Exhibit 1

US008165867B1

(12) **United States Patent**
Fish

(10) **Patent No.: US 8,165,867 B1**
(45) **Date of Patent: Apr. 24, 2012**

(54) **METHODS FOR TRANSLATING A DEVICE COMMAND**

(76) Inventor: **Robert D. Fish**, Tustin, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 649 days.

(21) Appl. No.: **10/466,202**

(22) PCT Filed: **Sep. 15, 2000**

(86) PCT No.: **PCT/US00/25613**
§ 371 (c)(1),
(2), (4) Date: **Sep. 6, 2006**

(87) PCT Pub. No.: **WO02/23389**
PCT Pub. Date: **Mar. 21, 2002**

(51) **Int. Cl.**
G06F 17/28 (2006.01)

(52) **U.S. Cl.** **704/3**; 704/251; 704/260; 704/271

(58) **Field of Classification Search** 704/3, 251, 704/260, 271
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,818,733	A *	10/1998	Hyuga	702/188
6,080,972	A *	6/2000	May	219/494
6,397,267	B1 *	5/2002	Chong, Jr.	710/1
6,633,235	B1 *	10/2003	Hsu et al.	340/825.69
6,654,378	B1 *	11/2003	Mahany et al.	370/401
2002/0069063	A1 *	6/2002	Buchner et al.	704/270

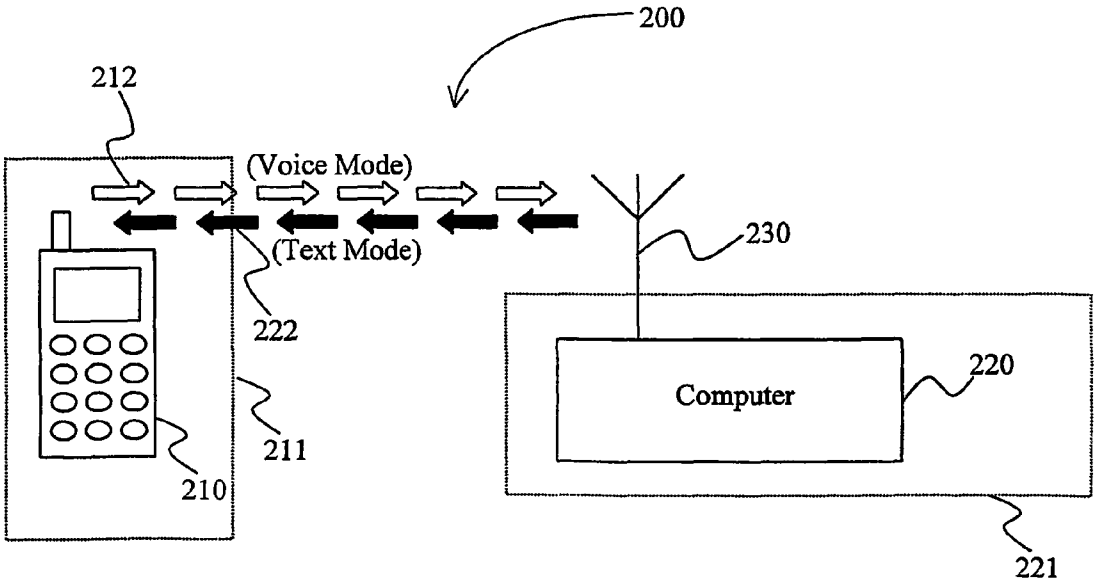
* cited by examiner

Primary Examiner — Leonard Saint Cyr
(74) *Attorney, Agent, or Firm* — Fish & Associates, PC

(57) **ABSTRACT**

An item of information is transmitted to a distal computer, translated to a different sense modality and/or, language and in substantially real time, and the translation is transmitted back to the location from which the item was sent. The device sending the item is preferably a wireless device, and more preferably a cellular or other telephone. The device receiving the translation is also preferably a wireless device, and more preferably a cellular or other telephone, and may advantageously be the same device as the sending device. The item of information preferably comprises a sentence of human speech having at least ten words, and the translation is a written expression of the sentence. All of the steps of transmitting the item of information, executing the program code, and transmitting the translated information preferably occurs in less than 60 seconds of elapsed time.

17 Claims, 1 Drawing Sheet



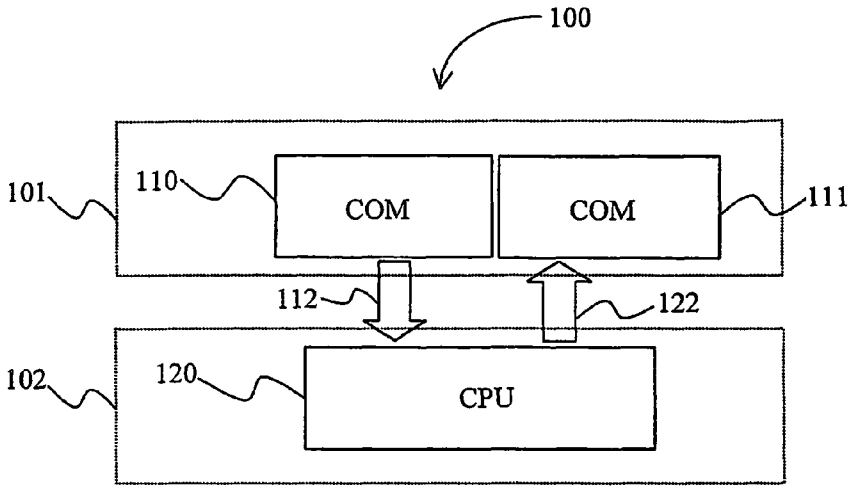


Figure 1

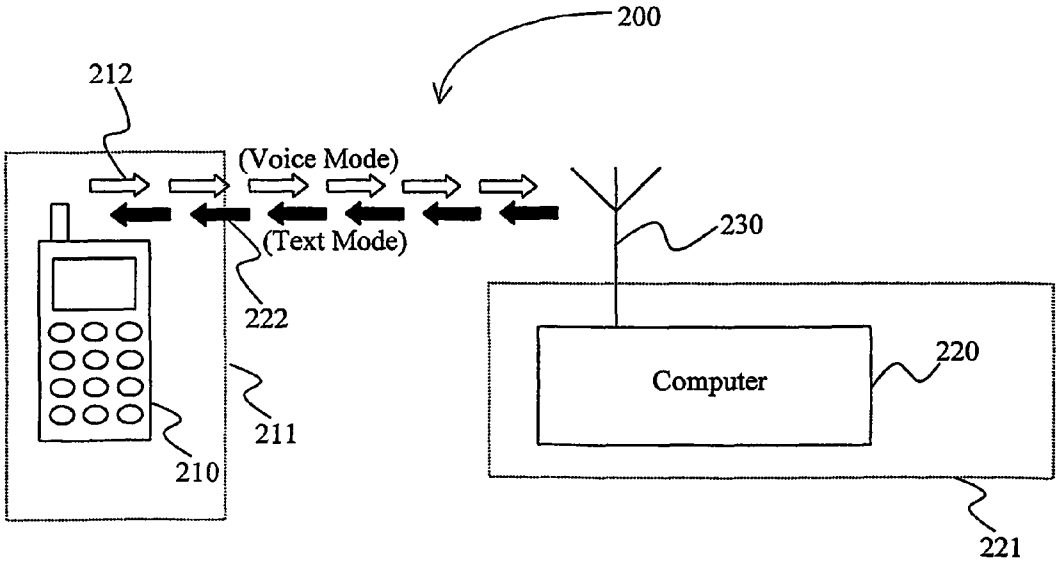


Figure 2

METHODS FOR TRANSLATING A DEVICE COMMAND

FIELD OF THE INVENTION

The field of the invention is remote computing.

BACKGROUND OF THE INVENTION

As processing speeds continue to improve and data storage becomes ever less expensive, many sophisticated applications that were previously only available on mainframe or desktop computers have been ported to laptop computers and other portable electronic equipment. Many applications have even been ported to hand held electronic devices as well, including hand held computers, digital telephones, personal digital assistants (PDAs), and so forth. For example, personal databases with limited search capabilities are now included in cellular phones, and word processing can now be performed in PDAs.

There are, however, several applications that are presently difficult or impossible to realize on hand-held electronic devices, and are only poorly realized even on larger systems such as desktop computers. Due to the large volumes of data involved, and the need to process at very high speeds, a particularly difficult application is voice recognition. Some attempts have been made in that direction, but all of them suffer from one or more disadvantages.

At the low end, limited word or phrase recognition capabilities are sometimes provided in cell phones. Such systems can usually recognize only a few words (e.g., the numerals 0-9, and specialized key words such as a person's name, or the commands "dial" or "open file patentapp.doc"). Such systems are particularly advantageous where only rudimentary recognition capabilities are needed, or where only very limited data storage capability or computing power is available. However, an obvious shortcoming of the word or phrase recognition systems is that the usability is limited to a small, preprogrammed vocabulary, and at most a few custom words. Moreover, word or phrase recognition systems often fail to recognize personal speech pattern or accents.

At the higher end, speech recognition programs are currently available for operation on laptop computers. As used herein both "speech recognition" and "word or phrase recognition" are considered to be categories of voice recognition. "Speech recognition", however, is limited to systems having a vocabulary of at least 200 words, and where individual words are interpreted in the context of surrounding words. For example, speech recognition would correctly interpret phrases such as "I have been to the beach" whereas a word or phrase recognition system may substitute "bean" for "been".

As with other computer software application, most of the development effort is being directed towards porting the more sophisticated speech recognition to smaller and smaller devices. It may well be that within a decade the goal of true speech recognition will be available on even hand-held electronic devices.

What is not presently appreciated, however, is that porting of sophisticated software to portable electronic devices may not be desirable. Cell phones, for example, need only relatively rudimentary electronics to support the required communications, and placing sophisticated storage and processing in cell phones may be a waste of money. Moreover, no matter how sophisticated the software and hardware becomes in hand held and other portable devices, there will always be

capabilities for different accents and languages, and perhaps even language translation capabilities. Still further, it is impractical to install voice recognition in all the myriad types of devices that may advantageously utilize voice recognition. For example, voice recognition may be useful in VCR and CD players, kitchen and other household appliances such as toasters and washing machines, automobiles and so forth.

Thus, while it has been known to translate information in a first sense modality and language into a second sense modality and language on a single local computer, it has not been appreciated to perform the translation in a "remote computing" manner, thereby concentrating the computing power in a cost effective manner. Consequently, there is a need to provide voice recognition capabilities, and especially speech recognition capabilities, to myriad electronic devices without actually installing all of the required hardware and software in all such devices.

SUMMARY OF THE INVENTION

The present invention provides systems and methods in which an item of information is transmitted to a distal computer, translated to a different sense modality and/or language, and in substantially real time, and the translation is transmitted back to the location from which the item was sent.

The device sending the item is preferably a wireless device, and more preferably a cellular or other telephone. The device receiving the translation is also preferably a wireless device, and more preferably a cellular or other telephone, and may advantageously be the same device as the sending device. The item of information preferably comprises a sentence of human speech having at least ten words, and the translation is a written expression of the sentence. All of the steps of transmitting the item of information, executing the program code, and transmitting the translated information preferably occurs in less than 60 seconds of elapsed time, and more preferably less than 30 seconds.

Various objects, features, aspects and advantages of the present invention will become more apparent from the following detailed description of preferred embodiments of the invention, along with the accompanying drawing.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. **1** is an exemplary schematic of a method of changing the sense modality of an information according to the inventive subject matter.

FIG. **2** is an exemplary embodiment of a method of changing the sense modality of an information according to the inventive subject matter.

DETAILED DESCRIPTION

As used herein, the term "sense modality" refers to the manner in which information is perceived by a human being. There are five sense modalities comprising sight, sound, taste, smell, and touch. Obviously, different aspects of information may be expressed in multiple sense modalities at the same time. A conversation between two people, for example, may be perceived as both sound (spoken language) and sight (hand gestures). Similarly, music can be perceived as both sound (auditorily perceived vibration) and touch (tactually perceived vibration).

Information in each of the five sense modalities can be expressed in numerous languages, with the term "language"

languages as well as various graphics languages. Exemplary text languages include the various character sets of human languages (Roman, Cyrillic, Chinese, etc), as well as computer languages (ASCII, HTTP, XML, Basic, Cobol, Pascal, C++, etc). Graphics "languages" include moving images, still images, painting, and so forth.

Even within a given language there are different styles, which are also referred to herein from time to time as styles. Character fonts (Arial, Courier, Gothic, Lucida, Times New Roman, various forms of handwriting, etc) comprise one type of style, and various sizings and spacings of characters comprise other styles. With respect to graphics there are styles here as well. Moving images, for example, can be styled as VCR or Beta video, or DVD. Similarly, still images can be styled as hard copy photographs, TIF, GIF, and other computer files.

The sense modality of sound is also deemed herein to include several languages, including the various spoken and written human languages, various languages of music (including, for example, classical music, rock music, punk music, and jazz), animal sounds, industrial sounds, transportation sounds, and electronic sounds such as beeps. Still other languages are contemplated as well, each of which may have several different styles. With the language of classical music, for example, some of the possible styles include baroque, modem, and so forth.

Technically, the sense modality of taste only includes four possible sensations, sweet, sour, salty and bitter. In our lexicon these would comprise the different languages of taste, with variations within each sensation comprising the different styles.

In our lexicon, the sense modality of smell includes the "languages" of florals, musks, foods, inorganics, etc.

In our lexicon, the sense modality of touch includes the "languages" of vibration, pressure, temperature, movement, texture, etc.

As can now be appreciated, the terms "sense modality", "language", and "style" are each used herein in a very specific manner. Sense modalities are distinguished one from another by the sense organ(s) primarily used to detect the information, while languages are different means of expression within a given sense modality. With a given sense modality and language, styles refer to variations in expressing information that can be achieved without changing the language.

All of these are distinguishable from the "medium", which is employed herein to mean the physical device upon which an item of information resides. A photographic image, for example, may reside on a piece of photographic paper, in which case the medium is the paper. The same image may also reside on computer disk, in which the medium is the disk. The image can also be transmitted via modem, in which case the medium may be a copper wire.

This is an important distinction because a change in medium does not necessarily mean a change in sense modality or style. For example, when a person talks on a portable telephone, the relevant item of information may be a spoken sentence. The sense modality would be sound, and the language may be that of English. The style may be very fast, slurred speech. The telephone translates the sounds into an analog or digital language for transmission through the medium of air, with the particular style depending upon the specific protocols of the service provider. Throughout the entire process, however, the sense modality is still considered to be sound because that is how a human being would perceive the information once it was converted back into an

changed between digital and analog, the information is still considered to maintain the same language and style.

There are many circumstances in which it is known to translate information between sense modalities, and between languages of the same or different sense modalities. For example, the jazz can be translated between written notes (sight modality, and possibly Western music transcription as the language) and notes played on an instrument (sound modality, with jazz as the language). Similarly, spoken English (sound modality, English language) can be translated between spoken German (sound modality, German language). Humans are quite adept at performing such translations internally, and as discussed above, computers are beginning to achieve a useful translation capability as well.

In all known instances of which the present inventor has knowledge, however, the information is never wirelessly transmitted to a distant computer for translation, translated at the distant computer (at least 20 kilometers away), wirelessly returned to the location from which it was sent ("locally", "local", and "location" all being defined as within a radius of 100 meters), and then expressed locally to the source, all in substantially real time (less than three minutes from initial transmission of the information to expression of the translated information). Examples follow:

- In laboratories that develop voice recognition software, it is presumably known to utilize a central computer for development work, and to access that computer using workstations wired into the central computer. That situation does not, however, involve wireless transmission, and the translating computer is not distal.
- A user loads voice recognition software on a desktop or laptop computer, telephones the computer to record a message, and then accesses that information from a distant computer. In that situation the operation does not occur in substantially real time. The user most likely records several minutes of speech using his telephone, and then downloads a text file translated from the speech using a laptop or other computer.
- One person transmits an e-mail to a recipient, and the recipient causes a computer to "read" the e-mail to him over the telephone. In that situation the total duration between transmitting of the e-mail and hearing it spoken is most likely not less than 60 seconds, and the message is most likely not heard locally to the place from which the e-mail was originally sent.
- A user employs a distal central computer for computational purposes. The user enters the equation x=156×2, asks the computer for the answer, and the computer immediately transmits back the answer. That situation falls outside the present invention because the distal computer evaluated the expression rather than translate what was sent to it. If the computer had returned the spoken words "x equals one hundred fifty six times two", then the computer would have returned a translation.
- A user has a cell phone that is connected to a music web site on the Internet. The user speaks the words "Beethoven's Fifth Symphony", and the web site transmits a portion of the symphony over the phone. This situation also falls outside the present invention because the distal computer evaluated the words rather than translated them. If the computer had returned the text "Beethoven's Fifth Symphony", then the computer would have returned a translation.
- A user employs his cell phone to secure a dictionary definition. He speaks a particular word, the cell phone trans-

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.