stroke Group Right in Class SHU are ranked by the order of stroke S5a and stroke S5b, and stroke S6. In the alphabetic stroke table, all strokes in the Compound Category of Class PIE are ranked by the order of stroke P1, and stroke P2. Within the primary categories of each stroke class, the point version of a particular primary stroke can either be recognized as the same rank as its corresponding primary stroke or be ranked after each of its corresponding primary stroke. In one implementation, stroke dTI is ranked after stroke HENG, stroke dSHU is ranked after stroke SHU, stroke dPIE is treated as the same rank as stroke PIE and stroke dNA is treated as the same rank as stroke NA.

With such a stroke ranking in place, a list of characters can be arranged in the order of their component stroke values. Such an ordinal list of characters can be used in various occasions where an ordinal list of characters is necessary, such as character lists in dictionaries. For those characters having same component stroke sequences, other language features, such as frequency of usage or pronunciation, can be used to sequentially list these characters. In one implementation of the present invention where a database of listed characters has to be prepared, the order of characters in the database can be arranged by their component stroke values.

There is a plurality of methods of implementing the present invention. Among these embodiments is a Latin alphabetic keyboard or a simplified numeric keypad, wherein the strokes are overlaid on the keys, as described above. The keyboard generates an output signal in response to the depression of a key, and the output signal is provided to a stroke interpreter. The stroke interpreter receives the output signals that are representative of the desired stroke (or stroke unit) and stores the strokes until it receives an output signal indicative of the end of a character or character input. The stroke interpreter combines all the received strokes to produce a character. This implementation is preferably designed to promote entry of a character through its devolution into component strokes, entered in the order in which a character is usually written. Optionally there is a conflict resolution module in the stroke interpreter to

determine if more than one character is mapped to a particular sequence of strokes. If the conflict resolution module detects a conflict, the candidate characters are provided to the user for selection to avoid inadvertently providing the wrong character. The keyboard having the dual positioning/mapping of the PIE/dPIE and NA/dNA strokes is additionally
5 able to be implemented as a stand alone product in addition to its ability to be implemented in the combination of keyboard and interpreter system. In an alternate embodiment, the stroke interpreter narrows the possible characters with each progressive stroke, so as to provide a form of predictive text input. An optional feature of this embodiment is that the possible characters are displayed on a portion of the screen as
10 the strokes are entered, so that the user can use a non-stroke key, or an alternate entry device, such as a mouse, to select the intended character without having to complete all the stroke entry of the character.

One skilled in the art will appreciate that the present invention can be implemented as either a keyboard having built in logic to process the stroke keys and
15 buffer them until the user indicates that the end of a character has been reached by using a non-stroke key. Alternatively, the invention can be implemented using a standard computing platform that receives input signals from a keyboard, each signal representative of a stroke associated with a key. A keystroke interpreter buffers the strokes and can narrow down the selection of characters stored in a database with each
20 stroke. This allows the user to type part of a character and then select the character from a pick list, or for the interpreter to employ predictive text input. In another embodiment a conflict resolution module provides the user with an indication that the strokes input match more than one known character. This preferably is performed by providing the user with a list of possible choices on a computer screen, and allowing the user to select
25 the desired character from a pick list.

The above described embodiments of the present invention are intended to be examples only. Alterations, modifications and variations may be effected to the particular
30 embodiments by those of skill in the art without departing from the scope of the invention, which is defined solely by the claims appended hereto.

What is claimed is:

1. A keyboard, for the entry of ideographic characters based on component strokes, comprising:
 - a plurality of stroke keys, each representing a component stroke, at least two of
 - 5 the plurality of stroke keys representing the same component stroke, each key for providing a signal representative of the stroke associated with the key; and
 - a non-stroke key, for providing a signal representative of the end of the entry of a character.
2. The keyboard of claim 1, wherein the plurality of stroke keys includes two sets of
- 10 two stroke keys, keys of the first of the two sets representing a first component stroke, and the keys of second of the two sets representing a second component stroke.
3. The keyboard of claim 2, wherein the keyboard is arranged to form two contiguous regions, horizontally adjacent to each other, the regions being a right and left sided region.
- 15 4. the keyboard of claim 3, wherein the keys of the first and second set are positioned so that one of the keys of each set is in the right sided region, and the other key of each set is in the left sided region.
5. The keyboard of claim 1, wherein the ideographic characters are Chinese characters, and the at least two of the plurality of stroke keys representing the same
- 20 component stroke represent a stroke selected from a list including stroke PIE, stroke dPIE, stroke NA and stroke dNA .
6. The keyboard of claim 1, including a plurality of non-stroke keys for providing a signal representative of the end of the entry of a character.
7. The keyboard of claim 1, wherein the keyboard has the layout of a QWERTY
- 25 keyboard and one of the at least two of the plurality of stroke keys representing the same component stroke resides in the row above the home row.
8. The keyboard of claim 7, wherein the other of the at least two stroke keys representing the same component stroke resides in the row below the home row.
9. The keyboard of claim 1, further including:

a stroke interpreter, operatively connected to the plurality of stroke keys and to the non-stroke key, for receiving and buffering the provided signals, and for selecting an ideographic character from a database based on the buffered signals representative of strokes when a signal representative of the end of the entry of a character is received.

- 5 10. The keyboard of claim 9, wherein the stroke interpreter includes means for both narrowing the number of characters for selection with each stroke; and selecting a character from the narrowed number of characters on the basis of a signal representative of one of the plurality of non-stroke keys.
- 10 11. The keyboard of claim 9, further including a conflict resolution module, operatively connected to the database and the stroke interpreter, for selecting an ideographic character from the database when the buffered signals are not uniquely associated with one character, the selection based on a signal representative of one of plurality of non-stroke keys received in response to a presentation of the characters in the database associated with the buffered signals.
- 15 12. The keyboard of claim 1, wherein the keyboard has the layout of a QWERTY keyboard and provides key mappings for stroke HENG/dTI, stroke SHU/dSHU, stroke PIE/dPIE and stroke NA/dNA as illustrated in Figure 11.
13. The keyboard of claim 1, wherein the keyboard has the layout of a QWERTY keyboard and provides key mappings as illustrated in Figure 12.
- 20 14. The keyboard of claim 1, wherein the keyboard has the layout of a QWERTY keyboard and provides key mappings as illustrated in Figure 14.
15. The keyboard of claim 1, wherein the keyboard has the layout of a numeric keypad, the keypad organised into three columns of three rows, the keypad having key mappings as illustrated in Figure 15.
- 25 16. The keyboard of claim 1, wherein the keyboard has the layout of a QWERTY keyboard and provides key mapping as illustrated in Figure 16.
17. The keyboard of claim 9, wherein the database contains a plurality of characters indexed according to the component stroke values in a sequence associated with each of the plurality of characters, the component stroke values assigned in accordance with the relative ranking values of strokes illustrated in Figure 1.
- 30

	Stroke Classification	Strokes	Stroke Titles	Sample Characters						
1	LING	○	LING	○						
2	Class HENG	Primary Category	—	HENG(TI)	一	二	三	七	万	
3			√	dTI	冰	江	河	湖	泊	
4		Compound Category	Group Left	7	H1	又	夕	久	冬	外
5				一	H2	欠	次	子	予	
6				3	H3	阴	阳	郭	那	
7				3	H4	建	延	及		
8				3	H5	扬	汤	场		
9				7	H6a	刀	力	刃		
10				7	H6b	口	国	日		
11				3	H7a	乃	仍	孕		
12				3	H7b	凸				

Figure 1A

	Stroke Classification			Strokes	Stroke Title	Sample Characters					
13	Class HENG	Compound Category	Group Right	乙	H8	乙	艺	忆			
14				飞	H9	飞	风	执			
15				凹	H10a	凹					
16				九	H10b	九	几				
17				话	H11	话	诒	谄			
18	Class SHU	Primary Category		十	SHU	十	上	个	五		
19				它	dSHU	它	宁	穴	农	冗	
20		Compound Category		Group Left	丁	S1	丁	了	小	寸	承
21					狗	S2	狗	狼	豕		
22					专	S3	专	传			
23					亏	S4a	亏	与	弓	弗	马
24					鼎	S4b	鼎	鼐			
25					Group Right		Group Right	山	S5a	山	亡
26		七	S5b	七				儿	己		
27		以	S6	以				比	瓦	叫	纠

Figure 1B

	Stroke Classification	Strokes	Stroke Title	Sample Characters					
28	Class PIE	Primary Category	丿	PIE	人	八			
29			㇀	dPIE	照	惊			
30		Compound Category	㇀	P1	女	巡			
31			㇁	P2	丝	红	允	乡	
32	Class NA	Primary Category	㇂	NA	人	入	八	衣	
33			㇃	dNA	凉	洪	六	照	兴
34		Compound Category	㇄	N1	弋	戈	我		

Figure 1C

4/19











Pairs	Strokes (Graphic Representation One)	Stroke Titles	Strokes (Graphic Representation Two)	Stroke Titles
1		H6a		H6b
2		H7a		H7b
3		H10a		H10b
4		S4a		S4b
5		S5a		S5b

Figure 2









	Primary Strokes	Stroke Titles	Primary Strokes (Point Version)	Stroke Titles
1		HENG(TI)		dTI
2		SHU		dSHU
3		PIE		dPIE
4		NA		dNA

Figure 3





	Key-strokes	Stroke Titles	Key-strokes (Point Version)	Stroke Titles
1		PIE		dPIE
2		NA		dNA

Figure 4

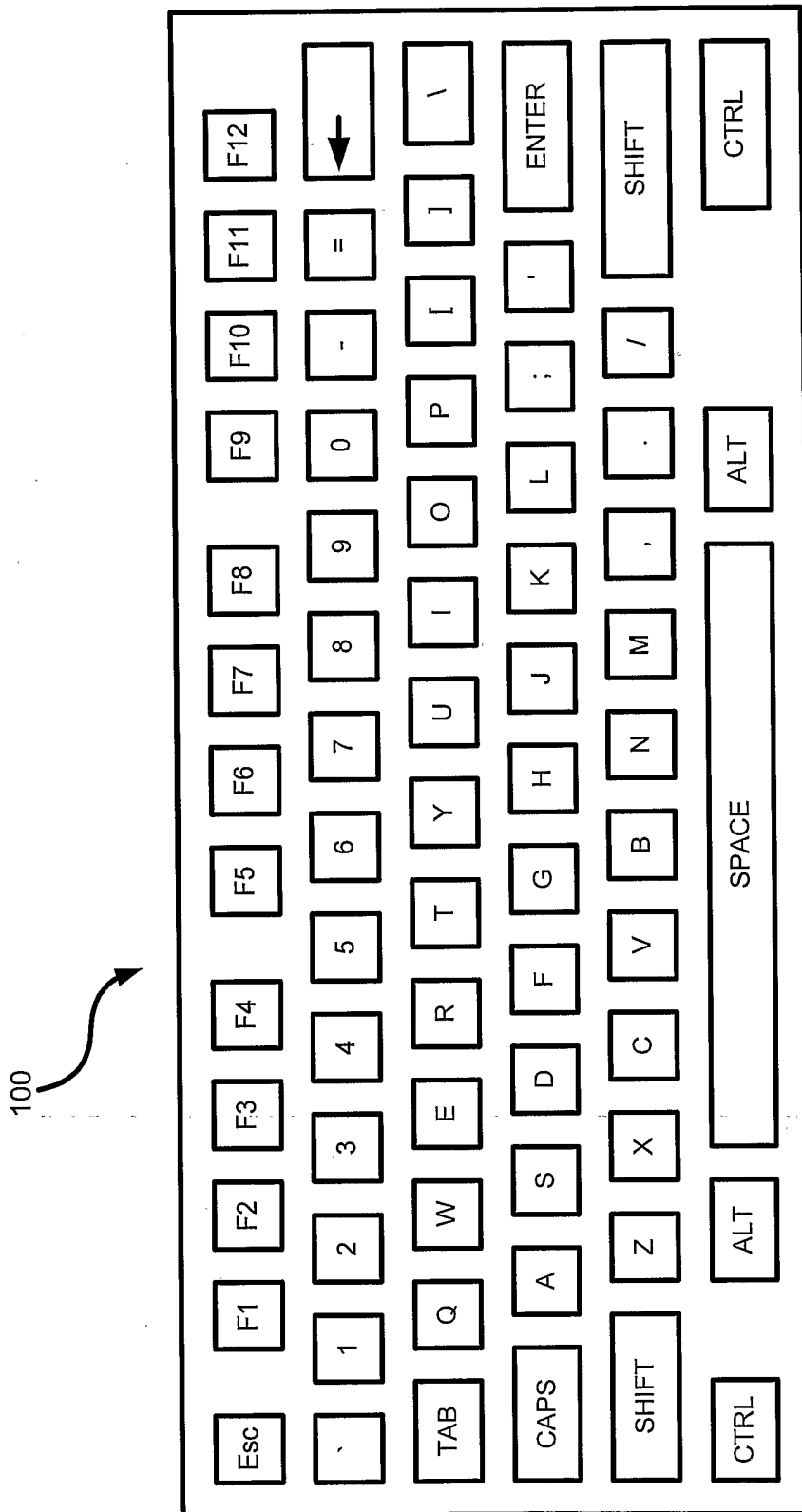


Figure 5 (prior art)

7/19

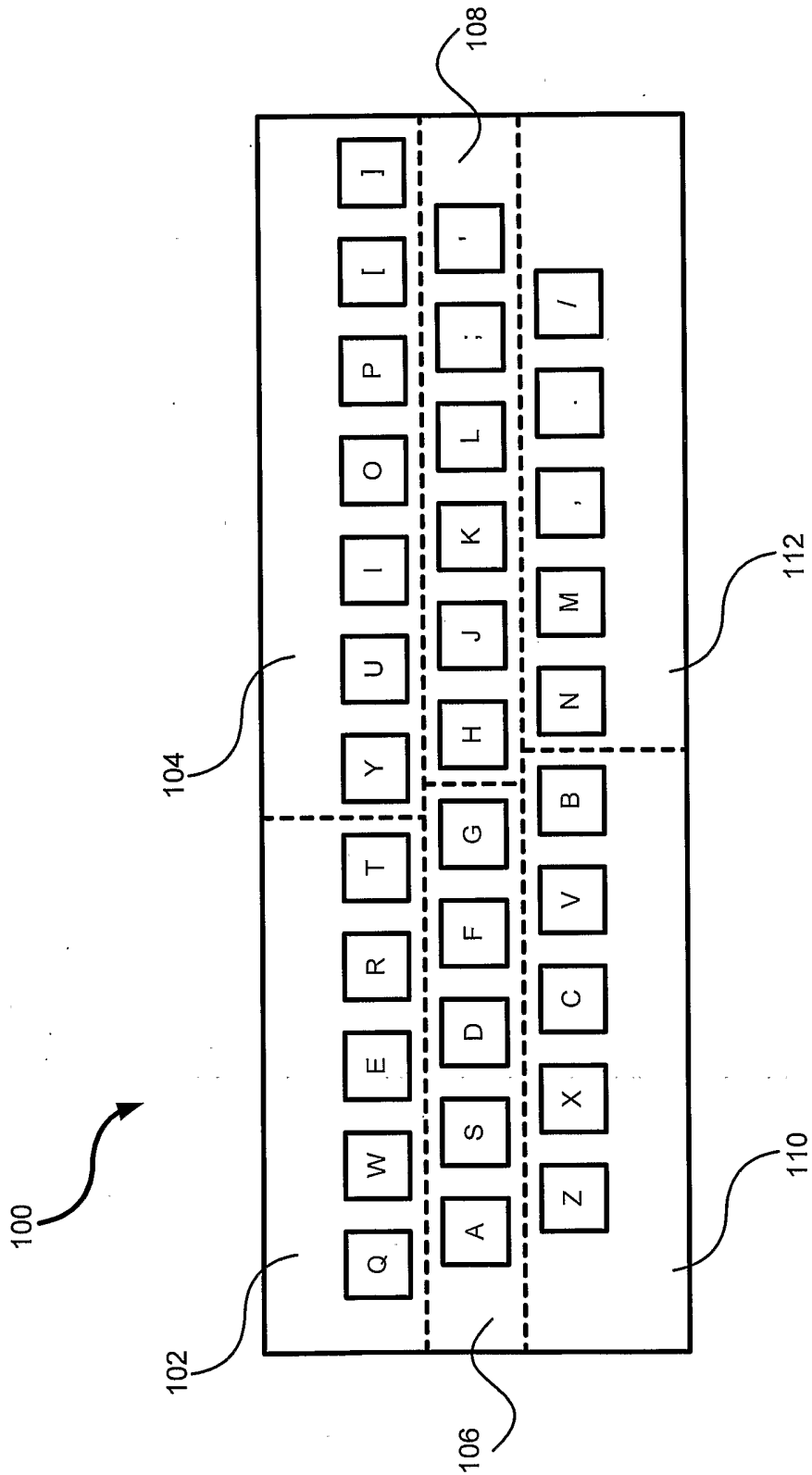


Figure 6 (prior art)

8/19

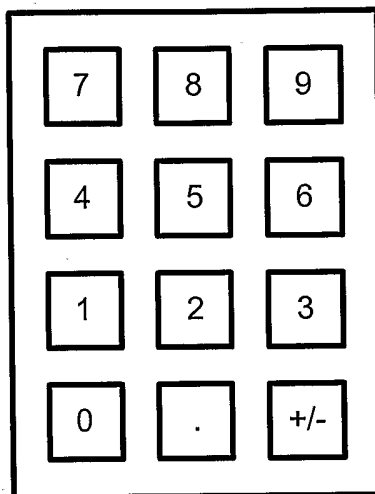


Figure 7 (prior art)

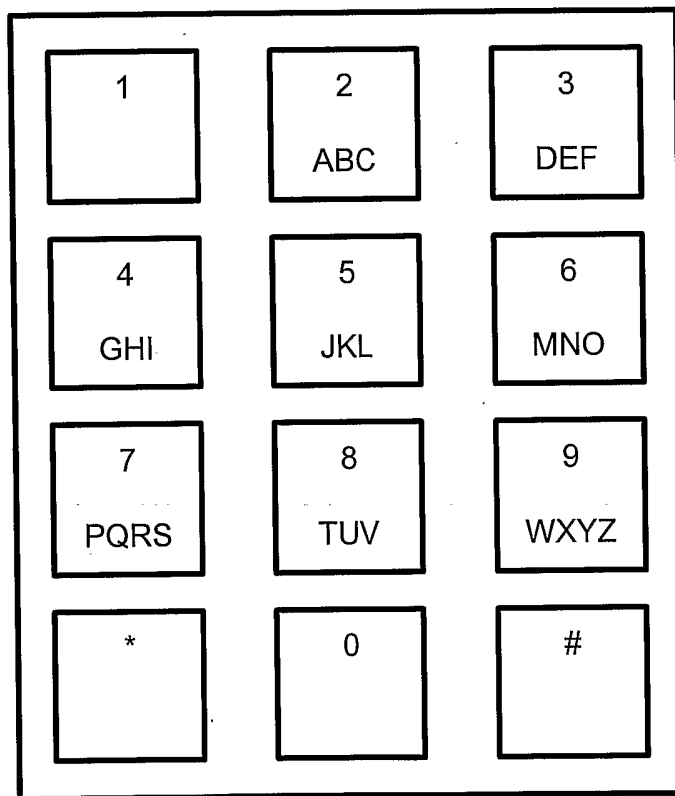


Figure 8 (prior art)

9/19

		Character Distribution		
		# of Characters per Group	% of Characters per Group	Cumulative % of Characters per Group
Character Groups by Number of Strokes per Character	1	5	0.07%	0.07%
	2	34	0.50%	0.58%
	3	79	1.17%	1.74%
	4	140	2.07%	3.81%
	5	203	3.00%	6.82%
	6	349	5.16%	11.98%
	7	531	7.85%	19.83%
	8	677	10.01%	29.83%
	9	752	11.12%	40.95%
	10	730	10.79%	51.74%
	11	702	10.38%	62.12%
	12	637	9.42%	71.54%
	13	524	7.75%	79.29%
	14	397	5.87%	85.16%
	15	311	4.60%	89.75%
	16	255	3.77%	93.52%
	17	174	2.57%	96.10%
	18	73	1.08%	97.18%
	19	77	1.14%	98.31%
	20	45	0.67%	98.98%
	21	23	0.34%	99.32%
	22	17	0.25%	99.57%
	23	13	0.19%	99.76%
	24	7	0.10%	99.87%
	25	6	0.09%	99.96%
26	1	0.01%	99.97%	
30	1	0.01%	99.99%	
	Total	6763	100%	

Figure 9

10/19

		Distribution and Frequency of Strokes											
Strokes	HENG /dTI	H1	H2	H3	H4	H5	H6a /b	H7a /b	H8	H9	H10a /b	H11	
Character Groups by Number of Strokes	1	1							1				
	2	10	1	1	1	1		8	1		2	1	
	3	58	5	4		1		19		2	1	2	
	4	122	15	5	2			42	1	2	3	7	5
	5	259	15	11	6			104	6	2	2	7	10
	6	557	25	15	15	8	4	167	1	7	8	16	14
	7	988	39	47	26	1	3	336	1	1	6	16	17
	8	1529	84	52	18	2	3	532		1	6	18	23
	9	2001	83	48	22	6	2	623	2	2	8	8	15
	10	2172	70	52	19	3	1	665	4		8	7	20
	11	2450	76	75	19	1	1	729		1	4	6	19
	12	2447	75	57	9	8	1	746	1		7	5	10
	13	2219	80	50	7	1		673	3		2	14	7
	14	1751	50	42	4	1		560	1		1	4	8
	15	1503	37	47	1	1		466	1		2	5	3
	16	1328	29	46	5			404			2	8	2
	17	924	27	35	1			307				6	
	18	425	15	11				135				2	1
	19	490	12	18	1			139				3	2
	20	304	3	7	1			93				1	
21	174	5	1				57						
22	128		2				39				1		
23	101	2					27						
24	68	2					15						
25	44	1	1				18						
26	12	1					3						
30	9						4						
	Total	22074	752	627	157	34	15	6911	22	19	60	138	157
	%	30.73%	1.05%	0.87%	0.22%	0.05%	0.02%	9.62%	0.03%	0.03%	0.08%	0.19%	0.22%

Figure 10A

11/19

		Distribution and Frequency of Strokes (continued)											
Strokes	SHU/ dSHU	S1	S2	S3	S4a /b	S5a /b	S6	PIE/ dPIE	P1	P2	NA/ dNA	N1	Total
1	1							1			1		5
2	9	3				6		13		1	10		68
3	31	9	1		4	10	1	41	4	7	36	1	237
4	71	10		1	4	26	6	128		7	101	2	560
5	183	23	2		8	33	9	171	2	17	140	5	1015
6	337	57	3	1	18	71	13	387	16	40	304	10	2094
7	620	82	8	1	17	128	41	674	19	58	579	9	3717
8	960	108	7	1	32	146	49	867	22	97	836	23	5416
9	1208	121	13	1	31	146	66	1172	27	102	1029	32	6768
10	1244	117	20		36	157	80	1290	32	93	1190	20	7300
11	1386	126	17	1	25	158	72	1246	34	90	1171	15	7722
12	1441	122	15		25	134	64	1207	23	93	1131	23	7644
13	1262	74	11		22	116	62	1070	25	71	1033	10	6812
14	1049	74	7		15	89	45	888	22	57	882	8	5558
15	902	68	10		6	67	30	768	8	36	699	5	4665
16	783	35	6		5	55	27	678	12	39	612	4	4080
17	579	29	15		9	43	21	482	3	25	446	6	2958
18	248	11	3		4	13	12	214	4	5	207	4	1314
19	303	15	3		2	23	10	222	1	7	211	1	1463
20	193	6	2		2	14	11	130	2	8	120	3	900
21	100	4	1			4	4	65			68		483
22	78		1		4	4	10	51	3	2	50	1	374
23	62	3				4	9	44	1	5	40	1	299
24	35	2						22			24		168
25	30	3				2	6	18	3	3	20	1	150
26	6							1			3		26
30	6							6			5		30
Total	13127	1102	145	6	269	1449	648	11856	263	863	10948	184	71826
%	18.28%	1.53%	0.20%	0.01%	0.37%	2.02%	0.90%	16.51%	0.37%	1.20%	15.24%	0.26%	100.0%

Figure 10B

12/19

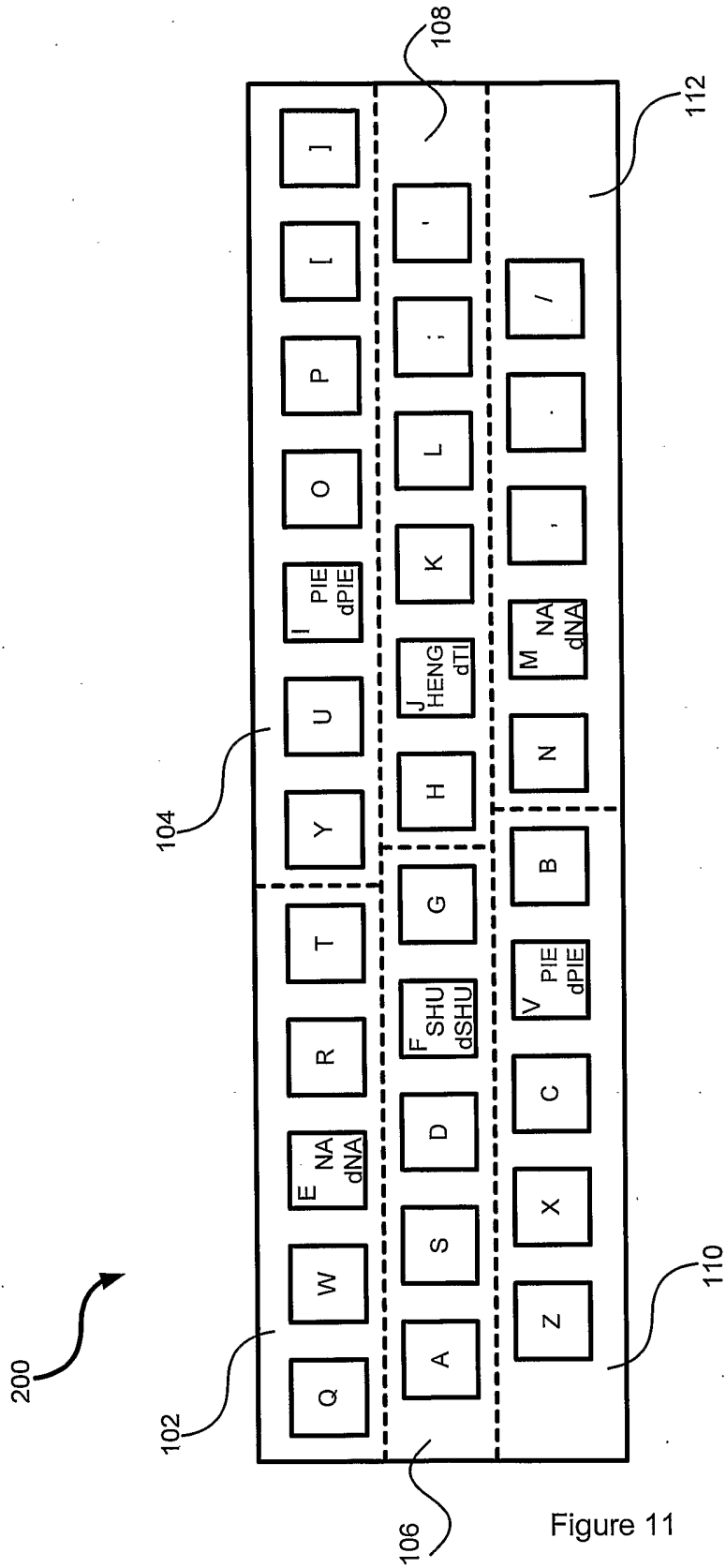


Figure 11

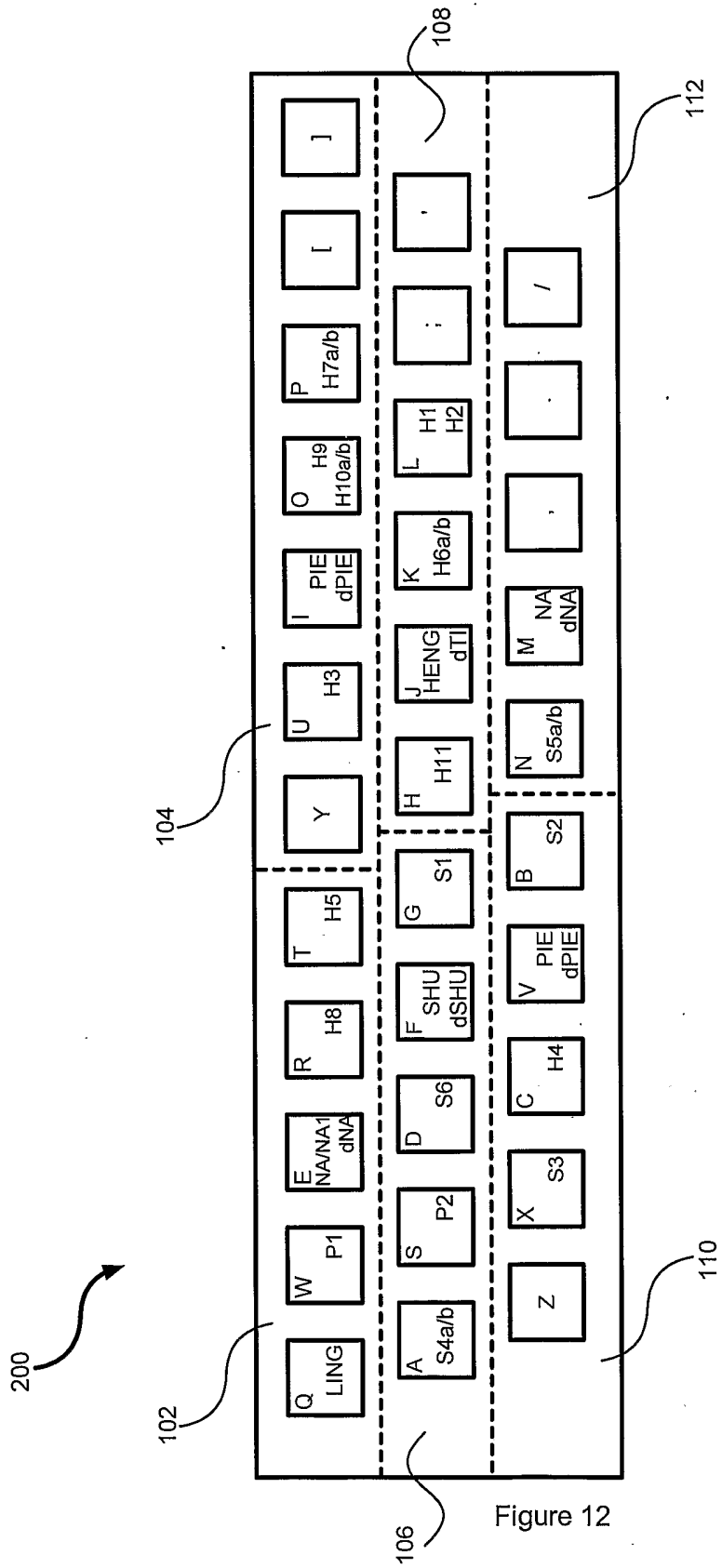


Figure 12

14/19

	Strokes	Stroke Titles	Standard QWERTY Keyboard Position Arrangements	Sample Characters and Relative Stroke Positions Used to Determine the Keyboard Position of Strokes
1	○	LING	Q	
2	—	HENG(TI)	J	米木术六兴文一
3	✓	dTI	J	彳彳
4	ㄚ	H1	L	又久夕之夕ㄥ
5	ㄣ	H2	L	欠了个疋皮声
6	ㄩ	H3	U	β
7	ㄣ	H4	C	及及
8	ㄣ	H5	T	汤场扬
9	ㄣ	H6a	K	力刀刁 冂冂巾冂万卫习勺
10	ㄣ	H6b	K	口ㄣ日尸巴贝五弓己马
11	ㄣ	H7a	P	乃
12	ㄣ	H7b	P	凸
13	乙	H8	R	乞
14	ㄣ	H9	O	飞风气迅执
15	ㄣ	H10a	O	凹
16	ㄣ	H10b	O	几九凡瓦
17	ㄣ	H11	H	讠 鳩顔

Figure 13A

15/19

	Strokes	Stroke Titles	Standard QWERTY Keyboard Position Arrangements	Sample Characters and Relative Stroke Positions Used to Determine the Keyboard Position of Strokes
18	丨	SHU	F	米木术个
19	灬	dSHU	F	𠃉𠃊
20	丿	S1	G	丁于了才 才于 小寸予
21	勹	S2	B	彳豕豸
22	㇇	S3	X	专
23	㇈	S4a	A	与弓马亏乌丐鸟号弗
24	㇉	S4b	A	鼎
25	㇊	S5a	N	匚凵山亡毋升叵
26	㇋	S5b	N	儿乚匕七己尤仓
27	丨	S6	D	钅比瓦切长氏以叫民印
28	丿	PIE	I and V	米木六兴文个义人八爻
29	灬	dPIE	I and V	𠃉𠃊
30	㇌	P1	W	女甬
31	㇍	P2	S	彡厶车东亥乡幺
32	㇎	NA	E and M	米木术文个义人八爻
33	㇏	dNA	E and M	米木六兴文义一彡
34	㇐	N1	M (and E)	弋戈氏戊我

Figure 13B

16/19

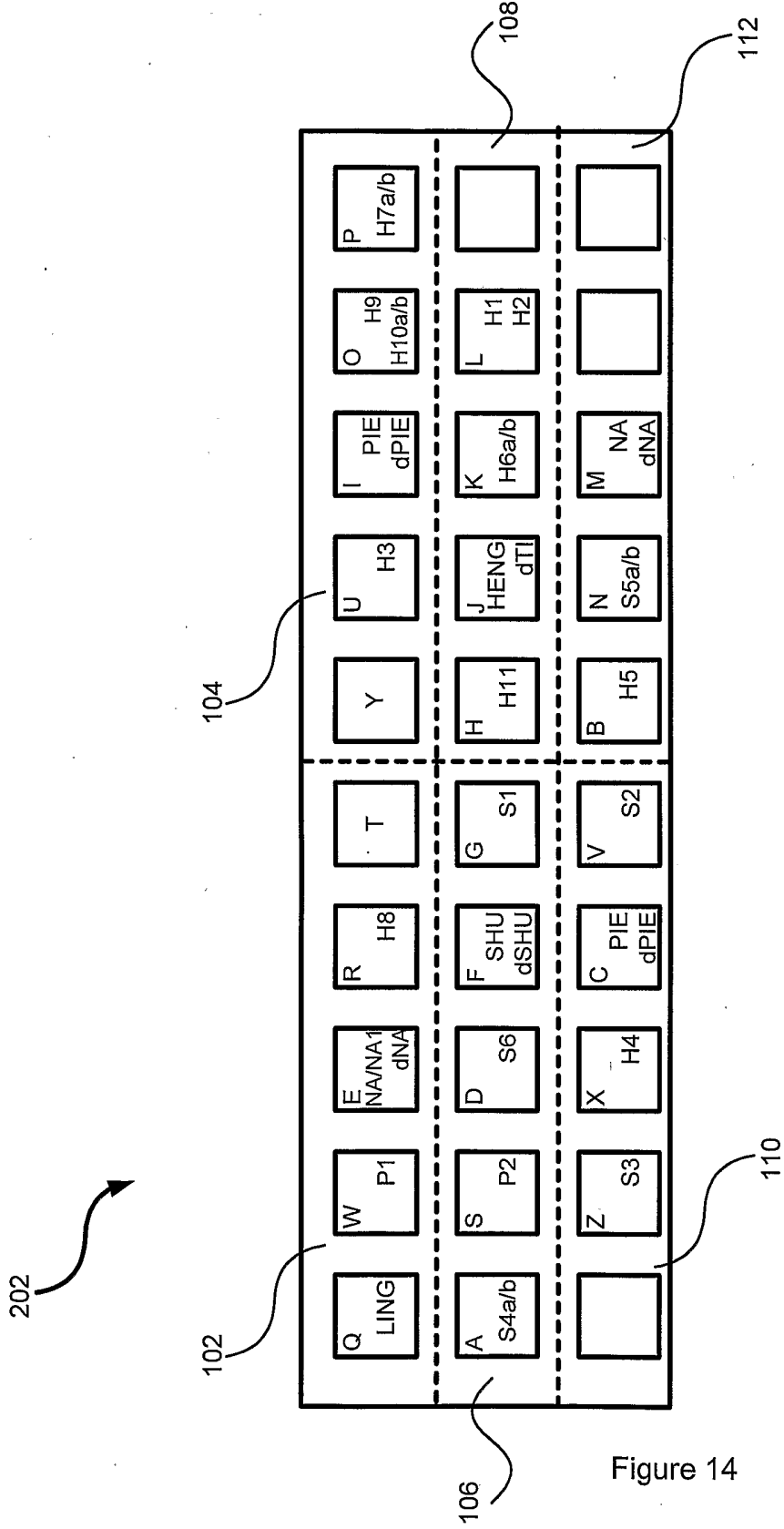
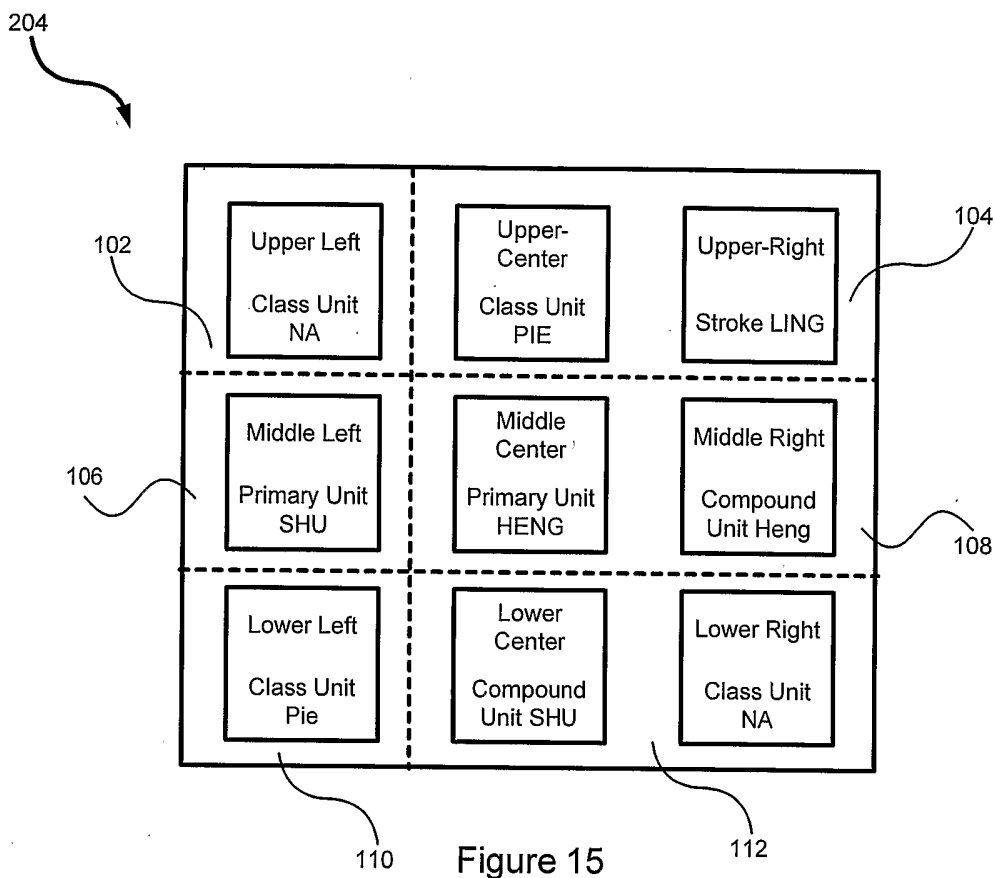


Figure 14

17/19



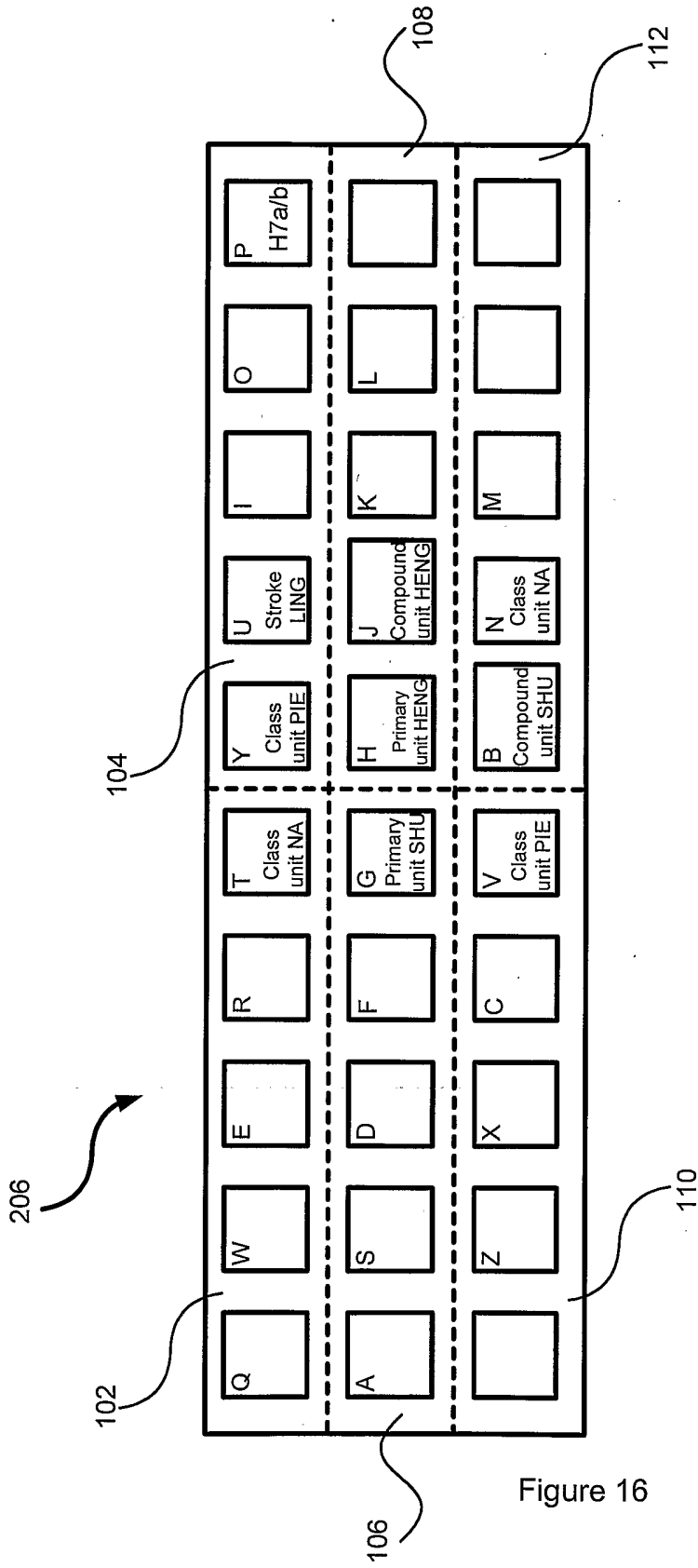


Figure 16

19/19

	English Terms/英文术语	Chinese Terms/中文术语
1	(Chinese) Character	汉字
2	Character-root	字根
3	Radical	部首
4	Stroke	笔划
5	Stroke Table	笔划表
6	Stroke Classification	笔划分类
7	Class	类
8	Class HENG	横类
9	Class SHU	竖类
10	Class PIE	撇类
11	Class NA	捺类
12	Category	部
13	Primary Category of strokes	单笔划部
14	Compound Category of strokes	复合笔划部
15	Group	组
16	Group Left of strokes	左笔划组
17	Group Right of strokes	右笔划组
18	Dian	点
19	Input Method	输入法
20	Relative Stroke Positions	相对笔划位置
21	Character Combination	词(组)

Figure 17