

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
7 April 2005 (07.04.2005)

PCT

(10) International Publication Number
WO 2005/031995 A1

(51) International Patent Classification⁷: H04B 1/38,
H04Q 7/20, H04M 1/00

(21) International Application Number:
PCT/US2004/030553

(22) International Filing Date:
17 September 2004 (17.09.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03124963.9 23 September 2003 (23.09.2003) CN

(71) Applicant (for all designated States except US): MO-
TOROLA, INC. [US/US]; 1303 East Algonquin Road,
Schaumburg, IL 60196 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ZHANG, Yaxin

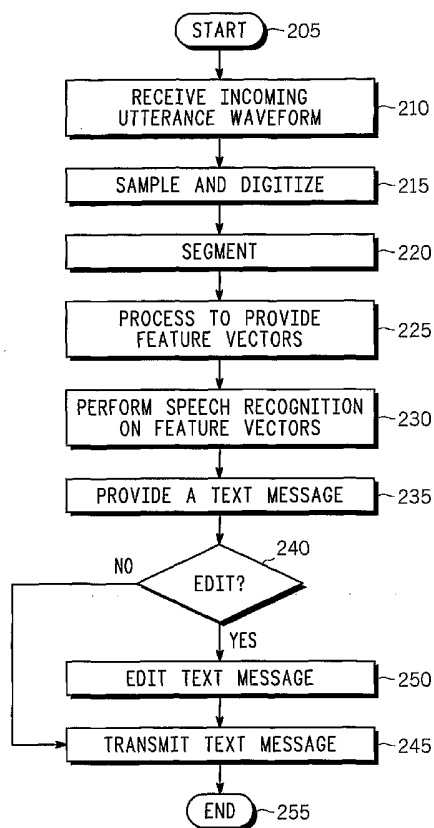
[AU/AU]; 28/2-8 Bridge Street, Hurstville, New South
Wales 2220 (AU). HE, Xin [CN/CN]; Room 6-203, 710
Yishan RD., Shanghai 200233 (CN). REN, Xiao-Lin
[CN/CN]; 402 No. 63 Building, 1880 Longyang Road,
Shanghai 201204 (CN). SUN, Fang [CN/CN]; Panyu Rd.
800/2/604, Shanghai 200030 (CN).

(74) Agents: NICHOLS, Daniel, K. et al.; 600 North U.S.
Highway 45, Libertyville, IL 60048 (US).

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, 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, NA, 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.

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR PROVIDING A TEXT MESSAGE



(57) Abstract: A method and apparatus for providing a text message includes receiving an utterance (Step 210) at an input of an electronic device (100). Speech recognition is then performed on the utterance (Step 230) guided by user-defined message templates stored in a memory (155) associated with the electronic device (100). Speech recognition is defined by matching the utterance with one of the templates to create a matching template. A text message is then provided from the matching template (Step 235).

WO 2005/031995 A1



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

— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

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.

Published:

— with international search report

METHOD AND APPARATUS FOR PROVIDING A TEXT MESSAGE

FIELD OF THE INVENTION

5 The invention relates to a method and apparatus for providing a text message using voice. The invention is particularly useful for, but not necessarily limited to, providing a text message using voice inputs processed on a portable electronic device having limited memory and computational capacity.

BACKGROUND OF THE INVENTION

10 Short text messaging, often using the Short Messaging Service (SMS) format, is a very popular application in wireless communications. Billions of short text messages are sent each month, usually from one mobile phone to another. Such text messages are popular for a number of reasons. The messages are generally a fraction of the cost of a one-minute mobile telephone call and they do not require an engaged tone to send or to receive. Therefore the messages can be created and sent at a time that is convenient to the sender, and received and read at a time that is convenient to the recipient.

20 Text messages are generally created by typing characters into the keypad of a mobile telephone. However using such small, non-quiry keypads to compose a message can be awkward and generally requires more time than would be needed using a full-size query keyboard. But of course it is impractical to have a full size keyboard attached to a mobile phone. Thus there is a need for a more effective method of composing short text messages.

25 Further, although various types of speech recognition systems are well known, most are not suitable for use in portable electronic devices such as mobile phones. That is because prior art speech recognition systems generally require more processing power and memory than are available in portable electronic devices.

30 Prior art closed vocabulary speech recognition systems and methods employ a pre-defined, fixed vocabulary list. In use, the fixed vocabulary list

may be large but may not be exhaustive and therefore, for instance, a person's family name and the names of many locations would not be included. In contrast, open vocabulary speech recognition systems and methods have a variable vocabulary list to which new words and phrases may be added by a user or otherwise. However, current open vocabulary speech recognition systems and methods require relatively high computational overheads that may not be acceptable for portable electronic devices such as Personal Digital Assistants, radio-telephones and other portable devices.

In this specification, including the claims, the terms 'comprises', 'comprising' or similar terms are intended to mean a non-exclusive inclusion, such that a method or apparatus that comprises a list of elements does not include those elements solely, but may well include other elements not listed.

SUMMARY OF THE INVENTION

According to one aspect of the invention there is provided a method of providing a text message. The method includes the steps of receiving an utterance at an input of an electronic device. Speech recognition is then performed on the utterance guided by user-defined message templates stored in a memory associated with the electronic device, wherein speech recognition is defined by matching the utterance with one of the templates to create a matching template. A text message is then provided from the matching template.

At least one of the message templates may include a fixed language component.

At least one of the message templates may include a variable language component.

At least one of the message templates may include both a fixed and a variable language component.

The text message may be an SMS message.

The above method may also include the step of editing the user-defined message template by receiving typed characters from a keypad of the electronic device.

A component of the text message may be a transcription of the utterance.

The entirety of the text message may be a transcription of the utterance.

According to another aspect of the invention there is provided an
5 electronic device for providing a text message. The device includes a
microphone operative to receive an utterance; a non-volatile memory for
storing message templates; and a processor operative to perform speech
recognition of the utterance guided by the message templates, wherein
processor is operative to match the utterance with one of the templates to create
10 a matching template, and to provide a text message from the matching
template.

With respect to the electronic device, the message templates may also
include fixed or variable language components or both fixed and variable
language components.

15 With respect to the electronic device, the text message may be an SMS
message.

The electronic device may include a keypad operative for editing the
message template.

The electronic device may be operative to match the utterance with a
20 plurality of the templates and to calculate a likelihood score for each of the
templates.

BRIEF DESCRIPTION OF THE DRAWINGS

In order that the invention may be readily understood and put into
25 practical effect, reference will now be made to preferred embodiments as
illustrated with reference to the accompanying drawings in which:

Fig. 1 is a schematic block diagram of a radio telephone in
accordance with the present invention;

30 Fig. 2 is a flow diagram illustrating a method for providing,
editing and transmitting a text message in accordance with the present
invention;

Fig. 3 is a flow diagram that illustrates a method for providing a list of candidate message templates to a user in accordance with the present invention; and

5 Fig. 4 is a flow diagram illustrating a method for enabling a user to edit existing message templates and save new templates in a static programmable memory in accordance with the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS OF THE INVENTION

10 With reference to FIG. 1, there is illustrated a radio telephone 100 comprising a radio frequency communications unit 105 coupled to be in communication with a processor 110. Input/Output (I/O) interfaces in the form of a display 115, a keypad 120, a microphone 190, and a speaker 195 are also coupled to be in communication with the processor 110.

15 The processor 110 includes an encoder/decoder 125 with an associated Read Only Memory (ROM) 130 storing data for encoding and decoding voice or other signals that may be transmitted or received by the radio telephone 100. The processor 110 also includes a micro-processor 135 coupled, by a common data and address bus 140, to the encoder/decoder 125 and an associated
20 character Read Only Memory (ROM) 145, a Random Access Memory (RAM) 150, static programmable memory 155 and a removable SIM module 160. The static programmable memory 155 and SIM module 160 each can store, amongst other things, selected incoming text messages, a telephone book database, and, as described in more detail below, templates of outgoing text
25 messages. The microprocessor 135 has ports for coupling to the keypad 120, the display 115 and an alert module 165 that typically contains a speaker, vibrator motor and associated drivers. The character Read Only Memory 145 stores code for decoding or encoding text messages that may be received by the
communication unit 105, input at the keypad 120.

30 The radio frequency communications unit 105 is a combined receiver and transmitter having a common antenna 170. The communications unit 105 has a transceiver 175 coupled to antenna 170 via a radio frequency amplifier

180. The transceiver 175 is also coupled to a combined modulator/demodulator 185 that couples the communications unit 2 to the processor 110.

Referring to Fig. 2 there is a flow diagram illustrating one embodiment of the present invention including a method 200 for providing, editing and transmitting a text message using the radio telephone 100. The method 200 is invoked at a start step 205. At step 210 an utterance is received at an input, such as the microphone 190, of the telephone 100. The processor 110 then performs sampling and digitizing of the utterance waveform at step 215, then segmenting at a step 220 before processing to provide feature vectors representing the waveform at a step 225. It should be noted that steps 215, 220, and 225 are well known in the art and therefore do not require a detailed explanation.

Next, at step 230, speech recognition is performed on the feature vectors resulting from step 225. The speech recognition is guided by user-defined message templates stored in the static programmable memory 155 of the device 100. The message templates are described in more detail later in this specification. The method 200 then provides a text message to a user at step 235. The message may be provided to the user using one of the I/O interfaces such as the display 115 or the speaker 195 of the device 100.

After the message is provided to the user, the user is then able to decide whether to edit the message at step 240. If the user decides not to edit the message, the message is transmitted at step 245 in a message format such as SMS. However if the user decides at step 240 to edit the message, the message is edited at step 250 before being transmitted at step 245. In various embodiments of the present invention, the user may edit the message in several different ways including speaking edits into the speaker 195 or typing edits into the keypad 120. The method 200 then ends at step 255.

In an alternative embodiment of the present invention, after the speech recognition step 230 described above the provide a text message step 235 may include providing a user of the telephone 100 with a list of candidate message

templates from which the user may select the template that is most appropriate for the intended text message.

Fig. 3 is a flow diagram that illustrates a method 300 for providing such a list of candidate templates to a user. The method 300 is invoked at start step 305 when a user inputs a command into the keypad 120 or into the microphone 190. The method 300 first includes the processor 110 selecting at step 310 a message template from a list of available message templates. At step 315 the selected template is then compared with the feature vectors provided in step 225 of method 200. The processor 110 then calculates a likelihood score at step 320 that estimates the matching quality between aspects of the selected template and the feature vectors of the input utterance.

The processor 110 then determines at step 325 whether the likelihood score is above a set threshold. The threshold may be automatically calculated by processor 110, or it may be pre-set by a user of the telephone 100. If the likelihood score of the selected template is below the set threshold, the template is rejected at step 330. However if the likelihood score of the selected template is above the set threshold, then at step 335 the template is considered to be a reasonable match with the input utterance and the template is added to a list of candidate templates. Regardless whether the selected template is rejected or added to the list of candidate templates, the method 300 then proceeds to step 340 where the processor 110 determines whether all available templates have been evaluated. If all available templates have not been evaluated, at step 345 the method 300 selects the next message template and returns to step 315 where the next template is compared with the feature vectors of the input utterance. If all templates have been evaluated at step 340, the method 300 continues to step 350 and provides a list of all of the candidate templates to the user. The candidate templates may be provided to the user using one of the I/O interfaces such as the display 115 or the speaker 195 of the device 100. The method 300 then ends at step 355.

According to one embodiment of present invention, users of the telephone 100 are not limited to the use of templates supplied by a manufacturer of the device 100. Rather, users of the device 100 are able to

edit existing templates stored in the static programmable memory 155 to create their own personalized message templates.

Referring to Fig. 4, there is illustrated a method 400 for enabling a user to edit existing templates and save new templates in static programmable memory 155. The method 400 is invoked at start step 405 when a user inputs a command into the keypad 120 or into the microphone 190. At step 410 a list of existing templates is provided to the user of the device 100 through an I/O interface such as the display 115 or the speaker 195. The user then selects a desired message template at step 415 using an I/O interface such as the microphone 190 or the keypad 120. Next, the user edits the template at step 420, again using an I/O interface such as the microphone 190 or the keypad 120. Finally, at step 425 the user saves the edited template in static programmable memory 155. The method 400 then ends at step 430.

Other methods of editing the message templates are also within the scope of the present invention, including connecting the telephone 100 to a host computer using a communication channel such as a USB cable and then downloading or flashing edited templates to the static programmable memory 155.

The method of the present invention may further include message templates that comprise fixed and variable language components. The fixed language components are not changed when a user selects a template and transmits a message. However the variable language components may be changed by the user from message to message. The use of fixed and variable language components can greatly leverage the limited processing power and memory of the telephone 100.

For example, a particular template of a short text message concerning a meeting request might include the following: “Meet me at \$PLACE at \$TIME”. Here the fixed language components are underlined and the variable language components are capitalized and begin with “\$”. Different users of the template may then edit the variable such as \$PLACE to suit their particular circumstances. For example a university student might define the variable \$PLACE as:

\$PLACE = sp|library|dormitory|cafeteria, etc.

Whereas a lawyer might define the variable \$PLACE as:

5

\$PLACE = sp|office|courthouse|home, etc.

In the above, “sp” means a pause or no voice event, and “|” means the logic operator “OR”.

10 Another example of a message template that may be used in the present invention is “Happy \$FESTIVAL.” Here the variable language component \$FESTIVAL may be edited by the user to include:

\$FESTIVAL = sp|birthday|new year|thanksgiving, etc.

15

Using open vocabulary speech recognition, the phone 100 is able to recognize the edited variable language components entered by a user. Because the variable language components consist of discrete sets of variables, the speech recognition processing overhead and memory requirements are
20 minimized. The above method is thus particularly suited for devices having limited processing and memory resources such as mobile phones.

The use of templates including fixed and variable language components increases the efficiency of a speech recognition system for several reasons. First, the fixed language components of a particular template may generally be
25 recognized quickly and efficiently because there are only a modest number of templates saved in the static programmable memory 155 compared with the almost unlimited number of sentence permutations associated with natural language sentence structures. Second, the variable language components may also be recognized efficiently because the intra-sentence location of a variable
30 language component in a message template automatically identifies a discrete set of possible responses. For example, referring to the “Happy \$FESTIVAL” message template given above, the fixed language component “Happy” may

act as a signal such that the processor 110 knows that the subsequent voice input received at the microphone 190 will be the variable language component “\$FESTIVAL.”

5 Although the above embodiments of the present invention are described in relation to a radio telephone 100, the method and apparatus of the present invention could also include other electronic devices that provide text messages such as Personal Digital Assistants (PDAs).

10 Accordingly, the present invention simplifies the steps required for providing and transmitting a text message from a portable electronic device. A text message may be provided through voice inputs rather than through typed characters entered into a small keypad. Further the invention may include open vocabulary speech recognition to avoid the memory intensive requirements of prior art closed vocabulary speech recognition. Open vocabulary speech recognition uses speaker-independent sub-word acoustic models designed to cover all of the acoustic occurrences, or phonemes, of a language. Thus a user is not limited to a predefined vocabulary but can edit the variable language components as described above to include words not found in a dictionary, such as names and locations. The result is that the text messages provided by the present invention may be highly personalized.

15 20 The above detailed description provides preferred exemplary embodiments only, and is not intended to limit the scope, applicability, or configuration of the invention. Rather, the detailed description of the preferred exemplary embodiments provides those skilled in the art with an enabling description for implementing the preferred exemplary embodiments of the invention. It should be understood that various changes may be made in the function and arrangement of elements without departing from the spirit and scope of the invention as set forth in the appended claims.

WE CLAIM:

1. A method of providing a text message, said method comprising
5 the steps of:
receiving an utterance at an input of an electronic device;
performing speech recognition of said utterance guided by user-
defined message templates stored in a memory associated with said
electronic device, wherein speech recognition is defined by matching
10 said utterance with one of said templates to create a matching template;
and
providing a text message from said matching template.
2. The method of claim 1, wherein at least one of said message
15 templates comprises a fixed language component.
3. The method of claim 1, wherein at least one of said message
templates comprises a variable language component.
- 20 4. The method of claim 1, wherein at least one of said message
templates comprises both a fixed and a variable language component.
5. The method of claim 1, wherein said text message is an SMS
25 message.
6. The method of claim 1, further comprising the step of editing
said user-defined message template by receiving typed characters from a
keypad of said electronic device.
- 30 7. The method of claim 1, wherein a component of said text
message is a transcription of said utterance.

8. The method of claim 1, wherein the entirety of said text message is a transcription of said utterance.

9. An electronic device for providing a text message, said device
5 comprising:

a microphone operative to receive an utterance;

a non-volatile memory for storing message templates; and

a processor operative to perform speech recognition of said
utterance guided by said message templates, said processor operative to
10 match said utterance with one of said templates to create a matching
template and to provide a text message from said matching template.

10. The method of claim 9, wherein at least one of said message
templates comprises a fixed language component.

15 11. The method of claim 9, wherein at least one of said message
templates comprises a variable language component.

12. The method of claim 9, wherein at least one of said message
20 templates comprises both a fixed and a variable language component.

13. The device of claim 9, wherein said text message is an SMS
message.

25 14. The device of claim 9, further comprising a keypad operative for
editing said message template.

15. The device of claim 9, wherein said processor is operative to
match said utterance with a plurality of said templates and to calculate a
30 likelihood score for each of said templates.

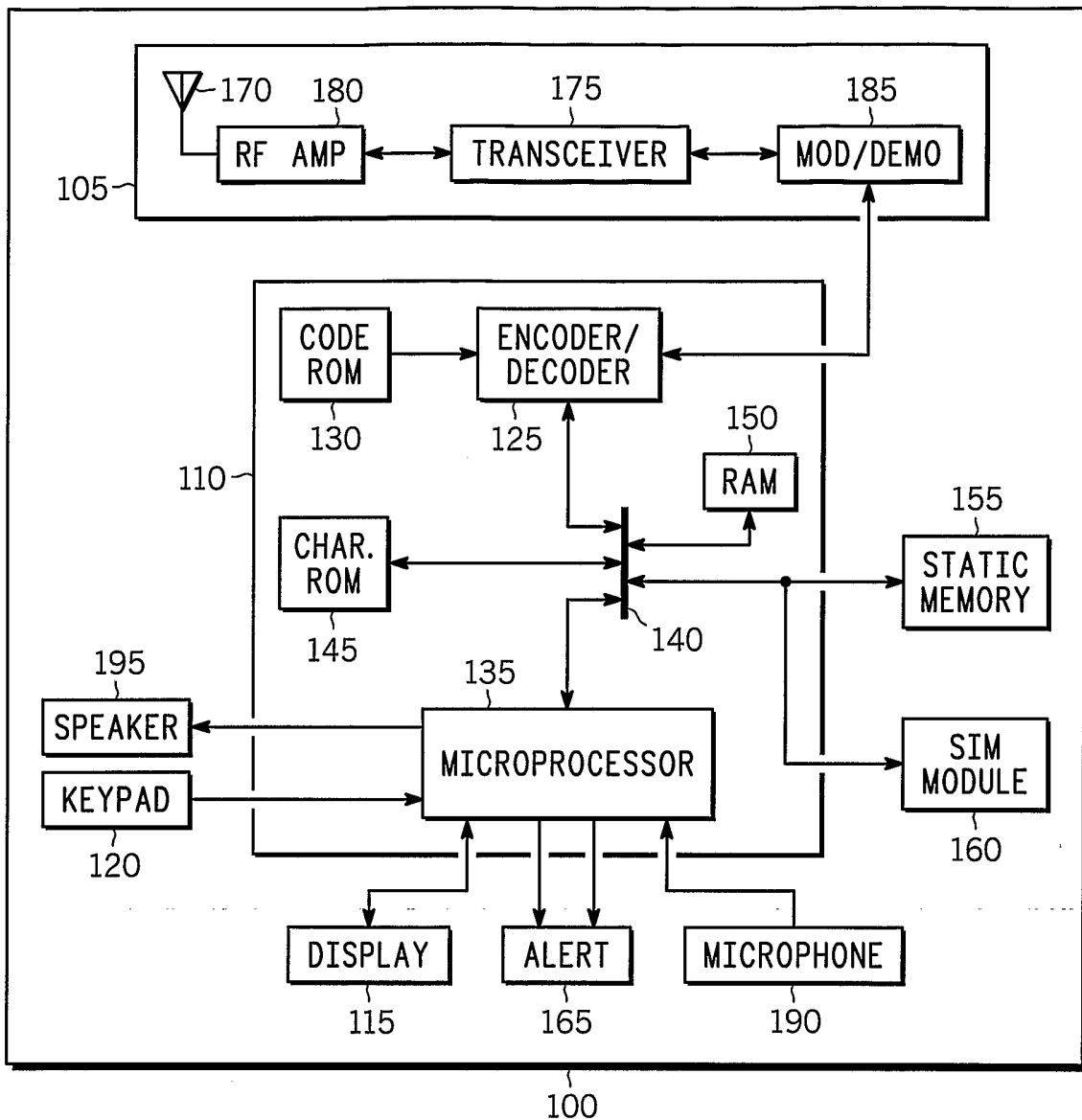
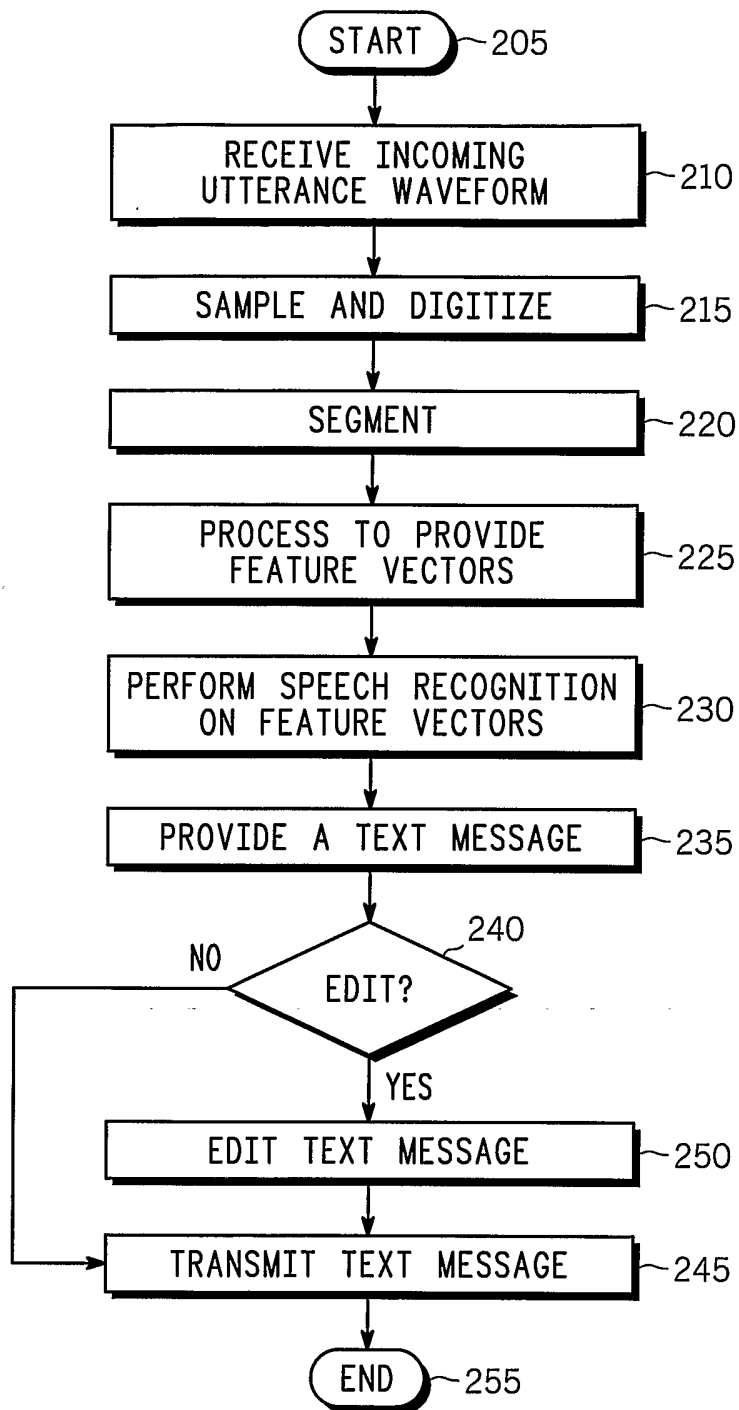


FIG. 1

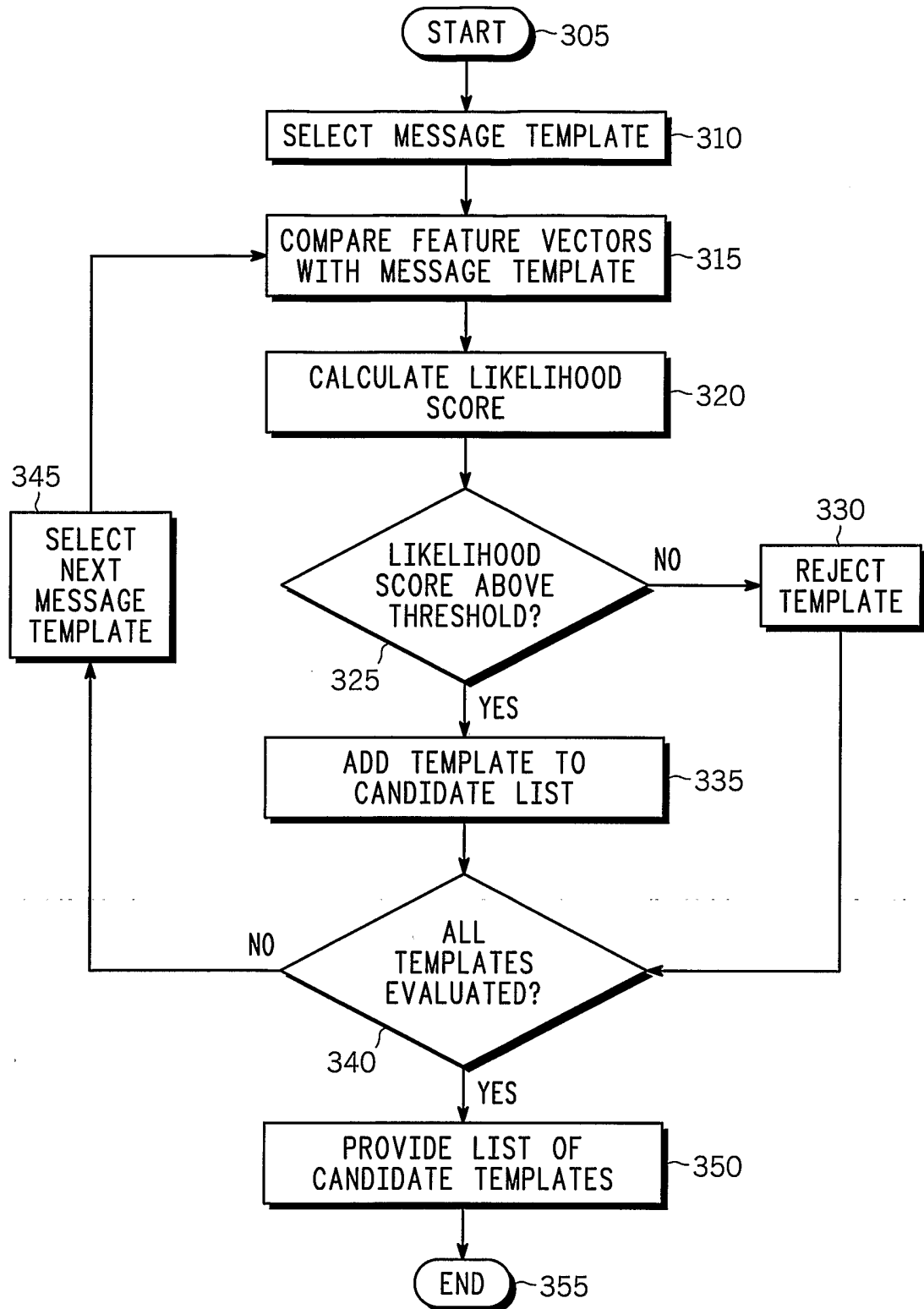
2/4



200

FIG. 2

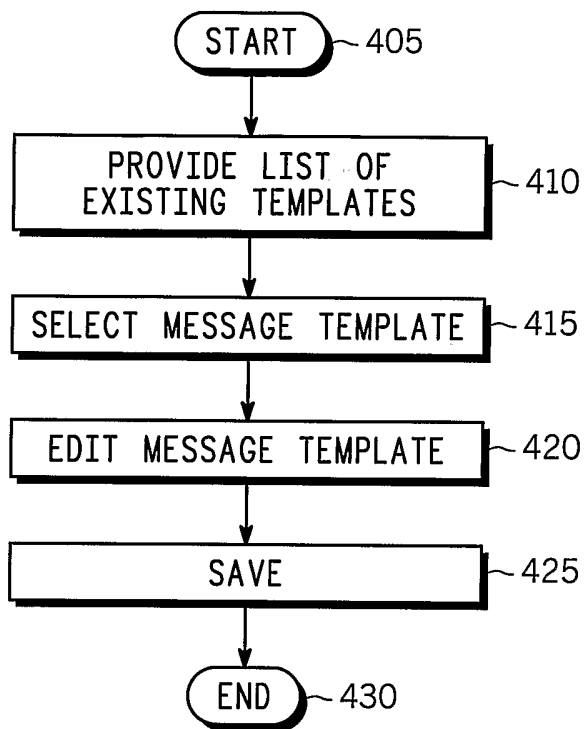
3/4



300

FIG. 3

4/4



400

FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US04/30553

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : H04B 1/38; H04Q 7/20; H04M 1/00
 US CL : 455/563, 466, 550.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 U.S. : 455/412.01, 466, 563; 379/ 88.01, 88.02, 88.03, 88.04

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
 NONE

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
 NONE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P --- Y, P	US 2004/0176139 A1 (WANG) 09 September 2004 (09.09.2004), page1 paragraph 14, page2 paragraphs 19 and 21, page3 paragraphs 25 and 27	1, 5-9, 13-15 ----- 2-4, 10-12
Y, E	US 6,795,808 B1 (STRUBBE et al) 21 September 2004 (21.09.2004), col3 lines 40-42, col4 lines 13-15 and 39-41	2-4, 10-12
A	US 6,526,292 B1 (HENRY) 25 February 2003 (25.02.2003), whole document	1-15
A, E	US 2004/0204115 A1 (GUSLER et al) 14 October 2004 (14.10.2000), whole document	1-15

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means		
"P" document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

Date of mailing of the international search report

10 December 2004 (10.12.2004)
 Name and mailing address of the ISA/US
 Mail Stop PCT, Attn: ISA/US
 Commissioner of Patents
 P.O. Box 1450
 Alexandria, Virginia 22313-1450
 Facsimile No. (703) 305-3230

25 JAN 2005
 Authorized officer
 Wayne Young
 Telephone No. (703)306-5424