Title: SYSTEM AND METHOD OF INCREASING THE RECOGNITION RATE OF SPEECH-INPUT INSTRUCTIONS IN REMOTE COMMUNICATION TERMINALS

Abstract: A method for enhancing the accuracy of speech-based dialing of remote communication terminals, and terminals incorporating the method, are disclosed. Analog speech input representative of a desired phone number is converted to a digital signal. An automatic speech recognition module identifies the digits and produces an output signal representative of the digits. A determining module applies a test to determine whether one or more digits in the phone number were not recognized by the conversion module. If the phone number includes unrecognized digits, a search module searches an associated memory module for phone numbers having digits that match the recognized digits of the phone number input by the user. Phone numbers from the memory that match may be presented to the user, either visually or audibly. If desired, the remote terminal may establish a connection with the phone number selected from the memory module.
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.
SYSTEM AND METHOD OF INCREASING THE RECOGNITION RATE OF SPEECH-INPUT INSTRUCTIONS IN REMOTE COMMUNICATION TERMINALS

BACKGROUND

The present invention relates to speech-input recognition in communication devices and more particularly to systems and methods for enhancing the accuracy of speech dialing systems in remote communication terminals.

Remote communication terminals such as, for example, mobile telephones are ubiquitous in many modern industrialized countries. Most remote communication terminals utilize a keypad as an input device. However, keypads suffer from certain drawbacks. Foremost, the use of keypads may require a user to direct his or her attention to the communication device, if only for a brief moment. In certain circumstances, such as when driving, this is considered undesirable.

Further, market forces continuously drive manufacturers to produce smaller remote telephone terminal devices, also referred to as handsets. Reducing the size of the terminal device renders keypad errors more likely, thereby reducing the accuracy of the keypad as an input device.

Manufacturers have implemented speech-based input devices adapted to receive a speech input, to recognize the input, and to perform an action based on the input. By way of example, U.S. Patent No. 4,959,850 to Kuniyoshi discloses a radio telephone apparatus that includes speech recognition capabilities for speech-based dialing of the phone. Similarly, U.S. Patents No. 5,042,063 to Sakanishi and No. 4,870,686 to Gerson et al. disclose a telephone apparatus that utilizes speech recognition capabilities to allow speech-based dialing. Speech recognition functions are also disclosed in the following references: U.S. Patents No. 5,917,891 to Will; No. 5,884,257 to Maekawa et al.; No. 5,651,056 to Eting et al.; No. 5,638,425 to Meador; No. 5,509,049 to Peterson; No. 5,495,553 to Jakatdar; and No. 5,303,299 to
Hunt et al.

However, speech recognition is a difficult task, particularly when the speech signal is combined with ambient noise from the surrounding environment, such as automobile noise or street noise. Inadequate enunciation and/or interference from ambient noise may render a user's speech unrecognizable to the device. In speech-based dialing applications, this may result in the telephone device dialing an incorrect number. Alternatively, the telephone device may prompt the user to repeat the unrecognized digit(s), or the entire digit sequence. Depending upon the accuracy of the speech recognition system, the user may be required to repeat numbers a significant percentage of the time, rendering the speech-based dialing feature less convenient for the user.

Accordingly, there is a need in the art for improved speech-based dialing systems and methods.

**SUMMARY**

The present invention addresses these and other problems by providing an apparatus and method for facilitating speech-based dialing of remote communication terminals, including mobile phones. According to the invention, a remote terminal is adapted to use information stored in a memory to enhance the accuracy of the speech-recognition routine. Preferably, the information includes a-priori information about phone numbers previously dialed from the remote terminal, which can be matched with phone numbers input by a speech-based dialing method to enhance the accuracy of the speech recognition system.

In one aspect, the invention provides a system for facilitating speech-based dialing of a communication device. The system comprises a conversion module for receiving speech input representative of an input character sequence and generating a signal representative of each character in the input character sequence, a determining module for determining whether the input character sequence includes
unrecognized characters, a memory module including a plurality of character sequences corresponding to network addresses, and a search module for searching the memory module for a character sequence having characters that correspond to recognized characters in the input character sequence. In use, if the conversion module is unable to convert one or more characters in the input character sequence, then the search module can search the memory module for one or more character sequences in the memory module having characters that match the recognized characters of the input character sequence.

In another aspect, the invention provides a method of facilitating speech-based calling in a communication device. The method comprises the steps of receiving a speech input representative of a desired character sequence, generating a signal representative of each character in the character sequence, determining whether the character sequence includes unrecognized characters, and if so, then searching a memory module for a matching character sequence having characters that correspond to recognized characters in the input character sequence, and generating a signal representative of a matching character sequence.

**BRIEF DESCRIPTION OF THE DRAWINGS**

These and other objects, features and advantages of the present invention will become more apparent upon reading this description, taken in conjunction with the accompanying drawings, wherein:

Fig. 1 is a block diagram of an exemplary GSM communication suitable for implementing the present invention;

Fig. 2 is a flow chart illustrating a method of facilitating speech-based calling in a communication device according to an embodiment of the invention; and

Fig. 3 is a schematic depiction of a remote communication terminal according to an embodiment of the invention.
DETAILED DESCRIPTION

Many digital wireless systems in use today utilize a time slotted access system. User information (e.g., speech) is segmented, compressed, packetized and transmitted in a pre-allocated time slot. Time slots can be allocated to different users, a scheme commonly referred to as Time Division Multiple Access (TDMA). Time Division Multiple Access (TDMA) communication systems, such as the Global System for Mobile communications (GSM) system in Europe, the Digital-Advanced Mobile Phone System (D-AMPS) system in North America, or the Personal Digital Cellular (PDC) system in Japan, allow a single radio frequency channel to be shared between multiple remote terminals, thereby increasing the capacity of the communication system.

The following exemplary embodiments are provided in the context of time division multiple access (TDMA) radiocommunication systems. However, those skilled in the art will appreciate that a TDMA methodology is described solely for purposes of illustration, and that the present invention is readily applicable to all types of access methodologies including frequency division multiple access (FDMA), TDMA, code division multiple access (CDMA) and/or hybrids thereof.

Operation of a cellular communication system in accordance with the GSM standard is described in European Telecommunication Standard Institute (ETSI) documents ETS 300 573, ETS 300 574, and ETS 300 578, which are hereby incorporated by reference. Therefore, the operation of an exemplary GSM system is only briefly described herein. Although the present invention is described in terms of exemplary embodiments in a GSM system, those skilled in the art will appreciate that the present invention could be used in other communication systems.

Referring to Fig. 1, a communication system 10 in which the present invention can be implemented is depicted. The system 10 is a hierarchical network with multiple levels for managing calls. Using a set of uplink and downlink radio...
frequencies, remote communication terminals 12 operating within the system 10 participate in calls using time slots allocated to them on these frequencies. At an upper hierarchical level, a group of Mobile Switching Centers (MSCs) 14 route calls from originators to destinations. In particular, these entities are responsible for setup, control and termination of calls. One of the MSCs 14, commonly referred to as a gateway MSC, handles communication with a Public Switched Telephone Network (PSTN) 18, or other public and private networks.

Each of the MSCs 14 are connected to one or more base station controllers (BSCs) 16. Under the GSM standard, the BSC 16 communicates with a MSC 14 under a standard interface known as the A-interface, which is based on the Mobile Application Part of CCITT Signaling System No. 7.

Each of the BSCs 16 controls one or more base transceiver stations (BTSs) 20. Each BTS 20 includes one or more transceivers (TRXs) (not shown) that use the uplink and downlink radio frequencies (RF channels) to serve a particular geographical area, such as one or more communication cells 21. The BTSs 20 primarily provide the RF links for the transmission and reception of data bursts to and from the remote stations 12 within their respective cells. In an exemplary embodiment, a number of BTSs 20 are incorporated into a radio base station (RBS) 22. The RBS 22 may be, for example, configured according to a family of RBS-2000 products, which products are offered by Telefonaktiebolaget LM Ericsson, the assignee of the present invention. For more details regarding exemplary remote station 12 and RBS 22 implementations, the interested reader is referred to U.S. Patent No. 5,909,469 to Frodigh et al., the disclosure of which is expressly incorporated here by reference.

Fig. 2 presents a schematic depiction of a remote terminal 200 adapted for use in accordance with the present invention. Remote terminal 200 is preferably a mobile phone for use in a digital TDMA cellular communication system, such as, for example, a GSM system, a PDC system, or a D-AMPS system. However, as
noted above, the present invention is applicable to all types of access systems, and can easily be applied in TDMA or CDMA systems, or hybrids thereof. Remote terminals are widely known and readily commercially available. Accordingly, only those aspects of remote terminal 200 that are pertinent to the present invention are described in detail. For additional information relating to remote terminals, the interested reader is referred to U.S. Patent No. 5,745,523 to Dent et al., the disclosure of which is incorporated here by reference.

Referring to Fig. 2, remote terminal 200 includes, in relevant part, a microphone 210 for receiving speech input from a user of the phone. Microphone 210 is connected to conversion module 220. Conversion module 220 may comprise an analog to digital (A/D) converter 224 for converting analog speech input to a digital signal. Conversion module 220 may also include an automatic speech recognition (ASR) module 228 for recognizing the speech of the user. Remote terminal 200 further includes a determining module 230 for determining whether a character spoken by the user was recognized by ASR module 228 with a desired degree of accuracy. Remote terminal 200 further includes a memory module 250 for storing character sequences that represent valid phone numbers, and a search module 240 for searching memory module 250. Remote terminal 200 also includes a connection module 260 for establishing a communication connection with a communication network such as, for example, a GSM network as depicted in Fig. 1. Remote terminal 200 further includes a suitable display 270 (e.g., an LED or LCD display) for displaying information to a user. One terminal with a suitable speech recognition module is the T28 commercially available from Ericsson.

It will be appreciated that some or all of modules 220-260 may be embodied in a suitable application specific integrated circuit (ASIC) or a programmed digital signal processor (DSP), or by a chip set comprising a plurality of ASICs. Electrical connections are formed between the respective modules 220-260 and other components of the remote terminal. For example, determining module 230 and
search module 240 are electrically connected to display 270, to speaker 280, and to connection module 260.

Additionally, in a preferred embodiment, an electrical connection between memory module 250 and connection module 260 allows memory module 250 to store telephone numbers associated with connections established by remote terminal 200. For example, each time a user enters a phone number in remote terminal 200, the number may be stored in memory module 250. In this manner, memory module 250 maintains a list of previously-dialed telephone numbers that can be used as a-priori information to enhance the accuracy of speech-based dialing, as described below.

Fig. 3 illustrates a method for speech-based dialing according to an embodiment of the invention. In brief overview, referring to Fig. 3, the method includes receiving a spoken character from a user, converting the character to a digital signal, and determining whether the character sequence is complete. If the character sequence is not complete, the system iteratively receives additional characters and converts the characters to a digital signal. After a complete character sequence has been received, the system determines whether the character sequence includes one or more unrecognized characters. If the character sequence does not include unrecognized characters, then the character sequence may be transmitted to a module (e.g., a connection module) that enables the phone to dial the number corresponding to the recognized character sequence. If the character sequence includes one or more unrecognized characters, then a search module is invoked. The search module compares the recognized digits in the character sequence with corresponding digits in character sequences in an associated memory to determine whether a character sequence in memory is a likely match with the character sequence input by the user. When a likely match is detected, the character sequence may be transmitted to a module that enables the phone to dial the number corresponding to the recognized character sequence. Alternatively, the character
sequence may be displayed or audibly presented to the user of the phone, who can indicate whether the character sequence does, in fact, match the desired character sequence. This process will be explained in greater detail below.

In an exemplary embodiment, the process set forth in Fig. 3 may be implemented in a remote communication terminal, e.g., a mobile phone, having a speech-based dialing feature. Referring to Fig. 3, at step 310 the speech-based dialing feature is activated and the remote terminal receives speech input representative of a first character in a character sequence. In the United States, the character preferably represents one digit of the well-known ten-digit dialing format (e.g., xxx-xxx-xxxx). However, it will be appreciated that the character sequence could be in a format adapted for a dialing system of a different geographic region, or, in a data application, could represent a network address in a data network (e.g., a URL or an IP address). Alternatively, the character sequence may represent commands addressed to the remote terminal, or a memory location that includes a number for speed dialing.

At step 320, the received character is converted to a digital signal representative of the character spoken by the user. Conversion may be accomplished using an analog-to-digital (A/D) converter in combination with a suitable ASR module. Many ASR modules implement statistical procedures for reporting reliability metrics of the determination made for a particular character. Desired reliability rates may be programmed into the ASR module's logic, or may be selectable by the user and input to the system as a parameter. ASR modules are known in the art, and particular details of the ASR module are not critical to the invention.

At step 330, a test is performed to determine whether the character sequence input is complete. For example, in the United States telephone system, which uses a ten character format, the character sequence may be considered complete at the entry of the tenth character. In an alternate embodiment, the determination step may use a
time-out procedure, such that the character sequence is assumed to be complete if a predetermined time elapses after the entry of a particular character. In another alternate embodiment, a user may actively indicate that the character sequence is complete, either by pressing a designated key or by speaking a designated code. One of ordinary skill in the art will recognize numerous other ways to detect the end of an input character sequence. If the character sequence is not complete, then steps 310 through 330 may be repeated until the character sequence is complete, or the user indicates a desire to cancel the speech input process.

After it is determined that the character sequence is complete, at step 340, a test is conducted to determine whether the character sequence includes one or more unrecognized characters. As used herein, the term “unrecognized character” shall refer to a character in the character sequence that is not validated by the ASR module. In one embodiment, the system may test to determine whether a reliability metric associated with one or more characters in the character sequence is less than a predetermined threshold (e.g., 95%, or 90%), and, if so, then the character sequence may be characterized as having unrecognized characters. Additional tests may also be applied. For example, if the reliability metric associated with two characters is less than a predetermined threshold, then the character sequence may be characterized as having unrecognized characters.

If the character sequence does not include unrecognized characters, then at step 380, the character sequence is dialed and remote terminal 200 attempts to establish a connection with the network.

If the character sequence includes unrecognized characters, then at step 350, a memory module associated with the remote terminal is searched to determine whether a character sequence in the memory module matches the recognized characters in the character sequence input by the user. If at step 360, a match is found, then the character sequence is retrieved from memory and optionally may be presented to the user, at step 370. In one embodiment, the character sequence is
visually presented to the user, such as by display on a LCD or other suitable display. In another embodiment, a speech synthesizer presents the character sequence to the user audibly. Upon receiving an indication of approval from the user, the character sequence is dialed at step 380.

It will be recognized that some or all of steps 310 through 380 may be performed by a suitable ASIC, DSP, or chip set, or by logic instructions operating on a general purpose processor.

Although the invention has been described in detail with reference to a few exemplary embodiments, those skilled in the art will appreciate that various modifications can be made without departing from the invention. Accordingly, the invention is defined only by the following claims which are intended to embrace all equivalents thereof.
What is claimed is:

1. A system for facilitating speech-dialing of a communication device, comprising:
   a conversion module for receiving speech input representative of an input character sequence and generating a signal representative of each character in the input character sequence;
   a determining module for determining whether the input character sequence includes unrecognized characters;
   a memory module including a plurality of character sequences corresponding to network addresses; and
   a search module for searching the memory module for a character sequence having characters that correspond to recognized characters in the input character sequence;
   such that, if the conversion module is unable to convert one or more characters in the input character sequence, then the search module can search the memory module for one or more character sequences in the memory module having characters that match the recognized characters of the input character sequence.

2. A system according to claim 1, wherein the conversion module comprises:
   an A/D converter for digitizing the received speech input signal.

3. A system according to claim 1, wherein the conversion module comprises:
   a speech recognition module for analyzing the digital signal and generating a signal indicative of a character sequence represented by the digital signal.

4. A system according to claim 1, wherein:
   the conversion module generates a signal representative of a confidence level
associated with the accuracy of the conversion; and
the determining module generates a signal indicative of whether the
confidence level is greater than a predetermined threshold.

5. A system according to claim 1, wherein:
the conversion module and the determining module are embodied within a
digital signal processor.

6. A system according to claim 1, further comprising:
an output module for generating a signal representative of a character
sequence in the memory.

7. A system according to claim 6, further comprising:
a display module for displaying the character sequence represented by the
signal generated by the output module.

8. A system according to claim 6, further comprising:
a module for audibly announcing the character sequence represented by the
signal generated by the output module.

9. A system according to claim 1, further comprising:
a connection module for establishing a connection with the character
sequence represented by the signal generated by the output module.

10. A method of facilitating speech-based calling in a communication device,
comprising the steps of:
receiving a speech input representative of a desired character sequence;
generating a signal representative of each character in the character
sequence;
determining whether the character sequence includes unrecognized characters, and if so, then searching a memory module for a matching character sequence having characters that correspond to recognized characters in the input character sequence; and
generating a signal representative of a matching character sequence.

11. A method according to claim 10, wherein the step of generating a signal representative of each character in the character sequence includes digitizing the received speech input signal.

12. A method according to claim 11, wherein the step of generating a signal representative of each character in the character sequence includes analyzing the digital signal and generating a signal indicative of a character sequence represented by the digital signal.

13. A method according to claim 10, wherein the step of generating a signal representative of each character in the character sequence includes generating a first signal representative of a confidence level associated with the accuracy of the conversion.

14. A method according to claim 13, wherein the step of determining whether the character sequence includes unrecognized characters includes comparing the confidence level to a predetermined threshold and generating a second signal indicative of whether the confidence level is greater than a predetermined threshold.

15. A method according to claim 10, further comprising displaying the character sequence represented by the signal generated by the output module.
16. A method according to claim 10, further comprising audibly announcing the character sequence represented by the signal generated by the output module.
FIG. 1
PRIOR ART
FIG. 2