

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
10 February 2005 (10.02.2005)

PCT

(10) International Publication Number
WO 2005/013596 A1

(51) International Patent Classification⁷: **H04M 3/56**

(21) International Application Number:
PCT/EP2004/051575

(22) International Filing Date: 22 July 2004 (22.07.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/626,050 24 July 2003 (24.07.2003) US

(71) Applicant (for all designated States except US): **INTERNATIONAL BUSINESS MACHINES CORPORATION** [US/US]; New Orchard Road, Armonk, NY 10504 (US).

(71) Applicant (for MG only): **IBM UNITED KINGDOM LIMITED** [GB/GB]; P.O. Box 41, North Harbour, Portsmouth, Hampshire PO6 3AU (GB).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DAVIS, Brent** [US/US]; 1101 Little Harbor Drive, Deerfield Beach, FL

33441 (US). **JAISWAL, Peeyush** [US/US]; 10511 Plain View Circle, Boca Raton, FL 33498 (US). **MANSEY, Pradeep** [US/US]; 1115 NW 116th Avenue, Coral Springs, FL 33071 (US).

(74) Agent: **LING, Christopher, John**; IBM United Kingdom Limited, Intellectual Property Law, Hursley Park, Winchester, Hampshire SO21 2JN (GB).

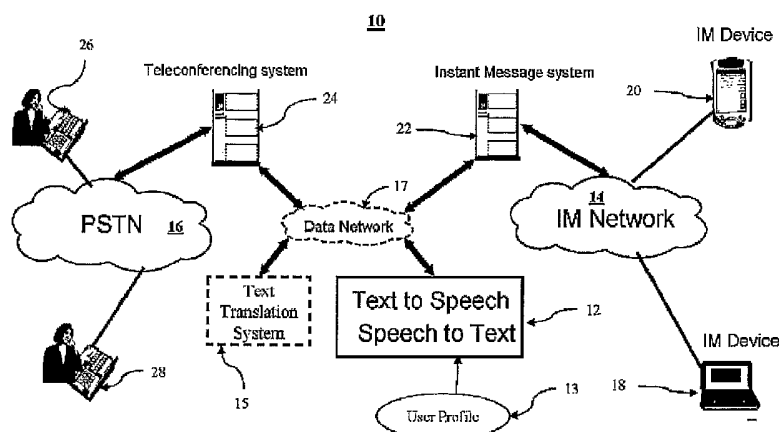
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, 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.

(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),

[Continued on next page]

(54) Title: CHAT AND TELECONFERENCING SYSTEM WITH TEXT TO SPEECH AND SPEECH TO TEXT TRANSLATION

Enhancing Conferencing Service with Tele-chat System



(57) Abstract: A system (10) and method (50) for enabling phone users to participate in an instant messaging based conference can include the steps of receiving (52) a speech input from a telephone (26 or 28) through a teleconferencing system (24), transcribing (54) the speech input to a first text message and transmitting (58) the first text message to a plurality of devices (18, 20, 26 or 28) coupled to an instant messaging network belonging to the instant messaging based conference. The method can further include the steps of receiving (60) a second text message from any one among the plurality of devices on the instant messaging based conference, converting (62) the second text message to a speech output, and transmitting (68) the speech output to the telephone via the teleconferencing system.



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).

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*

Description

CHAT AND TELECONFERENCING SYSTEM WITH TEXT TO SPEECH AND SPEECH TO TEXT TRANSLATION

Technical Field

- [001] This invention relates to the field of telecommunications and more particularly to a telephone conferencing system using real time messaging as well as text-to-speech and speech-to-text conversions.

Background Art

- [002] Conferencing using text-based instant messaging (IM) applications is frequently used as a means of collaboration among home users as well as among enterprises. Unfortunately, not everybody has access or is connected to a LAN or the Internet to participate in such text conferences. Mobile people and computer adverse people may not have access to a networked computer and keyboard to participate in an IM-based conference. Many of these users, while not connected, would still like to participate in an IM conference in a manner that is seamless and familiar to them.
- [003] Several systems attempt to bridge the gap between voice conferencing and instant messaging system, but such existing systems usually have limitations that prevent a truly user friendly experience in a real-time environment. For example, U.S. Patent No. 6,430,604 describes a method of delivering instant messages using cell phones and text pagers but using only text entry. Another patent, WO0135615A2, discusses a method of extending an IM system to telephone messaging systems where a user can log into their voice messaging system to communicate with users on their buddy list.
- [004] Examples of known systems using text-to-speech and speech-to-text include U.S. Patent Publication US2002/0069069 A1, where such system focuses on communications between participants that can and cannot hear voice conversations, or U.S. Patent No. 6,339,754 B1, where text-to-speech and speech-to-text technologies coupled with language translation enable chat and voice conferencing, or U.S. Patent Nos. 6,385,586 B1 or 6,292,769 B1, where text-to-speech and speech-to-text technologies are used to improve language translation between two or more spoken (different language) communications.
- [005] Although there are numerous systems using text-to-speech and speech-to-text technologies, none are ideally suited for augmenting voice (and text) chat over data transmission protocols, wherein such protocols can include chat/instant messaging (IM) and messaging protocols such as SMS. None of the existing systems provide a method of delivering voice messages to the intended recipient in the native format of the recipient's device in the language understood by the recipient while also providing

a real-time collaboration system that does not necessarily require a voice messaging system to gain access to a conference. Thus, a need exists for a system and method that can overcome the detriments described above.

Disclosure of Invention

- [006] Embodiments in accordance with the invention provide a new technique for enhancing a real-time chat channel to enable telephone users to participate in a instant messaging conference.
- [007] In a first aspect of the invention, a method for enabling phone users to participate in an instant messaging based conference includes the steps of receiving a speech input from a telephone through a teleconferencing system, transcribing the speech input to a first text message and transmitting the first text message to a plurality of devices coupled to an instant messaging network belonging to the instant messaging based conference. The method can further include the steps of receiving a second text message from any one among the plurality of devices on the instant messaging based conference, converting the second text message to a speech output, and transmitting the speech output to the telephone via the teleconferencing system.
- [008] In a second aspect of the invention, a system for enabling phone users to participate in an instant messaging based conference can include an input port for receiving a calling party's speech input via a teleconferencing system, a speech-to-text converter for converting the calling party's speech input to a text message for transmission to an instant messaging system, and a text-to-speech converter for converting text messages received from the instant messaging system to a speech output for transmission to the teleconferencing system. The system can further include a telephone coupled to the teleconferencing system and an instant messaging device.
- [009] In a third aspect of the invention, a computer program has a plurality of code sections executable by a machine for causing the machine to perform certain steps. The steps can include the steps of receiving a speech input from a telephone through a teleconferencing system, transcribing the speech input to a first text message, transmitting the first text message to a plurality of devices coupled to an instant messaging network belonging to the instant messaging based conference, receiving a second text message from any one among the plurality of devices on the instant messaging based conference, converting the second text message to a speech output, and transmitting the speech output to the telephone via the teleconferencing system.
- [010] Preferably when converting the second text message a voice signature, associated with any one among the plurality of device coupled to the instant messaging network, is used to provide the speech output with a personalised voice at the telephone.
- [011] Optionally the second text message is converted to the speech output by using text-

to-speech conversion.

- [012] Optionally the first text message is translated to another language in order to provide a translated first text message.
- [013] Optionally the second text message is translated into another language to provide a translated second text message for subsequent speech output.
- [014] Preferably the text message comprises is transmitted as a text.
- [015] Preferably the second text message is converted using text-to-speech synthesis. This may be performed by a speech synthesiser which optionally uses a voice signature of the called party in producing audible output.
- [016] Optionally a telephone is coupled to the teleconferencing system. Optionally the system further comprises an instant messaging device such as a personal digital assistant, a laptop computer, and a smartphone. Preferably the instant messaging device has a display for displaying the text message from the calling party and/or text messages from the instant messaging device.
- [017] Optionally if either text message is translated it is transmitted to either an instant messaging device as text output or to a telephone coupled to the teleconferencing system as a speech output.
- [018] Preferably text streams are received and transmitted over an instant messaging/chat system in substantially real-time.
- [019] Optionally text streams are received and transmitted over a messaging system using data transmission protocols.
- [020] Optionally a user profile is used for converting at least one among text messages from an instant messaging device into a customized speech output for transmission to the calling party and text messages from the calling party to alternate text messages as defined by a user.

Brief Description of the Drawings

- [021] The invention will now be described by way of example only, with reference to a preferred embodiment thereof, as illustrated in the accompanying drawings, in which:
- [022] FIG. 1 is a flow diagram illustrating an exemplary telecommunications system illustrating an enhanced conferencing system using instant messaging; and
- [023] FIG. 2 is a flow chart illustrating a method for enabling phone users to participate in an instant messaging based conference.

Best Mode for Carrying Out the Invention

- [024] Embodiments in accordance with the invention can provide a solution for enabling telephone users to participate in an IM-based conference. In a typical IM based conference, all participants are connected to an IM server over a data network and every participant's text message is broadcasted to all the parties in the conference. In

accordance with one embodiment, users can call into a system using their wireline or wireless phone, hear messages typed in by the IM participants and can participate by speaking their messages which can be transcribed into text and broadcast to the IM participants. Such a system can synthesize text messages into speech, transcribe text speech into text and essentially bridge an IM system and a teleconferencing system. Additionally, the system can be personalized by the user to provide a rich end-user experience.

[025] A system 10 for enabling phone users to participate in an instant messaging based conference can include a device 12 that serves as a bridge between a teleconferencing system 24 and an instant messaging system 22. The device 12 can be directly coupled between the teleconferencing system 24 and the instant messaging system 22 or coupled between such systems via an optional data network 17 as shown. Operationally, a conventional phone (26 or 28) on a network 16 such as a PSTN can couple and provide input to and receive input from the device 12 via the teleconferencing system 24. When the phone (26 or 28) provides speech input meant for devices (18 or 20) and their corresponding users on an IM conference, the device 12 can transcribe the speech input to a text message that can be broadcast to all or some of the devices in the IM conference. The devices 18 and 20 can be personal digital assistants, laptop computers, desktop computers, smartphone or essentially any computing device capable of receiving and displaying text messages. The devices 18 and 20 can be coupled to the IM conference via an IM network 14. The device 12 can transmit the text messages to such IM conference participants via an IM system or server 22 and the IM network 14.

[026] A conventional phone (26 or 28) participating in the IM conference can also receive IM messages from other devices in the form of a synthesized speech output. For example, a user inputting text on IM device (18 or 20) would transmit their text message to the device 12 via the IM network 14 and the IM system 22. The device 12 can convert the text message to speech and forward or transmit the speech to the phone 26 or 28 via the system 24 and network 16. Optionally, user profiles 13 (with voice prints or other indicia or a particular user) for devices 18 and 20 (and perhaps users of convention phones that have provided some form of identification when entering the IM conference) can enhance the user experience on conventional phones by reconstructing speech having a simulated voice print of the sending party.

[027] Yet another option would allow for language translation of the text received or converted at device 12. Thus, a user profile 13 corresponding to an IM device 18 can direct the device 12 to translate text received in one language to be transmitted to the device 18 in another language using an optional text translation system 15 coupled to the device 12 for example. Likewise, a phone 26 having a user profile 13 can direct

text messages intended for the phone 26 to be translated into another language (using the optional text translation system 15 for example) before speech synthesis so that the user at phone 26 hears speech in a preferred language.

[028] With reference to FIG. 2, a flow chart illustrates a method 50 of enabling a phone user to participate in a IM-based conference. Operationally, the system 10 as shown in FIG. 1 would receive a speech input from a telephone through a teleconferencing system at step 52. At step 54, the speech input can be transcribed into a first text message. Optionally, at step 56, the first text message can be translated to another language to provide a translated first text message. The user profile 13 of FIG. 1 can be used to set this additional capability if desired. The first text message can be transmitted at step 58 to a plurality of devices coupled to an instant messaging network belonging to the IM based conference. The first text message can be transmitted as a text stream.

[029] With reference again to FIG. 2, the system can receive at step 60 a second text message from any one among the plurality of devices on the IM-based conference. At step 62, the system can convert the second text message to a speech output, preferably by using text-to-speech conversion or synthesis. Once again, the system can optionally translate the second text message to another language to provide a translated second text message for subsequent speech output at step 64. Another option at step 66 enables the system to use a voice signature associated with any one among the plurality of devices coupled to the IM network to provide a speech output with a personalized or customized voice heard at the telephone. Finally, at step 68, the speech output can be transmitted to the telephone via the teleconference system.

[030] It should be understood that the present invention can be realized in hardware, software, or a combination of hardware and software. The present invention can also be realized in a centralized fashion in one computer system, or in a distributed fashion where different elements are spread across several interconnected computer systems. Any kind of computer system or other apparatus adapted for carrying out the methods described herein is suited. A typical combination of hardware and software can be a general purpose computer system with a computer program that, when being loaded and executed, controls the computer system such that it carries out the methods described herein.

[031] The present invention also can be embedded in a computer program product, which comprises all the features enabling the implementation of the methods described herein, and which when loaded in a computer system is able to carry out these methods. Computer program or application in the present context means any expression, in any language, code or notation, of a set of instructions intended to cause a system having an information processing capability to perform a particular function

either directly or after either or both of the following: a) conversion to another language, code or notation; b) reproduction in a different material form.

Claims

- [001] A method for enabling phone users to participate in an instant messaging based conference, comprising the steps of: receiving a speech input from a telephone through a teleconferencing system; transcribing the speech input to a first text message; transmitting the first text message to a plurality of devices coupled to an instant messaging network belonging to the instant messaging based conference; receiving a second text message from any one among the plurality of devices on the instant messaging based conference; converting the second text message to a speech output; and transmitting the speech output to the telephone via the teleconferencing system.
- [002] The method of claim 1, wherein the step of converting the second text message further comprises the step of using a voice signature associated with any one among the plurality of device coupled to the instant messaging network to provide the speech output with a personalized voice at the telephone.
- [003] The method of claim 1, wherein the method further comprises the step of translating the first text message to another language to provide a translated first text message.
- [004] The method of claim 1, wherein the method further comprises the step of translating the second text message to another language to provide a translated second text message for subsequent speech output.
- [005] A system for enabling phone users to participate in an instant messaging based conference, comprising: an input port for receiving a calling party's speech input via a teleconferencing system; a speech-to-text converter for converting the calling party's speech input to a text message for transmission to an instant messaging system; and a text-to-speech converter for converting text messages received from the instant messaging system to a speech output for transmission to the teleconferencing system.
- [006] The system of claim 5, wherein the system further comprises an instant messaging device selected from the group of devices comprising a personal digital assistant, a laptop computer, and a smartphone.
- [007] The system of claim 5, wherein the system further comprises a translator for translating the text message into another language for transmission to at least one among an instant messaging device as text and to a telephone coupled to the teleconferencing system as a speech output.
- [008] The system of claim 5, wherein the system further comprises a text-to speech synthesizer which uses a voice signature of the called party in producing the audible output.

- [009] The system of claim 5, wherein the system further comprises a user profile for converting at least one among text messages from an instant messaging device into a customized speech output for transmission to the calling party and text messages from the calling party to alternate text messages as defined by a user.
- [010] A machine-readable storage, having stored thereon a computer program having a plurality of code sections executable by a machine for causing the machine to perform the method of any one of claims 1 to 4.

[Fig. 001]

Enhancing Conferencing Service with Tele-chat System

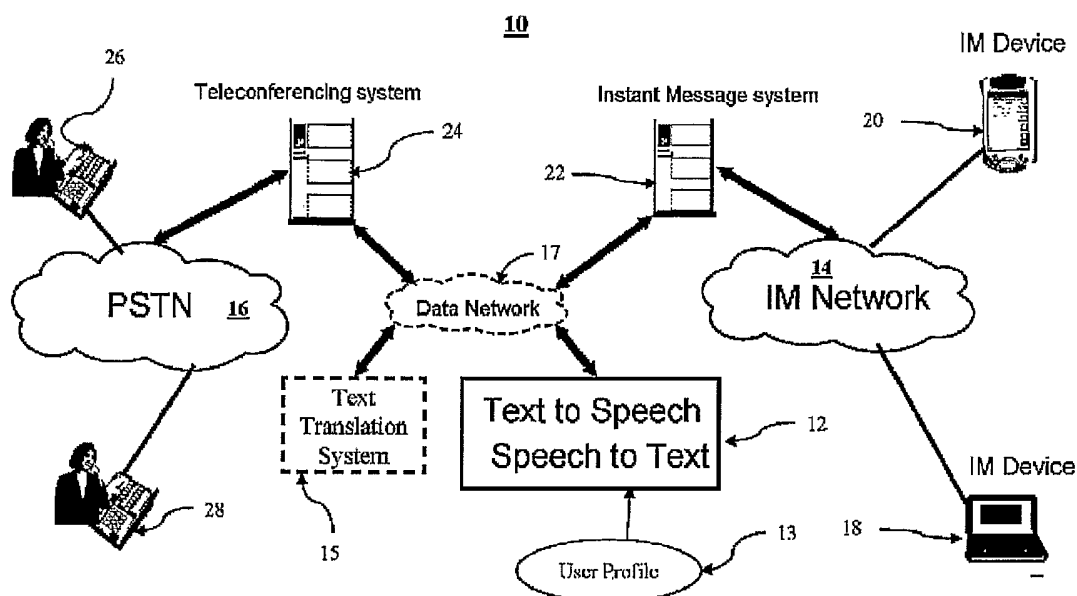


FIG. 1

[Fig. 002]

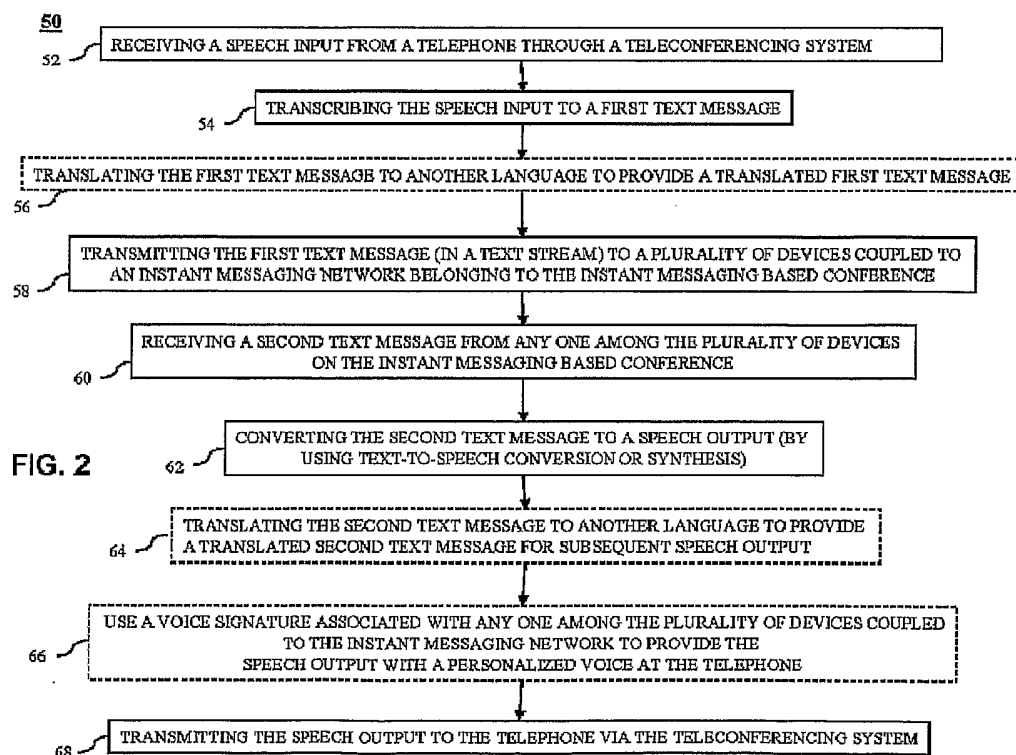


FIG. 2

INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP2004/051575

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04M3/56

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)

IPC 7 H04M G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 848 134 A (KOGA TADAHARU ET AL) 8 December 1998 (1998-12-08)	1,5,6,10
Y	abstract figure 1 column 1, line 1 - column 3, line 16	2-4,7-9
X	CA 2 281 147 A (AT & T CORP) 26 February 2001 (2001-02-26) abstract	1,5,6,10
Y	US 2002/193994 A1 (HANSON BRIAN ET AL) 19 December 2002 (2002-12-19) abstract	2,8
Y	WO 02/43360 A (LPS ASSOCIATES LLC) 30 May 2002 (2002-05-30) page 18, line 23 - page 19, line 19	3,4,7,9



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"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)

"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

"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

"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

"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.

"&" document member of the same patent family

Date of the actual completion of the international search

15 November 2004

Date of mailing of the international search report

26/11/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Tillgren, M

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP2004/051575

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
US 5848134	A	08-12-1998	JP	9214486 A	15-08-1997
CA 2281147	A	26-02-2001	CA	2281147 A1	26-02-2001
US 2002193994	A1	19-12-2002	WO	02080140 A1	10-10-2002
WO 0243360	A	30-05-2002	AU	3941102 A	03-06-2002
			WO	0243360 A2	30-05-2002