



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
02.05.2012 Bulletin 2012/18

(51) Int Cl.:
G10L 13/04 (2006.01)

(21) Application number: **10306196.6**

(22) Date of filing: **29.10.2010**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME

(72) Inventor: **Hernandez Martinez, Julian**
28041 Madrid (ES)

(74) Representative: **Cabinet Plasseraud**
52, rue de la Victoire
75440 Paris Cedex 09 (FR)

(71) Applicant: **FRANCE TELECOM**
75015 Paris (FR)

(54) **Method of and apparatus for providing audio data corresponding to a text**

(57) A method of and apparatus for receiving a request from a communication device for access to a webpage comprising a text portion; receiving from a first content provider server the webpage comprising the text portion; identifying the text portion; embedding into the webpage a link for providing audio data corresponding to the text portion; transmitting the webpage embedded with the link to the communication device; receiving from the communication device a request for audio data corresponding to the text portion; generating audio data corresponding to the text portion using a text to speech converter; transmitting the audio data to said communication device; wherein during the step of transforming the text portion into audio data a preliminary audio data content, is provided to the communication device so that the preliminary audio data content can be played on the communication device while the audio data corresponding to the text portion is being generated and streamed to the communication device.

responding to the text portion; generating audio data corresponding to the text portion using a text to speech converter; transmitting the audio data to said communication device; wherein during the step of transforming the text portion into audio data a preliminary audio data content, is provided to the communication device so that the preliminary audio data content can be played on the communication device while the audio data corresponding to the text portion is being generated and streamed to the communication device.

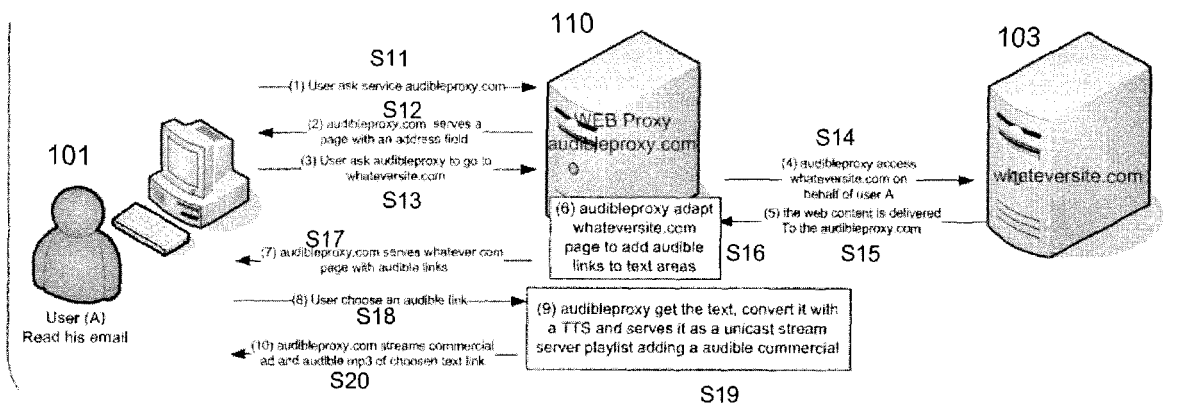


FIG.3

Description

Field of the Invention

[0001] The present invention relates in general to a method and apparatus for providing audio data corresponding to a text. Particularly but not exclusively the invention relates to a method and apparatus for providing audio data corresponding to at least one text portion of a web page.

Background of the Invention

[0002] Web pages are documents or information resources provided by a content server or computer that can be accessed via the Internet via a web browser and displayed on a user terminal. Web pages typically include portions of a text and other data content providing information to the user. In order to be able to retrieve the textual information the user must be capable of reading the text. Not all users however are capable of reading text provided on such web pages; blind or visually impaired users or users with reading difficulties, for example may not be capable of reading the text displayed on the user terminal.

[0003] Software applications have been developed for providing the user with access to the information displayed on a terminal screen by means of text-to-speech convertors or by means of a Braille display. Such software applications are stored on the terminal and can only convert text data to audio data as soon as it has been received at the terminal thereby generating a delay in the user receiving the data in a comprehensible format. US 2006/0111911 describes a method and apparatus for generating audio files from web pages in which the audio data corresponding to a web page may be generated by a server remote to the user terminal in response to a request by the user, and then transmitted to the user terminal. However, delays in buffering and streaming the audio data result in a waiting time for the user to receive the audio data content.

Summary of the Invention

[0004] Accordingly, in order to better address one or more of the foregoing concerns, a first aspect of the invention provides a method of providing audio data corresponding to a text, the method comprising: receiving a request from a communication device for access to a webpage comprising a text portion; receiving from a first content provider server the webpage comprising the text portion; identifying the text portion; embedding into the webpage a link for providing audio data corresponding to the text portion; transmitting the webpage embedded with the link to the communication device; receiving from the communication device a request for audio data corresponding to the text portion; generating audio data corresponding to the text portion using

a text to speech convertor; transmitting the audio data to said communication device; wherein during the step of generating audio data from the text portion a preliminary audio data content, is provided to the communication device so that the preliminary audio data content can be played on the communication device while at least a portion of the audio data corresponding to the text portion is being generated and streamed to the communication device.

[0005] A second aspect of the invention provides a network device, such as a server, for providing audio data corresponding to a text, the network device comprising: a transceiver for receiving a request from a communication device for access to a webpage comprising a text portion and for receiving from a first content provider server the webpage comprising the text portion; a processor for identifying the text portion and for embedding into the webpage a link for providing audio data corresponding to the text portion; a text to speech convertor for generating audio data corresponding to the text portion; a buffer for buffering the generated audio data in response to the link being activated; an audio data streamer for transmitting a preliminary audio data content, to the communication device while at least a portion of the audio data corresponding to the text portion is being generated and buffered so that the preliminary audio data content can be played on the communication device before the audio data corresponding to the text portion is transmitted

[0006] In embodiments of the invention:

- the preliminary audio data content comprises audio advertising data for promoting a product or service.
- audio advertising data related to the content of the web page is selected. For example, the text of the text portion is analysed for selection of the audio advertising data related to the content of the web page.
- the language of the or each text portion is detected prior to generating the audio data;
- the preliminary audio data content and the audio data is merged in a play list.

[0007] At least parts of the methods according to the invention may be computer implemented. The methods may be implemented in software on a programmable apparatus. They may also be implemented solely in hardware or in software, or in a combination thereof.

[0008] Since at least parts of the present invention can be implemented in software, the present invention can be embodied as computer readable code for provision to a programmable apparatus on any suitable carrier medium. A tangible carrier medium may comprise a storage medium such as a floppy disk, a CD-ROM, a hard disk drive, a magnetic tape device or a solid state memory device and the like. A transient carrier medium may include a signal such as an electrical signal, an electronic signal, an optical signal, an acoustic signal, a magnetic signal or an electromagnetic signal, e.g. a microwave or RF signal.

Brief Description of the Drawings

[0009] Embodiments of the invention will now be described, by way of example only, and with reference to the following drawings in which:-

Figure 1 is a schematic diagram of the architecture of a system for providing audio data corresponding to a text according to at least one embodiment of the invention;

Figure 2 is a block diagram illustrating some components of a proxy server for providing audio data according to some embodiments of the invention;

Figure 3 is a communication diagram of a method of providing audio data corresponding to a text according to a particular embodiment of the invention;

Figure 4 is a flow chart of steps of a method of providing audio data corresponding to a text according to a particular embodiment of the invention; and

Figure 5 is a flow chart of steps of a method of providing audio data corresponding to a text according to a particular embodiment of the invention.

Detailed description

[0010] A first embodiment of a method of providing audio data corresponding to a text according to at least one embodiment of the invention will be described with reference to Figures 1 to 5.

[0011] Figure 1 illustrates a network system in which embodiments of the invention may be implemented. The network system 100 comprises a user terminal 101 operable to receive and display a web page, a content server provider 103 for providing data content of the web page, and a proxy server 110 for providing audio data content corresponding to the text portions of the web page. The entities are interconnected via an internet network 120.

[0012] It will be understood that in the context of the present invention the user terminal 101 may be any type of fixed or mobile data communication terminal capable of interacting with a network to receive a web page and being configured to display the web page on a display screen of the terminal. The user terminal 101 will also be provided with an audio data processing module and loud speaker for playing back audio data. With reference to Figure 2 the proxy server 110 comprises a text to speech (TTS) engine 111 for converting textual data into audio data, an audio streaming device 112 for buffering and streaming audio data for transmission a memory 113 for storing audio data, a network interface 114 for receiving and transmitting data and a processor 115. Advertising audio data for promoting a product or service may be stored in the memory

[0013] With reference to Figures 3 to 5 in step S11 of the method of the first embodiment of the invention the user sends a request from the user terminal 101 to the proxy server 110 to request the service for provision of audio data corresponding to a web page. In response to the request, in step S12 the proxy server 110 transmits a request page with an address field for the user to specify the web page he or she wishes to access. The user fills in the address field with the address of a web page (in this example www.whateversite.com) hosted by content server 103 and transmits the request page to the proxy server 103 in step S13. It will be appreciated that in alternative embodiments of the invention the user may identify the web address in the initial request sent to the proxy server 110, and the request may take a number of different forms, for example by directly filling in fields of a dedicated web page, by transmitting an email, etc.

[0014] In step S14 the proxy server 110 accesses the data content server 103 hosting the requested web page of whateversite.com. In step S15 the data content server 103 delivers the web data content of the requested web page to the proxy server 110. The web page contains a number of text portions in its data content. After receiving the web page content the proxy server 110 identifies the text portions of the web page and in step S16 embeds into the web page a link to each text portion for providing audio data corresponding to the respective text portion when requested. The web page with the one or more embedded links is transmitted from the proxy server 110 to the user terminal 101 in step S17. In step S18 the user can select a link for providing an audio version of a text portion of the web page by clicking on the corresponding link. In response to clicking of the corresponding link the proxy server 110 receives a request for audio data corresponding to the text portion selected by the link. In step S19 the TTS module 111 of the proxy server 110 begins to convert the selected text portion into audible voice data and starts filling an audio data buffer of the audio stream module 112 for streaming the audio data to the user terminal 101 in step S20. While the audible voice data is being buffered and streamed, advertising audio data stored in the memory 113 of the proxy server 110 is transmitted to the user terminal 101. In some embodiments of the invention advertising audio data related to the content of the web page or the selected text portion of the web page is selected from a database of advertising audio data stored in the memory 113. The advertising audio data is then played back on the user terminal 101 while the user is waiting for the audible voice data corresponding to the selected text portion of the web page. The advertising audio data and the audible voice data corresponding to the selected text portion may be merged together in a play list and streamed as a playlist to the user terminal 101.

[0015] After the advertising audio data has been played back the audible voice data received at the user terminal is played back. The user may then select another text portion of the web page or request audio data content

of another web page.

[0016] With reference to Figure 5 the operation of the proxy server 110 will be described in more detail. In step S22 the proxy server 110 receives the text source requested by the user from data content server 103. The portion or portions of text of the web page are identified in the web page data content and the text data is adapted for input into the text to speech convertor 111. In step S24 the language of the text is detected for the purposes of the text to speech conversion. The language may be, for example, English, French, German, or any other language. When the user clicks on the link to obtain audible voice data of the text portion, in step S25A the text portion is analysed, for example by analysing keywords, in order to select advertising data content corresponding to the text portion. The most relevant advertising data is selected and associated with the respective text portion to provide preliminary audio file A. The advertising data in audio format may be stored in the memory 113 of the proxy server 110 as an audio file and an appropriate URL may be found for the advertising data content in the audio stream module 112. In step S25B which can be performed in parallel to step S25A a second audio file B (e.g. in MP3 format) and a URL to store the audio file in the streaming module 112 are assigned for the audio data corresponding to the text portion. In step S26B, the proxy server 110 starts to generate by means of the TTS module 111 the audible voice data corresponding to the text portion in mp3 file B in the assigned URL of the audio stream module 112. The voice data generated will be in the language detected for the text. In step S26A a play list generator merges the audio advertising data linked to the text portion - preliminary audio data file A, with the audible voice data version of the text portion - audio data file B into a playlist, for example a M3U play list, for streaming to the user terminal 101. The audio file A of the advertising audio data is streamed to the user in step S27A while at least part of the audio data file B for the audible voice data corresponding to the text portion is being generated and stored in step S27B. In step S28 the audible voice data of file B which has been generated is streamed to the user following streaming of the advertising audio data - file B. It will be appreciated that while a portion of the audible voice data already generated is being streamed to the user device 1010 further audible voice data corresponding to the text portion may still be generated by the text to speech module 111. The user receives the audio advertising data before the audible voice data of the text portion and the audio advertising data is played back to the user while the user awaits the requested audio data of the text portion.

[0017] In some embodiments of the invention the generated audible voice data corresponding to a selected text portion may be stored in the memory 113 of the proxy server so that it may be accessed in the case where the proxy server receives another request for audio data content corresponding to that text portion.

[0018] The methods and apparatus according to the

embodiments of the invention enable an end user with reading or vision impediments to listen to web text content without the need for screen reader software to be installed on the user terminal which may slow down the processing efficiency of the user terminal.

[0019] Moreover, the methods and apparatus according to the embodiments of the invention enable the delay time between a request for audible data corresponding to a text portion being made, and the audible data being received to be effectively managed and used. The user can be entertained while awaiting the requested audio content while income can be generated by providing an advertising service.

[0020] The method according to the embodiments of the invention can find applications where text content of web pages can be made available to users with reading or visual impairments in an audible format.

[0021] Although the present invention has been described hereinabove with reference to specific embodiments, the present invention is not limited to the specific embodiments, and modifications will be apparent to a skilled person in the art which lie within the scope of the present invention.

[0022] For instance, while in the foregoing examples the preliminary audio data transmitted to the user comprises audio advertising data content, it will be appreciated that in alternative embodiments of the invention the preliminary audio data content may comprises other audio data content for entertaining the user.

[0023] It will also be appreciated that in embodiments of the invention some steps of the process may be carried out prior to the user clicking on the link for obtaining the audio data. For example the steps of S23 to S25B may be carried out before or after the user clicks on the link for providing the audio version of the text portion.

[0024] Many further modifications and variations will suggest themselves to those versed in the art upon making reference to the foregoing illustrative embodiments, which are given by way of example only and which are not intended to limit the scope of the invention, that being determined solely by the appended claims. In particular the different features from different embodiments may be interchanged, where appropriate.

[0025] In the claims, the word "comprising" does not exclude other elements or steps, and the indefinite article "a" or "an" does not exclude a plurality. The mere fact that different features are recited in mutually different dependent claims does not indicate that a combination of these features cannot be advantageously used. Any reference signs in the claims should not be construed as limiting the scope of the invention.

Claims

1. A method for providing audio data corresponding to a text, the method comprising:

- receiving a request from a communication device for access to a webpage comprising a text portion;
 receiving from a first content provider server the webpage comprising the text portion;
 identifying the text portion;
 embedding into the webpage a link for providing audio data corresponding to the text portion;
 transmitting the webpage embedded with the link to the communication device;
 receiving from the communication device a request for audio data corresponding to the text portion;
 generating audio data corresponding to the text portion using a text to speech convertor;
 transmitting the audio data to said communication device;
 wherein during the step of generating the audio data from the text portion a preliminary audio data content, is provided to the communication device so that the preliminary audio data content can be played on the communication device while at least a portion of the audio data corresponding to the text portion is being generated and streamed to the communication device.
2. A method according to claim 1 wherein the preliminary audio data content comprises audio advertising data for promoting a product or service.
 3. A method according to claim 2, further comprising searching for audio advertising data related to the content of the web page.
 4. A method according to claim 3 wherein the text of the text portion is analysed for selection of the audio advertising data related to the content of the web page.
 5. A method according to any one of the preceding claims further comprising determining the language of the or each text portion.
 6. A method according to any one of the preceding claims further comprising merging the preliminary audio data content and the audio data in a play list.
 7. A method according to any one of the preceding claims, further comprising preliminary steps of generating a request page with an address field and transmitting the request page to a user terminal for requesting the webpage.
 8. A network device for providing audio data corresponding to a text, the network device comprising:
 - a transceiver for receiving a request from a communication device for access to a webpage comprising a text portion and for receiving from a first content provider server the webpage comprising the text portion;
 - a processor for identifying the text portion and for embedding into the webpage a link for providing audio data corresponding to the text portion;
 - a text to speech convertor for generating audio data corresponding to the text portion;
 - a buffer for buffering the generated audio data in response to the link being activated;
 - an audio data streamer for transmitting a preliminary audio data content, to the communication device while at least a portion of the audio data corresponding to the text portion is being generated and buffered so that the preliminary audio data content can be played on the communication device before the audio data corresponding to the text portion is transmitted
 9. A device according to claim 8 wherein the preliminary audio data content comprises audio advertising data for promoting a product or service.
 10. A device according to claim 8 or 9, further comprising searching means for searching for audio advertising data related to the content of the web page.
 11. A device according to any one of claims 8 to 10, further comprising language identifications means for determining the language of the or each text portion.
 12. A device according to any one of claims 8 to 11, further comprising a play list generator for merging the preliminary audio data content and the audio data in a play list
 13. A computer program product for a data-processing device, the computer program product comprising a set of instructions which, when loaded into the data-processing device, causes the device to perform the steps of the method as claimed in any of claims 1 to 7.
 14. A computer-readable medium carrying one or more sequences of instructions of the computer program product of claim 13.

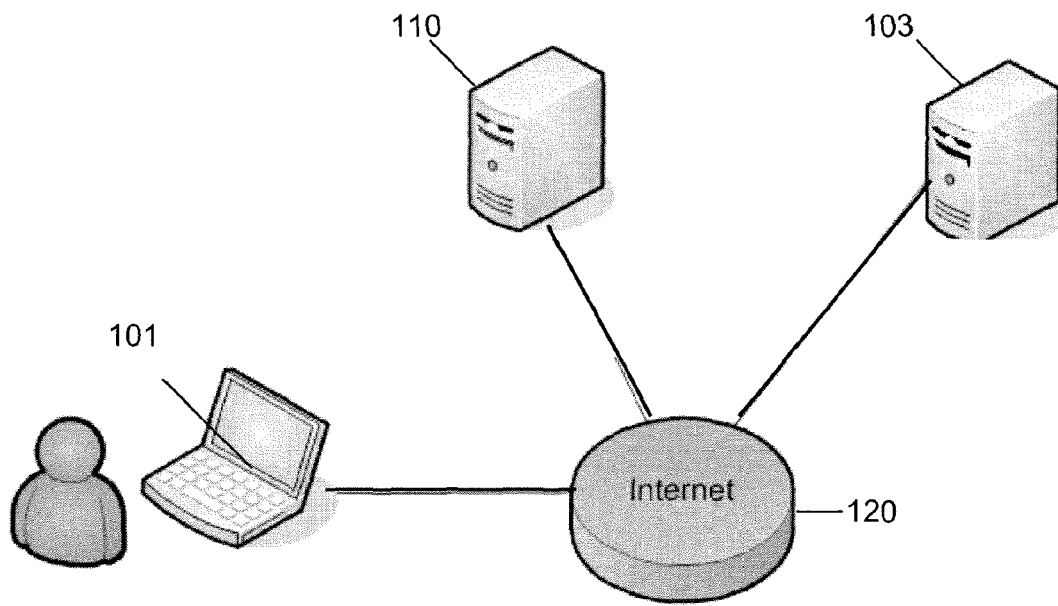


FIG. 1

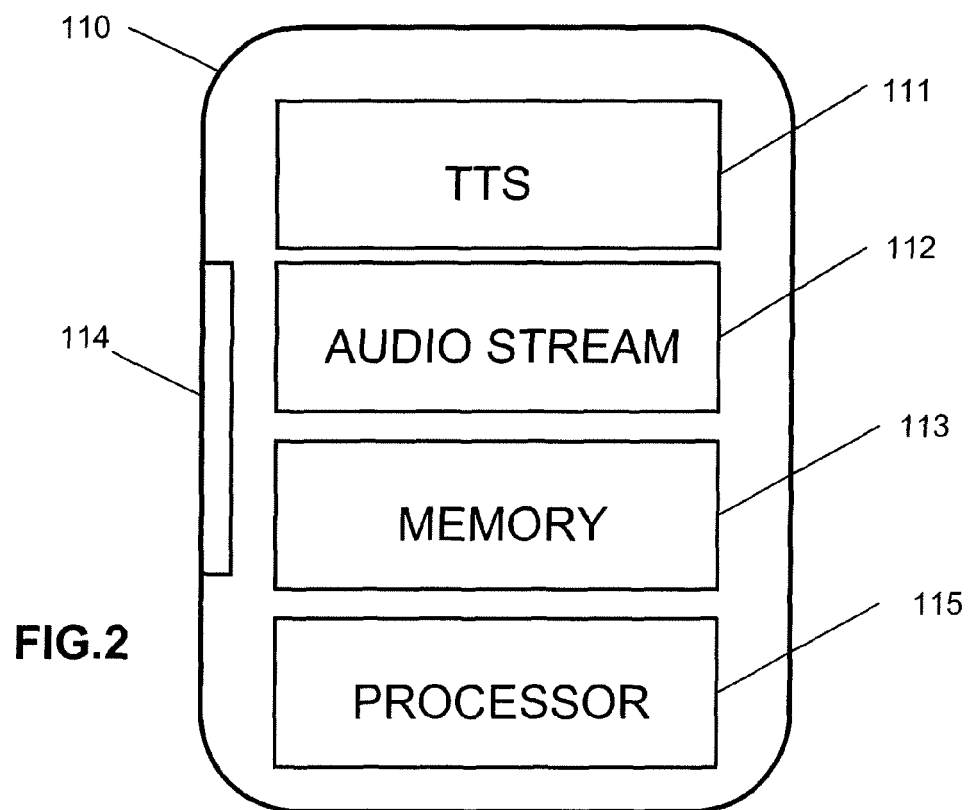


FIG.2

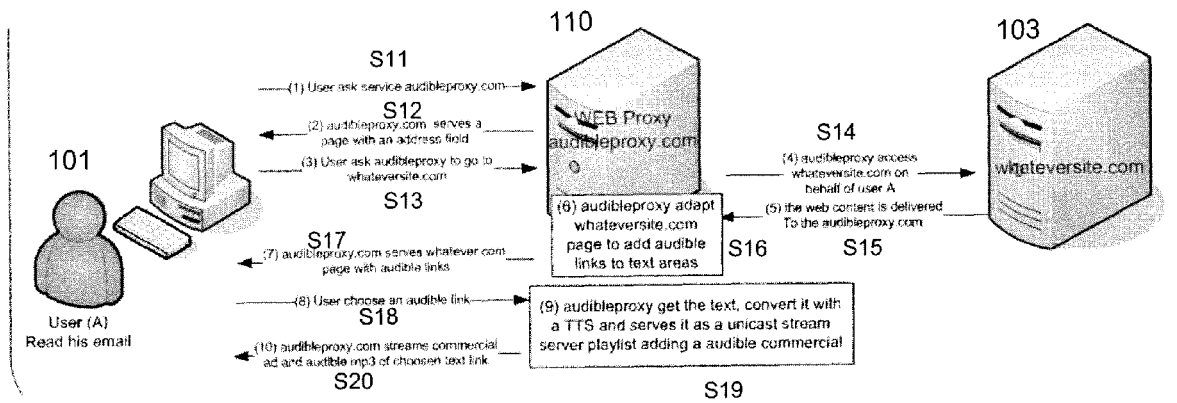


FIG.3

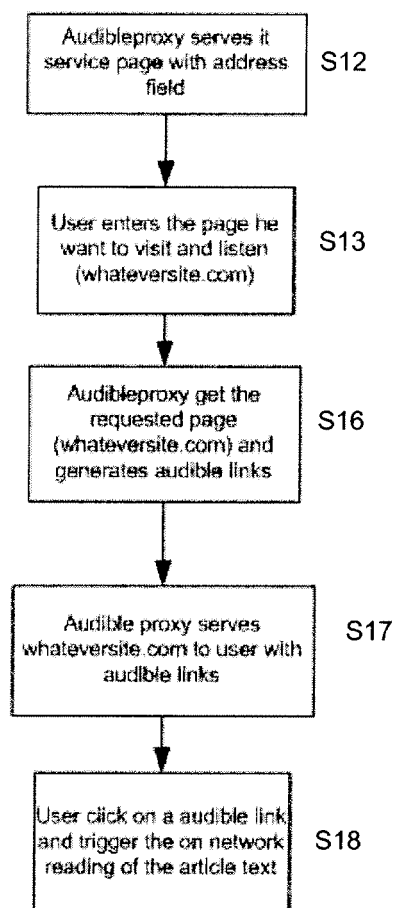


FIG.4

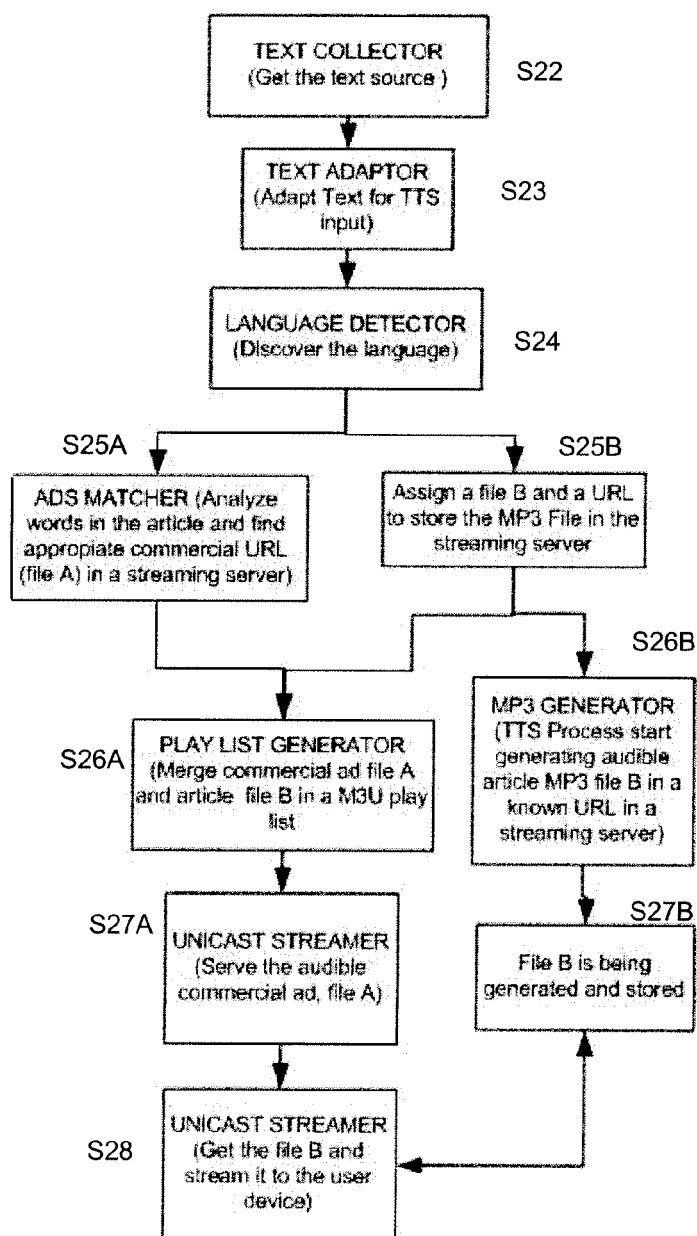


FIG.5



EUROPEAN SEARCH REPORT

Application Number
EP 10 30 6196

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
A,D	US 2006/111911 A1 (MORFORD TIMOTHY B [US]) 25 May 2006 (2006-05-25) * abstract; figures 2, 12 * * paragraph [0092] - paragraph [0094] * -----	1-14	INV. G10L13/04
A	US 2008/059189 A1 (STEPHENS JAMES H [US]) 6 March 2008 (2008-03-06) * abstract; figure 2 * * paragraphs [0021], [0089] * * paragraph [0095] - paragraph [0100] * -----	1-14	
			TECHNICAL FIELDS SEARCHED (IPC)
			G10L
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 25 February 2011	Examiner Zimmermann, Elko
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

1
EPO FORM 1503 03.92 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 10 30 6196

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

25-02-2011

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2006111911 A1	25-05-2006	NONE	
US 2008059189 A1	06-03-2008	NONE	

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20060111911 A [0003]