(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 073 036 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
14.12.2005 Bulletin 2005/50

(51) Int Cl.[7]: **G10L 13/04**

(21) Application number: **00306355.9**

(22) Date of filing: **26.07.2000**

(54) **Parsing of downloaded documents for a speech synthesis enabled browser**

Analyse von heruntergeladenen Dokumenten für einen Browser mit Sprachsynthese-Fähigkeit

Analyse de documents téléchargés pour un navigateur équipé d'un dispositif de synthèse de la parole

(84) Designated Contracting States:
DE FR GB

(30) Priority: **30.07.1999 JP 21721999**

(43) Date of publication of application:
31.01.2001 Bulletin 2001/05

(73) Proprietor: **CANON KABUSHIKI KAISHA**
Tokyo (JP)

(72) Inventors:
- **Otsuka, Mitsuru**
  Ohta-ku, Tokyo (JP)
- **Miyazaki, Yasuko**
  Ohta-ku, Tokyo (JP)
- **Kamiyama, Shinichi**
  Ohta-ku, Tokyo (JP)
- **Aso, Takashi**
  Ohta-ku, Tokyo (JP)

(74) Representative:
Beresford, Keith Denis Lewis et al
BERESFORD & Co.
16 High Holborn
London WC1V 6BX (GB)

(56) References cited:
EP-A- 0 542 628
EP-A- 0 848 373
WO-A-97/23973
GB-A- 2 336 974
US-A- 5 899 975
US-A- 5 915 001

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 073 036 B1

Description

[0001] The present invention relates to a speech processing method and apparatus. The invention has particular, although not exclusive relevance to the retrieval of data including speech data, which comprises control data and phoneme data identifying phoneme symbols, from a remote server over a communications channel.

[0002] Web browsing software exists for allowing users to browse web pages which are stored on servers located throughout the Internet. There are currently two techniques which allow the browsing software to output speech relevant to the web page. In particular, either the remote server can store and send the user recorded voice data which is relevant to the web page, or the user's browser may include a text to speech synthesizer which converts the hypertext in the web page into speech, see e.g. US-A-5899975.

[0003] The problem with sending recorded speech waveform data is that speech waveform data files are relatively large and they must be transmitted over the Internet to the user. Further, each time the user selects a new web page, new speech waveform data files may need to be downloaded. Since this type of speech waveform data is much larger than the hypertext associated with the web page, this will slow down the browser's operation. This is not a problem where the user's browser converts the hypertext of the web page into speech using a text to speech synthesizer. However, with this technique, the synthesized speech can have a poor quality since it depends upon the accuracy of the text to speech converter used by the synthesizer in the web browser.

[0004] According to one aspect, the present invention provides a web browser which is operable to process web pages received from the Internet in order to determine whether or not there is an associated speech data file stored for the web page in a server on the Internet and, if there is, means for retrieving the speech data file from that server.

[0005] The invention also provides a data retrieval system which can process received text data to determine whether or not there is an associated speech data file and, if there is, means for retrieving the speech data file a remote storage location.

[0006] Exemplary embodiments of the present invention will now be described with reference to the accompanying drawings in which:

Figure 1 is a block diagram illustrating a server, a client and a computer network which allows the client to browse web pages stored on the server;

Figure 2a is a flow diagram illustrating the processing steps performed during a browsing operation using the system shown in Figure 1;

Figure 2b illustrates the data format of a web page which can be read out by the client from the server shown in Figure 1 and which includes a speech command which informs the client that speech data is stored for this web page;

Figure 2c illustrates another data format of a web page which can be read out by the client from the server shown in Figure 1 and which includes a speech command which informs the client that speech data is stored for this web page;

Figure 2d illustrates another data format of a web page which can be read out by the client from the server shown in Figure 1 and which includes a speech command which informs the client that speech data is stored for this web page;

Figure 3 illustrates the phoneme symbols used in the speech data and, next to each symbol, an example of how the phoneme is used;

Figure 4 illustrates prosody control data which is embedded within the phoneme data and which controls the pronunciation of the phoneme data using a speech synthesizer;

Figure 5 illustrates the form of speaker control data which is transmitted from the server to the client and which controls the speed and pitch of the speech generated by a synthesizer forming part of the client shown in Figure 1;

Figure 6 illustrates two examples of text and the corresponding speech data and prosody data transmitted by the server;

Figure 7 illustrates a further example of text for display and corresponding speech data transmitted by the server to the client which includes both prosody control data and speaker control data;

Figure 8 is a block diagram illustrating a client server system similar to that shown in Figure 1 but which includes a user interface which allows a user to set parameters of the speaker control data;

Figure 9 is a flow diagram illustrating the processing steps performed during a browsing operation using the system shown in Figure 8;

Figure 10 illustrates the processing steps performed by the speech synthesizer under control of user commands input via the user interface;

Figure 11a illustrates a displayed user interface which allows a user of the client to control the synthesizing operation of the speech synthesizer;

Figure 11b illustrates another displayed user interface which allows a user of the client to control the synthesizing operation of the speech synthesizer;

Figure 12 is a table illustrating the correlation between key-strokes and their corresponding function in a keyboard controlled user interface for allowing a user to control the synthesizing operation of the speech synthesizer;

Figure 13 is a block diagram illustrating a client server system similar to that shown in Figure 8 but which allows a user to input commands via the user interface to move to another part of the speech data or to other linked data;

Figure 14 is a flow diagram illustrating the processing steps performed during a browsing operation using the system shown in Figure 13;

Figure 15 is a flow diagram illustrating the processing steps performed in controlling the operation of the speech synthesizer in moving to another part of the speech data or to other linked data;

Figure 16 illustrates an example of text for display and corresponding speech data transmitted by the server to the client, which speech data includes links to other data files in addition to the prosody control data and the speaker control data;

Figure 17 is a table illustrating the correlation between key-strokes and their corresponding function in a keyboard controlled user interface used to control the browsing operation of the system shown in Figure 13;

Figure 18 is a block diagram of a client server system similar to that shown in Figure 1, in which the client does not include a speech synthesizer but which includes a module for receiving a speech synthesizer from the server; and

Figure 19 is a flow diagram illustrating the processing steps performed during a browsing operation using the system shown in Figure 18.

FIRST EMBODIMENT

[0007] Figure 1 is a block diagram showing an Internet browsing system embodying the present invention. Although the invention is described with reference to an Internet browsing system, the invention may be used in any hypertext system which includes a first and second apparatus which are connected together via a network or other communications channel and in which data from the second apparatus can be retrieved by the first apparatus and can be output as speech.

[0008] As shown in Figure 1, the system comprises a client 1, a server 3, a communications network 5 for linking the client 1 to the server 3. The system is generally called a client server system. Both the client 1 and the server 3 comprise the modules illustrated in Figure 1 and operate under control of a stored program in order to carry out the process steps shown in Figure 2. The control program for each of the client and server may be stored in a memory (not shown) at the time of manufacture or it may be read out from a storage medium detachable from the client or server or downloaded from a remote terminal connected to the client or server via a computer network.

[0009] The operation of the client server system shown in Figure 1 will now be described referring also to Figure 2. As shown in Figure 2, in step s101, an input address for a web page (which is a hypertext data file) is input by a user of the client 1. The input web address is stored in the address memory 110 and identifies the server which stores the web page and the file name of the web page that the client wishes to download. The input web address is then transmitted, in step s102, by a web address transmitting unit 101 to the server 3, where it is received, in step s151, by a web page communicating unit 102. The web page communicating unit 102 then reads out the web page from the storage unit 103 corresponding to the received web page address and transmits it, in step s152, back to the client 1. The web page receiving unit 104 in the client 1 receives, in step s103, the web page transmitted from the server 3 and a web page display unit 105 in the client 1 develops image and character data on the basis of the received web page and displays it on a display (not shown) in step s104. Then, in step s105, a speech data require unit 106 in the client 1 processes the received web page and determines whether it comprises a requirement for speech data. In particular, the speech data require unit 106 processes the received web page to determine whether or not it includes a predetermined speech command which identifies that a speech data file exists which is associated with the received web page.

[0010] Figures 2b, 2c and 2d, illustrate the form of different types of web pages which may be received in step s103 by the client 1. As shown, each of the web pages includes a header portion 11 which defines, among other things, the size and position of a window in which the web page will be displayed; a text portion 13 which includes the text to be displayed and a speech command portion 15 which identifies that there is speech data associated with the current web page. As shown in Figures 2b, 2c and 2d, the speech command portion 15 may occur before or after the text portion 13 and will comprise at least the file name 17 of the file holding the associated speech data. In this embodiment, the speech command also includes an address portion 19 which identifies where the speech data file is stored. In most cases, the address 19 will identify the same server 3 which transmitted the current web page being viewed by the user.

However this is not essential and the address portion 19 may identify a different server. Further, as illustrated in Figure 2d, the speech command may include the file names 17a and 17b of more than one speech data file and the corresponding address portions 19a and 19b. In this embodiment, the address portion 19 also identifies where the speech data file 17 is to be found within the identified server.

[0011] If at step s105, the client 1 determines that there is a requirement for speech data, then the processing proceeds to step s106, otherwise it proceeds to step s109. In step s106, a speech data require unit 106 transmits a request for the speech data. This request for speech data includes the file name 17 for the speech data file and its address 19. A speech data communicating unit 107 located within the server 3 receives, in step s153, the speech data request transmitted from the client 1 and then retrieves the appropriate speech data file from the speech data storage unit 108 and then transmits it back to the client 1 in step s154. A speech synthesizer 109 located within the client 1 receives, in step s107, the speech data file transmitted from the server 3, transforms the speech data into a synthesized speech signal in step s108 and outputs the synthesized speech signal to a loudspeaker (not shown).

[0012] The processing then proceeds to step s109 where the client 1 determines whether or not the browsing process is finished. If it has, then the processing ends, otherwise the processing returns to step s101 where the process is continued for a new input address. In this embodiment, the browser is a computer program run on a conventional computer system and which is displayed as a window on the display of the computer system. Therefore, in order to finish the browsing process, the user of the client 1 can simply close the window currently running the browser program.

[0013] In this embodiment, the speech data within the speech data file transmitted from the server 3 to the client 1 includes phoneme data (identifying phoneme symbols) which defines the sound signals to be output by the speech synthesizing unit 109; prosody control data which is used to control the stress and intonation of the sounds defined by the phoneme data; and speaker control data which controls the pitch, speed and type of speaker (male or female) of the speech signals generated by the speech synthesizing unit 109. Figure 3 illustrates the phoneme symbols used in this embodiment which are transmitted from the server 3 to the client 1. Figure 3 also gives, next to each symbol, an example of how the corresponding phoneme is used. As shown in Figure 3, the phoneme symbols are split into three groups - those relating to simple vowels 25, those relating to diphthongs 27 and those relating to consonants 29. Figure 4 illustrates the prosody control data used in this embodiment, which, in use, is embedded within the transmitted phoneme data. As shown in Figure 4, the prosody control data includes symbols which identify boundaries between syllables and words and symbols for controlling the stress applied to vowels and other symbols which control the pronunciation of the phonemes. Figure 5 illustrates the form of the speaker control data used in this embodiment. The top half 31 of Figure 5 illustrates the way in which the speaker control data is used, whilst the lower half 33 of Figure 5 illustrates the parameters that can be set. As shown, the parameters include the speed at which the speech is generated (which can be set to a value of 1 to 4); the pitch of the generated speech (which can be set to a value indicative of a shift from a predetermined standard speaker's pitch); and the type of speaker (which can either be set to male or female).

[0014] Figure 6 illustrates the phoneme data together with the embedded prosody control data generated for the example text: "New York City." and "This is a text-to-speech system.". As shown, the prosody control data used in these examples include the use of a primary stress control symbol "1" which causes the speech synthesizer 109 to place more stress on the subsequent vowel and a period control symbol "." which causes the speech synthesizer 109 to add a pause.

[0015] Figure 7 illustrates a further example which shows text data for display 39 together with corresponding speech data 41 which includes the use of both the above described prosody control data and speaker control data. As shown, the speaker control data may be inserted within the phoneme data so that, for example, the type of speaker outputting the speech may be changed during the speech synthesizing operation.

[0016] As those skilled in the art will appreciate, since the speech data file which is transmitted from the server 3 to the client 1 only includes phoneme data and appropriate control data, the speech data file will be much smaller than a corresponding recorded speech waveform file and it can therefore be retrieved from the server 3 more quickly. Further, since the data for controlling the way in which the speech synthesizer 109 synthesizes the speech is also transmitted from the server 3, it can be set in advance by the owner of the web site. Therefore, the web site owner can set the control data so that the client's speech synthesizer 109 will output the speech in a predetermined manner.

SECOND EMBODIMENT

[0017] A second embodiment of the client server system shown in Figure 1 will now be described with reference to Figures 8 to 12. Figure 8 is a block diagram illustrating the client server system of this embodiment and Figure 9 is a flow diagram illustrating the processing steps performed by this client server system during a browsing operation. As those skilled in the art will appreciate by comparing Figures 8 and 9 with Figures 1 and 2a, the main difference between this embodiment and the first embodiment is in the provision of a speech synthesizer user interface 209 within the client 1. The description of this second embodiment will therefore be

restricted to the interaction of this speech synthesizer user interface unit 209 with the other modules of the system.

[0018] In this embodiment, when the speech data require unit 106 determines that the received web page includes a speech command identifying that a speech data file exists which is associated with the current web page being viewed, it starts the speech synthesizer user interface unit 209 as well as transmitting the requirement for speech data to the server 3. Once the speech data communicating unit 107 located within the server 3 retrieves the appropriate speech data file and transmits it back to the client 1, the speech synthesizer 109 synthesizes a corresponding speech signal under control of the speech synthesizer user interface unit 209.

[0019] Figure 11a shows an example of a graphical user interface which the speech synthesizer user interface 209 outputs to the display (not shown) of the client 1, and which allows the user of the client to set various control parameters and input various commands used to control the operation of the speech synthesizer 109. In particular, the graphical user interface illustrated in Figure 11a allows the user to start, stop, pause and restart the operation of the speech synthesizer 109, to change the pitch and speed of the synthesized speech signal and to change the type of speaker etc.

[0020] Figure 10 is a flow chart illustrating the way in which commands and settings input via the speech synthesizer user interface 209, controls the operation of the speech synthesizer 109. As shown, in step s20901, the speech synthesizer 109 receives a command input by the user via the user interface 209. The speech synthesizer 109 then determines, in step s20902 if the input command is a start command. If it is, then the processing proceeds to step S20903 where the speech synthesizer 109 begins to output a synthesized speech signal from the received speech data and the processing returns to step s20901 where the speech synthesizer 109 awaits the next input command.

[0021] If at step s20902, the input command is not the start command, then the processing proceeds to step s20904 where the speech synthesizer 109 determines if the input command is a stop command. If it is, then the processing proceeds to step s20905 where the speech synthesizer 109 stops outputting the synthesized speech signal and the processing returns to step s20901 where the speech synthesizer 109 awaits the next input command.

[0022] If at step s20904, the input command is not the stop command, then the processing proceeds to step s20906 where the speech synthesizer 109 determines whether or not the input command is a pause command. If it is, then the processing proceeds to step s20907 where the speech synthesizer 109 pauses outputting the synthesized speech signal at the current location in the speech data file and then the processing returns to step s20901 where the speech synthesizer 109 awaits the next input command.

[0023] If at step s20906, the speech synthesizer 109 determines that the input command is not the pause command, then processing proceeds to step s20908 where the speech synthesizer 109 determines whether or not the input command is a restart command. If it is, then the processing proceeds to step s20909 where the speech synthesizer 109 restarts outputting the synthesized speech signal from the current location (corresponding to the location where the outputting of the synthesized speech was paused) in the speech data file and then the processing returns to step s20901 where the speech synthesizer awaits the next input command.

[0024] If at step s20908, the speech synthesizer 109 determines that the input command is not the restart command, then the processing proceeds to step s20910 where the speech synthesizer 109 determines if the input command is a command to change the pitch of the synthesized speech. If it is, then the processing proceeds to step s20911 where the pitch of the generated speech signal is changed on the basis of the pitch level set by the user. The processing then returns to step s20901 where speech synthesizer 109 awaits the next input command.

[0025] If at step s20910, the speech synthesizer 109 determines that the input command is not a change of pitch command, then the processing proceeds to step s20912 where the speech synthesizer 109 determines if the input command is for changing the speed at which the speech is being synthesized. If it is, then the processing proceeds to step s20913 where the speech synthesizer 109 changes the speed at which the speech is being synthesized on the basis of the speed level set by the user. The processing then returns to step s20901 where the speech synthesizer awaits the next input command.

[0026] If at step s20912, the speech synthesizer 109 determines that the input command is not a speed change command, then processing proceeds to step s20914 where the speech synthesizer 109 determines if the input command is a command to change the type of speaker. If it is, then the processing proceeds to step s20915 where the settings of the speech synthesizer 109 are changed in accordance with the type of speaker set by the user. The processing then returns to step s20901 where the next input command is awaited.

[0027] If at step s20914, the speech synthesizer 109 determines that the input command is not a command to change the type of speaker, then the processing proceeds to step s20916 where the speech synthesizer 109 determines whether or not the input command is to end the synthesizing operation. If it is, then the processing ends, otherwise the processing returns to step s20901 where the speech synthesizer 109 awaits the next input command.

[0028] In this embodiment, the speech synthesizer user interface 209 is a graphical user interface such as the two example interfaces shown in Figure 11. As shown, the interfaces include a control button 10 for in-

putting a command to start a synthesizing operation; a control button 11 for stopping a synthesizing operation and a control button 12 for controlling the pausing and restarting of a synthesizing operation, each of which is activated by a user controlled cursor (not shown) in a known manner. The interfaces also include a menu select button 13 for changing the type of speaker from a male speaker to a female speaker and vice versa and two sliders 14 and 15 for controlling the pitch and speed of the synthesized speech respectively. In the interface shown in Figure 11a, a progress slider 16 is also provided to show the user the progress of the speech synthesizer in generating synthesized speech signals for the received speech data file.

[0029] As those skilled in the art will appreciate as an alternative to providing a graphical user interface, the interface may be provided through a numerical or general keyboard forming part of the client, provided the relation between the keys and the commands is defined in advance. Figure 12 is a table illustrating the correlation between keys of a numeric keyboard and a general keyboard which may be programmed in advance to allow the user to control the speech synthesizer 109 using the keys of the keyboards.

[0030] As those skilled in the art will appreciate, with the client of this second embodiment, the user can change the default settings of the control parameters used to control the speech synthesizing operation set by the web site owners. In this way, the user of the client can customize the way in which the speech synthesizer synthesizes the speech in accordance with the user's preferences.

THIRD EMBODIMENT

[0031] A third embodiment of the client server system shown in Figure 1 will now be described with reference to Figures 13 to 17. Figure 13 is a block diagram illustrating the client server system of this embodiment and Figure 14 is a flow diagram illustrating the processing steps performed by this client server system during a browsing operation. As those skilled in the art will appreciate by comparing Figures 8 and 9 with Figures 13 and 14, the main difference between this embodiment and the second embodiment is the arrow coming from the speech synthesizer 109 to the address transmitting unit 101.

[0032] This allows addresses for other speech data or other hypertext data which is embedded within the received speech data file to be extracted and the corresponding data retrieved. The description of this third embodiment will therefore be restricted to the processing which is performed in this situation.

[0033] Referring to Figure 14, after the speech synthesizer 109 begins to synthesize speech in step s108, the processing proceeds to step s310 where the speech synthesizer 109 determines whether or not a hypertext address is found and output from the received speech data. If there is not, then the processing proceeds to step s109 as before. On the other hand, if an address is output, then the processing returns to step s101 where the output address is input to the address transmitting unit 101 so that the linked web page is accessed.

[0034] In this embodiment, in addition to generating a synthesized speech signal corresponding to the received speech data on the basis of commands within the speech data and/or commands input by the user via the user interface 209, the speech synthesizer 109 also responds to input commands designating a move to another part of the received speech data or to a linked web page. The way in which this is performed will now be described with reference to Figures 15 and 16. In particular, Figure 15 is a flow chart illustrating the processing steps performed by the speech synthesizer 109 in determining whether or not to move to another part of the speech data or to linked data. Figure 16 shows an example of speech data which includes link data 162. Figure 16 also shows the corresponding text which is displayed to the user and which illustrates to the user that there is a link (by virtue of the underlined text). As shown in the speech data portion of Figure 16, the link data 162 includes and address 162-1 which identifies the location of related information (which may be a related web site or it may be a further file stored in the same web site) and speech data 162-2 of a message for the user (which in this example is: "For Canon-group companies pages, push M key now").

[0035] Returning to Figure 15, the process shown starts when the user inputs in step s30901, one of the commands shown in Figure 17 whilst the speech synthesizer 109 is synthesizing speech. As shown in Figure 17, the commands include "next link", "previous link" and "go to the link". Once an input command 13 is received in step s30901, the processing proceeds to step s30902 where the speech synthesizer 109 determines whether or not the input command is the command to move to the next link. If it is, then the processing proceeds to step s30903 where the speech synthesizer 109 searches forward from its current location within the speech data file, for the next link 162 and restarts the synthesizing operation from the speech data 162-2 within that link 162. The processing then returns to step s30901 where the next input command is awaited. In this way, the user can cause the speech synthesizer to skip portions of the speech data being synthesized.

[0036] If at step s30902, the speech synthesizer 109 determines that the input command is not a command to go to the next link, then the processing proceeds to step s30904 where the speech synthesizer 109 determines whether or not the input command is a command to return to the previous link. If it is, then the processing proceeds to step s30905 where the speech synthesizer 109 searches backward from its current location for the previous link and then restarts the synthesizing operation from the speech data 612-2 within that link 162. The processing then proceeds to step s30901 where the

next input command is awaited.

[0037] If at step s30904, the speech synthesizer 109 determines that the input command is not a command to return to the previous link, then the speech synthesizer 109 determines whether or not the input command was to go to the link. If it is, then the speech synthesizer retrieves the hypertext address 162-1 from the link data and outputs the address in step s30907 to the address transmitting unit 101. The processing then ends. If, however, at step s30906 the speech synthesizer 109 determines that the input command is not to go to the link, then the processing returns to step s30901 where the next input command is awaited.

[0038] Therefore, in this way, the user can control the synthesizing of the current speech data file and can cause the client to access a linked hypertext address to retrieve further text data and/or further speech data.

[0039] In order to inform the user that the speech data 162-2 currently being synthesized corresponds to a link 162, the speech synthesizer 109 can be configured so that speech data 162-2 which forms part of a link is synthesized with a different speed or pitch or with a different type of speaker. If the speaker type is to change, then this can simply be set to be the opposite of the speaker type for speech being generated from speech data which is not part of a link.

FOURTH EMBODIMENT

[0040] A fourth embodiment of the client server system shown in Figure 1 will now be described with reference to Figures 18 and 19. Figure 18 is a block diagram illustrating the client server system of this embodiment and Figure 19 is a flow diagram illustrating the processing steps initially performed during a browsing operation. The main difference between this embodiment and the first embodiment is that the client 1 in this embodiment does not include a speech synthesizer. Therefore, prior to or at the same time that the client 1 transmits a request for the speech data, it transmits a request to the server 3 to send the client 1 a speech synthesizer. The way in which this is performed will now be described in more detail.

[0041] Referring to Figure 19, after the received hypertext has been displayed to the user in step s104, the processing proceeds to step s405 where the client 1 determines whether or not there is a speech synthesizer module in the client. If there is, then the processing shown in Figure 19 ends and the processing returns to step s105 shown in Figure 2, where the downloading of the speech data is performed in the manner described above.

[0042] If, however, at step s405 the client determines that there is no speech synthesizer, then processing proceeds to step s406 where a speech data require unit 406 transmits a request for a speech synthesizer module to the server 3. This request is received at step s453 by the speech synthesizer module transmitting unit 407. The speech synthesizer module transmitting unit 407 then retrieves an appropriate speech synthesizer module from the storage unit 408 and transmits it, in step s454 back to the client 1. The transmitted speech synthesizer is received in step s407 by a speech synthesizer receiving unit 409. The processing then proceeds to step s408 where the received speech synthesizer is initialized and set into a working condition. The processing shown in Figure 19 then ends and the processing then returns to step s105 shown in Figure 2.

[0043] In some situations, although the client may have a speech synthesizer stored in, for example, a hard disk, it may not be currently running. In this case, rather than downloading the speech synthesizer from the server 3, the client may simply load the stored speech synthesizer into the working memory and set the synthesizer into a working condition.

[0044] In the above embodiments, the web page received from the server included both text data and a speech command which identifies that a speech data file exits for the received web page. This speech command included both a file name and an address for this speech data file. In an alternative embodiment, the speech command may only include a file name for the speech data file. In this case, the client may be programmed to assume that the speech data file is stored in the server from which it downloaded the web page. Therefore, in such an embodiment, the request for the speech data file would be transmitted to the server using the address stored in the address memory 110.

[0045] In the third embodiment described above, the user interface allowed a user to move to different portions of a speech data file being synthesized and to retrieve a further web page and/or a further speech data file. In particular the user interface allows the user to input a command to move from the current point in the speech data file to a next or to a previous link. In this way, the links were used as control characters within the speech data files in addition to providing a link to another web page or another speech data file. In an alternative embodiment, the speech data file may be arranged in paragraphs or sections, with each paragraph or section having a control header which can be used to control movement of the speech synthesizing operation through the speech data file. For example, in such an embodiment, the user may input a command to move to a next paragraph or to a previous paragraph.

Claims

1. A data retrieval apparatus (1) comprising:

means (104) for receiving text data;

means (106) for processing the received text data to determine whether or not it includes data identifying a speech data file which is associated with the received text data, **character-**

ized in that the speech data file includes phoneme data and control data for use by a speech synthesizer in synthesizing a speech signal corresponding to the phoneme data under control of the control data; and
means (109) responsive to said processing means, for retrieving the speech data file.

2. An apparatus according to claim 1, further comprising means for transmitting a request for the text data to a remote terminal which stores the text data.

3. An apparatus according to claim 1 or 2, comprising storing means for storing one or more predetermined commands identifying speech data and wherein said processing means is operable to compare said stored predetermined commands with the received text data.

4. An apparatus according to any preceding claim, wherein said retrieving means comprises means for transmitting a request for said speech data file to a storage location and means for receiving said speech data file from said storage location.

5. An apparatus according to any preceding claim, further comprising a user interface for allowing a user of said apparatus to set one or more control parameters of said control data.

6. An apparatus according to any preceding claim, wherein said control data comprises prosody control data for controlling pronunciation of said phoneme data.

7. An apparatus according to any preceding claim, wherein said control data comprises speaker control data for controlling speaker related parameters of said speech.

8. An apparatus according to claim 7, wherein said speaker related parameters include the speed at which the speech is synthesized.

9. An apparatus according to claim 7 or 8, wherein said speaker related parameters include the pitch at which said speech is synthesized.

10. An apparatus according to any of claims 7 to 9, wherein said speaker related parameters include whether the speech is synthesized as a male voice or a female voice.

11. An apparatus according to any preceding claim, wherein said speech data file includes one or more links to other data and wherein said apparatus comprises means, responsive to a command input by a user, for retrieving said other data.

12. An apparatus according to claim 11, wherein said one or more links include an address portion which identifies the storage location of said other data and speech data associated with said link.

13. An apparatus according to any preceding claim, wherein said retrieving means is operable to retrieve a speech synthesizer from a remote terminal for synthesizing speech signals corresponding to said speech data.

14. An apparatus according to any of claims 1 to 12 further comprising a speech synthesizer operable to receive said phoneme data and said control data and operable to output a speech signal corresponding to the phoneme data under control of said control data.

15. An apparatus according to claim 14, further comprising a loudspeaker for generating speech sounds corresponding to the speech signal synthesized by said speech synthesizer.

16. An apparatus according to claim 14 or 15, wherein said retrieval means is operable to retrieve set up data for controlling a set up procedure for said synthesizer prior to the generation of said speech signal corresponding to the speech data.

17. An apparatus according to any of claims 14 to 16, further comprising a user interface for allowing a user to move to different portions of a speech data file to be synthesized.

18. An apparatus according to claim 17 when dependent upon claim 11, wherein said user interface allows a user to move to a next link or to a previous link within said speech data file.

19. An apparatus according to any preceding claim, further comprising output means for outputting the received text data to a user.

20. An apparatus according to claim 19, wherein said output means comprises a display.

21. An apparatus according to claim 19 or 20, wherein said processing means is operable to process said text data after said output means has output said text data to said user.

22. An apparatus according to any preceding claim, wherein said text data forms part of a hypertext data file.

23. An apparatus according to any preceding claim, wherein said text data forms part of a web page.

24. An apparatus according to any preceding claim, wherein said text data includes data identifying a storage location of said speech data file and wherein said retrieving means is operable to retrieve said speech data file from said storage location.

25. A web browser comprising a data retrieval apparatus according to any preceding claim and a display for displaying the received text data and a loudspeaker for outputting synthesized speech.

26. A data retrieval system comprising one or more computer terminals storing at least text data and wherein at least one of said terminals also stores speech data corresponding to some of the stored text data, a data retrieval apparatus according to any preceding claim for retrieving text data and speech data from said terminals and a communications network for linking said retrieval apparatus with said one or more terminals and through which retrieved text data and speech data passes.

27. A system according to claim 26, wherein said one or more computer terminals are servers, wherein said communications network forms part of the Internet and said data retrieval apparatus comprises a web browser.

28. A data retrieval method comprising the steps of:

receiving (S103) text data;
processing (S105) the received text data to determine whether or not it includes data identifying a speech data file which is associated with the received text data, **characterized in that** the speech data file includes phoneme data and control data for use by a speech synthesizer in synthesizing speech corresponding to the phoneme data under control of the control data; and
retrieving (S106; S107) the speech data file.

29. A method according to claim 28, further comprising the step of transmitting a request for the text data to a remote terminal which stores the text data.

30. A method according to claim 28 or 29, comprising the step of storing one or more predetermined commands identifying speech data and wherein said processing step compares said stored predetermined commands with the received text data.

31. A method according to any of claims 28 to 30, wherein said retrieving step comprises the steps of transmitting a request for said speech data file to a storage location and receiving said speech data file from said storage location.

32. A method according to any of claims 28 to 31, further comprising the step of receiving control parameters from a user interface

33. A method according to any of claims 28 to 32, wherein said control data comprises prosody control data for controlling pronunciation of said phoneme data.

34. A method according to any of claims 28 to 33, wherein said control data comprises speaker control data for controlling speaker related parameters of said speech.

35. A method according to claim 34, wherein said speaker related parameters include the speed at which the speech is synthesized.

36. A method according to claim 34 or 35, wherein said speaker related parameters include the pitch at which said speech is synthesized.

37. A method according to any of claims 34 to 36, wherein said speaker related parameters include whether the speech is synthesized as a male voice or a female voice.

38. A method according to any of claims 28 to 37, wherein said speech data file includes one or more links to other data and wherein said method comprises the step of retrieving said other data in response to a command input by a user.

39. A method according to claim 38, wherein said one or more links include an address portion which identifies the storage location of said other data and speech data associated with said link.

40. A method according to any of claims 28 to 39, wherein said retrieving step retrieves a speech synthesizer from a remote terminal for synthesizing speech signals corresponding to said speech data.

41. A method according to any of claims 28 to 39, further comprising the steps of synthesizing and outputting a speech signal corresponding to the phoneme data under control of said control data.

42. A method according to claim 41, further comprising the step of outputting said synthesized speech signal to a loudspeaker.

43. A method according to claim 41 or 42, wherein said retrieval step retrieves set up data for controlling a set up procedure for a speech synthesizer prior to the generation of said speech signal corresponding to the speech data.

44. A method according to any of claims 41 to 43, further comprising the step of receiving an input command from a user to move to different portions of a speech data file to be synthesized.

45. A method according to claim 44 when dependent upon claim 37, wherein said input command is a command to move to a next link or to a previous link within said speech data file.

46. A method according to any of claims 28 to 45, further comprising the step of outputting the received text data to a user.

47. A method according to claim 46, wherein said output step outputs said text data to a display.

48. A method according to claim 46 or 47, wherein said processing step is performed after said output step has output said text data to said user.

49. A method according to any of claims 28 to 48, wherein said text data forms part of a hypertext data file.

50. A method according to any of claims 28 to 49, wherein said text data forms part of a web page.

51. A method according to any of claims 28 to 50, wherein said text data includes data identifying a storage location of said speech data file and wherein said retrieving step retrieves said speech data file from said storage location.

52. A data retrieval method comprising the steps of:

at a first computer terminal:
receiving text data;
processing the received text data to determine whether or not it includes data identifying a speech data file which is associated with the received text data and which includes phoneme data and control data for use by a speech synthesizer in synthesizing a speech signal corresponding to the phoneme data under control of the control data; and
requesting a second remote computer terminal to send said speech data file to said first computer terminal; and
at said second remote computer terminal:
receiving said request for said speech data file;
retrieving said speech data file in accordance with said request; and
transmitting said retrieved speech data file to said first computer terminal.

53. A method according to claim 52 performed over the Internet.

54. A storage medium storing processor implementable instructions for controlling a processor to implement the method of any one of claims 28 to 53.

55. Processor implementable instructions for controlling a processor to implement the method of one of claims 28 to 53.

Patentansprüche

1. Datenabrufvorrichtung (1) mit
einer Einrichtung (104) zum Empfangen von Textdaten,
einer Einrichtung (106) zur Verarbeitung der empfangenen Textdaten zur Bestimmung, ob sie Daten enthalten oder nicht, die eine Sprachdatendatei identifizieren, die mit den empfangenen Textdaten verbunden ist,
dadurch gekennzeichnet, dass
die Sprachdatendatei Phonemdaten und Steuerdaten zur Verwendung durch einen Sprachsynthetisierer beim Synthetisieren eines den Phonemdaten entsprechenden Sprachsignals unter Steuerung der Steuerdaten enthält, und **gekennzeichnet durch**
eine Einrichtung (109), die auf die Verarbeitungseinrichtung anspricht, um die Sprachdatendatei abzurufen.

2. Vorrichtung nach Anspruch 1, ferner mit einer Einrichtung zum Senden einer Anforderung der Textdaten zu einem entfernten Endgerät, das die Textdaten speichert.

3. Vorrichtung nach Anspruch 1 oder 2, mit einer Speichereinrichtung zur Speicherung eines oder mehrerer vorbestimmter Befehle, die Sprachdaten identifizieren, wobei die Verarbeitungseinrichtung zum Vergleichen der gespeicherten vorbestimmten Befehle mit den empfangenen Textdaten eingerichtet ist.

4. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Abrufeinrichtung eine Einrichtung zum Senden einer Anforderung der Sprachdatendatei zu einem Speicherort und eine Einrichtung zum Empfangen der Sprachdatendatei von dem Speicherort umfasst.

5. Vorrichtung nach einem der vorhergehenden Ansprüche, ferner mit einer Benutzerschnittstelle, die einem Benutzer der Vorrichtung die Einstellung eines oder mehrerer Steuerparameter der Steuerdaten erlaubt.

6. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Steuerdaten Satzrhythmussteu-

erdaten zur Steuerung der Aussprache der Phonemdaten umfassen.

7. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Steuerdaten Sprechersteuerdaten zur Steuerung sprecherbezogener Parameter der Sprache umfassen.

8. Vorrichtung nach Anspruch 7, wobei die sprecherbezogenen Parameter die Geschwindigkeit enthalten, mit der die Sprache synthetisiert wird.

9. Vorrichtung nach Anspruch 7 oder 8, wobei die sprecherbezogenen Parameter die Tonhöhe enthalten, mit der die Sprache synthetisiert wird.

10. Vorrichtung nach einem der Ansprüche 7 bis 9, wobei die sprecherbezogenen Parameter Informationen darüber enthalten, ob die Sprache als männliche Stimme oder als weibliche Stimme synthetisiert wird.

11. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Sprachdatendatei eine oder mehrere Verknüpfungen zu anderen Daten enthält, und wobei die Vorrichtung eine Einrichtung zum Abrufen der anderen Daten umfasst, die auf einen von einem Benutzer eingegebenen Befehl anspricht.

12. Vorrichtung nach Anspruch 11, wobei die eine oder mehreren Verknüpfungen einen Adressabschnitt enthalten, der den Speicherort der anderen Daten und mit der Verknüpfung verbundene Sprachdaten identifiziert.

13. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Abrufeinrichtung zum Abrufen eines Sprachsynthetisierers von einem entfernten Endgerät zum Synthetisieren von den Sprachdaten entsprechenden Sprachsignalen eingerichtet ist.

14. Vorrichtung nach einem der Ansprüche 1 bis 12, ferner mit einem Sprachsynthetisierer zum Empfangen der Phonemdaten und der Steuerdaten und zur Ausgabe eines den Phonemdaten entsprechenden Sprachsignals unter Steuerung der Steuerdaten.

15. Vorrichtung nach Anspruch 14, ferner mit einem Lautsprecher zur Erzeugung von Sprachlauten, die dem durch den Sprachsynthetisierer synthetisierten Sprachsignal entsprechen.

16. Vorrichtung nach Anspruch 14 oder 15, wobei die Abrufeinrichtung zum Abrufen von Aufbaudaten zur Steuerung eines Aufbauvorgangs für den Synthetisierer vor der Erzeugung des den Sprachdaten entsprechenden Sprachsignals eingerichtet ist.

17. Vorrichtung nach einem der Ansprüche 14 bis 16, ferner mit einer Benutzerschnittstelle, um einem Benutzer die Bewegung zu verschiedenen Abschnitten einer zu synthetisierenden Sprachdatendatei zu ermöglichen.

18. Vorrichtung nach Anspruch 17, wenn dieser von Anspruch 11 abhängt, wobei die Benutzerschnittstelle einem Benutzer die Bewegung zu einer nächsten Verknüpfung oder einer vorhergehenden Verknüpfung in der Sprachdatendatei ermöglicht.

19. Vorrichtung nach einem der vorhergehenden Ansprüche, ferner mit einer Ausgabeeinrichtung zur Ausgabe der empfangenen Textdaten zu einem Benutzer.

20. Vorrichtung nach Anspruch 19, wobei die Ausgabeeinrichtung eine Anzeige umfasst.

21. Vorrichtung nach Anspruch 19 oder 20, wobei die Verarbeitungseinrichtung zur Verarbeitung der Textdaten eingerichtet ist, nachdem die Ausgabeeinrichtung die Textdaten zu dem Benutzer ausgegeben hat.

22. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Textdaten einen Teil einer Hypertextdatendatei bilden.

23. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Textdaten einen Teil einer Web-Seite bilden.

24. Vorrichtung nach einem der vorhergehenden Ansprüche, wobei die Textdaten Daten enthalten, die einen Speicherort der Sprachdatendatei identifizieren, und wobei die Abrufeinrichtung zum Abrufen der Sprachdatendatei von dem Speicherort eingerichtet ist.

25. Web-Browser mit einer Datenabrufvorrichtung nach einem der vorhergehenden Ansprüche und einer Anzeige zum Anzeigen der empfangenen Textdaten und einem Lautsprecher zur Ausgabe synthetisierter Sprache.

26. Datenabrufsystem mit einem oder mehreren Computerendgeräten, die zumindest Textdaten speichern, wobei zumindest eines der Endgeräte auch einigen der gespeicherten Textdaten entsprechende Sprachdaten speichert, einer Datenabrufvorrichtung nach einem der vorhergehenden Ansprüche zum Abrufen von Textdaten und Sprachdaten von den Endgeräten und einem Kommunikationsnetz zum Verknüpfen der Abrufvorrichtung mit dem einen oder den mehreren Endgeräten, wobei abgerufene Textdaten und Sprachdaten über das Kom-

munikationsnetz laufen.

27. System nach Anspruch 26, wobei das eine oder die mehreren Computerendgeräte Server sind, wobei das Kommunikationsnetz einen Teil des Internet bildet und die Datenabrufvorrichtung einen Web-Browser umfasst.

28. Datenabrufverfahren mit den Schritten

Empfangen (S103) von Textdaten,

Verarbeiten (S105) der empfangenen Textdaten zur Bestimmung, ob sie eine Sprachdatendatei identifizierende Daten enthalten oder nicht, die mit den empfangenen Textdaten verbunden ist,

dadurch gekennzeichnet, dass

die Sprachdatendatei Phonemdaten und Steuerdaten zur Verwendung durch einen Sprachsynthetisierer beim Synthetisieren von den Phonemdaten entsprechender Sprache unter Steuerung der Steuerdaten enthält, und **gekennzeichnet durch**

Abrufen (S106; S107) der Sprachdatendatei.

29. Verfahren nach Anspruch 28, ferner mit dem Schritt des Sendens einer Anforderung der Textdaten zu einem entfernten Endgerät, das die Textdaten speichert.

30. Verfahren nach Anspruch 28 oder 29, mit dem Schritt der Speicherung eines oder mehrerer vorbestimmter Befehle, die Sprachdaten identifizieren, wobei der Verarbeitungsschritt die gespeicherten vorbestimmten Befehle mit den empfangenen Textdaten vergleicht.

31. Verfahren nach einem der Ansprüche 28 bis 30, wobei der Abrufschritt die Schritte des Sendens einer Anforderung der Sprachdatendatei zu einem Speicherort und Empfangen der Sprachdatendatei von dem Speicherort umfasst.

32. Verfahren nach einem der Ansprüche 28 bis 31, ferner mit dem Schritt des Empfangens von Steuerparametern von einer Benutzerschnittstelle.

33. Verfahren nach einem der Ansprüche 28 bis 32, wobei die Steuerdaten Satzrhythmussteuerdaten zur Steuerung der Aussprache der Phonemdaten umfassen.

34. Verfahren nach einem der Ansprüche 28 bis 33, wobei die Steuerdaten Sprechersteuerdaten zur Steuerung sprecherbezogener Parameter der Sprache umfassen.

35. Verfahren nach Anspruch 34, wobei die sprecherbezogenen Parameter die Geschwindigkeit enthalten, mit der die Sprache synthetisiert wird.

36. Verfahren nach Anspruch 34 oder 35, wobei die sprecherbezogenen Parameter die Tonhöhe enthalten, mit der die Sprache synthetisiert wird.

37. Verfahren nach einem der Ansprüche 34 bis 36, wobei die sprecherbezogenen Parameter Informationen darüber enthalten, ob die Sprache als männliche Stimme oder als weibliche Stimme synthetisiert wird.

38. Verfahren nach einem der Ansprüche 28 bis 37, wobei die Sprachdatendatei eine oder mehrere Verknüpfungen zu anderen Daten enthält, und wobei das Verfahren den Schritt des Abrufens der anderen Daten im Ansprechen auf einen von einem Benutzer eingegebenen Befehl umfasst.

39. Verfahren nach Anspruch 38, wobei die eine oder die mehreren Verknüpfungen einen Adressenabschnitt enthalten, der den Speicherort der anderen Daten und mit der Verknüpfung verbundene Sprachdaten identifiziert.

40. Verfahren nach einem der Ansprüche 28 bis 39, wobei der Abrufschritt einen Sprachsynthetisierer von einem entfernten Endgerät zum Synthetisieren von den Sprachdaten entsprechenden Sprachsignalen abruft.

41. Verfahren nach einem der Ansprüche 28 bis 39, ferner mit den Schritten des Synthetisierens und Ausgebens eines den Phonemdaten entsprechenden Sprachsignals unter Steuerung der Steuerdaten.

42. Verfahren nach Anspruch 41, ferner mit dem Schritt des Ausgebens des synthetisierten Sprachsignals zu einem Lautsprecher.

43. Verfahren nach Anspruch 41 oder 42, wobei der Abrufschritt Aufbaudaten zur Steuerung eines Aufbauvorgangs für einen Sprachsynthetisierer vor der Erzeugung des den Sprachdaten entsprechenden Sprachsignals abruft.

44. Verfahren nach einem der Ansprüche 41 bis 43, ferner mit dem Schritt des Empfangs eines eingegebenen Befehls von einem Benutzer zum Bewegen zu verschiedenen Abschnitten einer zu synthetisierenden Sprachdatendatei.

45. Verfahren nach Anspruch 44, wenn dieser von Anspruch 37 abhängt, wobei der eingegebene Befehl ein Befehl zum Bewegen zu einer nächsten Verknüpfung oder zu einer vorhergehenden Verknüpfung in der Sprachdatendatei ist.

46. Verfahren nach einem der Ansprüche 28 bis 45, ferner mit dem Schritt des Ausgebens der empfange-

nen Textdaten zu einem Benutzer.

47. Verfahren nach Anspruch 46, wobei der Ausgabeschritt die Textdaten zu einer Anzeige ausgibt.

48. Verfahren nach Anspruch 46 oder 47, wobei der Verarbeitungsschritt durchgeführt wird, nachdem der Ausgabeschritt die Textdaten zu dem Benutzer ausgegeben hat.

49. Verfahren nach einem der Ansprüche 28 bis 48, wobei die Textdaten einen Teil einer Hypertextdatendatei bilden.

50. Verfahren nach einem der Ansprüche 28 bis 49, wobei die Textdaten einen Teil einer Web-Seite bilden.

51. Verfahren nach einem der Ansprüche 28 bis 50, wobei die Textdaten Daten enthalten, die einen Speicherort der Sprachdatendatei identifizieren, und wobei der Abrufschritt die Sprachdatendatei von dem Speicherort abruft.

52. Datenabrufverfahren mit den Schritten
an einem ersten Computerendgerät:

Empfangen von Textdaten,
Verarbeiten der empfangenen Textdaten zur Bestimmung, ob sie Daten enthalten oder nicht, die eine Sprachdatendatei identifizieren, die mit den empfangenen Textdaten verbunden ist und Phonemdaten und Steuerdaten zur Verwendung durch einen Sprachsynthetisierer beim Synthetisieren eines den Phonemdaten entsprechenden Sprachsignals unter Steuerung der Steuerdaten enthält, und
Auffordern eines zweiten entfernten Computerendgeräts zum Senden der Sprachdatendatei zu dem ersten Computerendgerät, und
an dem zweiten entfernten Computerendgerät:

Empfangen der Anforderung der Sprachdatendatei,
Abrufen der Sprachdatendatei entsprechend der Anforderung und
Senden der abgerufenen Sprachdatendatei zu dem ersten Computerendgerät.

53. Verfahren nach Anspruch 52, wobei es über das Internet durchgeführt wird.

54. Speichermedium, das Prozessor-implementierbare Anweisungen zur Steuerung eines Prozessors zum Implementieren des Verfahrens nach einem der Ansprüche 28 bis 53 speichert.

55. Prozessor-implementierbare Anweisungen zur Steuerung eines Prozessor zum Implementieren des Verfahrens nach einem der Ansprüche 28 bis 53.

Revendications

1. Appareil (1) de récupération de données comportant :

un moyen (104) destiné à recevoir des données de texte ;
un moyen (106) destiné à traiter les données de texte reçues pour déterminer si elles comprennent ou non des données identifiant un fichier de données de parole qui est associé aux données de texte reçues, **caractérisé en ce que** le fichier de données de parole comprend des données de phonèmes et des données de commande pour une utilisation d'un synthétiseur vocal dans la synthèse d'un signal vocal correspondant aux données de phonèmes sous la commande des données de commande ; et
un moyen (109) qui, en réponse audit moyen de traitement, est destiné à récupérer le fichier de données vocales.

2. Appareil selon la revendication 1, comportant en outre un moyen destiné à transmettre une demande pour les données de texte à un terminal éloigné qui stocke les données de texte.

3. Appareil selon la revendication 1 ou 2, comportant un moyen de stockage destiné à stocker un ou plusieurs ordres prédéterminés identifiant des données vocales, et dans lequel ledit moyen de traitement peut être mis en oeuvre pour comparer lesdits ordres prédéterminés stockés aux données de texte reçues.

4. Appareil selon l'une quelconque des revendications précédentes, dans lequel ledit moyen de récupération comprend un moyen destiné à transmettre une demande pour ledit fichier de données vocales à un emplacement de stockage, et un moyen destiné à recevoir ledit fichier de données vocales depuis ledit emplacement de stockage.

5. Appareil selon l'une quelconque des revendications précédentes, comportant en outre une interface d'utilisateur pour permettre à un utilisateur dudit appareil d'établir un ou plusieurs paramètres de commande desdites données de commande.

6. Appareil selon l'une quelconque des revendications précédentes, dans lequel lesdites données de commande comprennent des données de commande de prosodies pour commander la prononciation

desdites données de phonèmes.

7. Appareil selon l'une quelconque des revendications précédentes, dans lequel lesdites données de commande comprennent les données de commande de parleur destinées à commander des paramètres de ladite parole associés à un parleur.

8. Appareil selon la revendication 7, dans lequel lesdits paramètres associés à un haut-parleur comprennent la vitesse à laquelle la parole est synthétisée.

9. Appareil selon la revendication 7 ou 8, dans lequel lesdits paramètres associés à un parleur comprennent la hauteur à laquelle ladite parole est synthétisée.

10. Appareil selon l'une quelconque des revendications 7 à 9, dans lequel lesdits paramètres associés à un parleur comprennent le fait de déterminer si la parole est synthétisée sous la forme d'une voix masculine ou d'une voix féminine.

11. Appareil selon l'une quelconque des revendications précédentes, dans lequel ledit fichier de données vocales comprend une ou plusieurs liaisons avec d'autres données, et dans lequel ledit appareil comporte un moyen qui, en réponse à un ordre introduit par un utilisateur, est destiné à récupérer lesdites autres données.

12. Appareil selon la revendication 11, dans lequel ladite une ou plusieurs liaisons comprennent une partie d'adresse qui identifie l'emplacement de stockage desdites autres données et des données vocales associées à ladite liaison.

13. Appareil selon l'une quelconque des revendications précédentes, dans lequel ledit moyen de récupération peut être mis en oeuvre pour récupérer un synthétiseur vocal à partir d'un terminal éloigné pour synthétiser des signaux vocaux correspondant auxdites données vocales.

14. Appareil selon l'une quelconque des revendications 1 à 12, comportant en outre un synthétiseur vocal pouvant être mis en oeuvre pour recevoir lesdites données de phonèmes et lesdites données de commande, et pouvant être mis en oeuvre pour délivrer en sortie un signal vocal correspondant aux données de phonèmes sous la commande desdites données de commande.

15. Appareil selon la revendication 14, comportant en outre un haut-parleur destiné à générer des sons vocaux correspondant au signal vocal synthétisé par ledit synthétiseur vocal.

16. Appareil selon la revendication 14 ou 15, dans lequel ledit moyen de récupération peut être mis en oeuvre pour récupérer des données d'établissement destinées à commander une procédure d'établissement pour ledit synthétiseur avant la génération dudit signal vocal correspondant auxdites données vocales.

17. Appareil selon l'une quelconque des revendications 14 à 16, comportant en outre une interface d'utilisateur pour permettre à un utilisateur de se déplacer vers différentes parties d'un fichier de données vocales devant être synthétisées.

18. Appareil selon la revendication 17 lorsqu'elle dépend de la revendication 11, dans lequel ladite interface d'utilisateur permet à un utilisateur de se déplacer jusqu'à une liaison suivante ou jusqu'à une liaison précédente dans ledit fichier de données vocales.

19. Appareil selon l'une quelconque des revendications précédentes, comportant en outre un moyen de sortie destiné à délivrer en sortie à un utilisateur les données de texte reçues.

20. Appareil selon la revendication 19, dans lequel ledit moyen de sortie comprend un visuel.

21. Appareil selon la revendication 19 ou 20, dans lequel ledit moyen de traitement peut être mis en oeuvre pour traiter lesdites données de texte après que ledit moyen de sortie a délivré en sortie audit utilisateur les données de texte.

22. Appareil selon l'une quelconque des revendications précédentes, dans lequel lesdites données de texte font partie d'un fichier de données d'hypertexte.

23. Appareil selon l'une quelconque des revendications précédentes, dans lequel lesdites données de texte font partie d'une page web.

24. Appareil selon l'une quelconque des revendications précédentes, dans lequel lesdites données de texte comprennent des données identifiant un emplacement de stockage dudit fichier de données vocales, et dans lequel ledit moyen de récupération peut être mis en oeuvre pour récupérer ledit fichier de données vocales à partir dudit emplacement de stockage.

25. Navigateur comportant un appareil de récupération de données selon l'une quelconque des revendications précédentes et un visuel destiné à afficher les données de texte reçues, et un haut-parleur destiné à délivrer en sortie une parole synthétisée.

26. Système de récupération de données comportant un ou plusieurs terminaux informatiques stockant au moins des données de texte et dans lequel au moins l'un desdits terminaux stocke aussi des données vocales correspondant à certaines des données de texte stockées, un appareil de récupération de données selon l'une quelconque des revendications précédentes pour récupérer des données de texte et des données vocales à partir desdits terminaux et un réseau de communications pour relier ledit appareil de récupération audit un ou auxdits plusieurs terminaux et par l'intermédiaire duquel les données de texte et les données vocales récupérées passent.

27. Système selon la revendication 26, dans lequel ledit un ou lesdits plusieurs terminaux informatiques sont des serveurs, dans lequel ledit réseau de communications fait partie de l'Internet et ledit appareil de récupération de données comporte un navigateur.

28. Procédé de récupération de données comprenant les étapes qui consistent :

à recevoir (S103) des données de texte ;
à traiter (S105) les données de texte reçues pour déterminer si elles comprennent ou non des données identifiant un fichier de données vocales qui est associé aux données de texte reçues,

caractérisé en ce que le fichier de données vocales comprend des données de phonèmes et des données de commande destinées à être utilisées par un synthétiseur vocal dans la synthétisation de paroles correspondant aux données de phonème sous la commande des données de commande ; et

à récupérer (S106 ; S107) le fichier de données vocales.

29. Procédé selon la revendication 28, comprenant en outre l'étape de transmission d'une demande pour les données de texte à un terminal éloigné qui stocke les données de texte.

30. Procédé selon la revendication 28 ou 29, comprenant l'étape de stockage d'un ou plusieurs ordres prédéterminés identifiant des données vocales et dans lequel ladite étape de traitement compare lesdits ordres prédéterminés stockés aux données de texte reçues.

31. Procédé selon l'une quelconque des revendications 28 à 30, dans lequel ladite étape de récupération comprend les étapes qui consistent à transmettre une demande pour ledit fichier de données vocales à un emplacement de stockage et à recevoir ledit fichier de données vocales depuis ledit emplacement de stockage.

32. Procédé selon l'une quelconque des revendications 28 à 31, comprenant en outre l'étape de réception de paramètres de commande depuis une interface d'utilisateur.

33. Procédé selon l'une quelconque des revendications 28 à 32, dans lequel lesdites données de commande comprennent des données de commande de prosodies destinées à commander la prononciation desdites données de phonèmes.

34. Procédé selon l'une quelconque des revendications 28 à 33, dans lequel lesdites données de commande comprennent des données de commande de parleur destinées à commander des paramètres associés à un parleur de ladite parole.

35. Procédé selon la revendication 34, dans lequel lesdits paramètres associés à un parleur comprennent la vitesse à laquelle la parole est synthétisée.

36. Procédé selon la revendication 34 ou 35, dans lequel lesdits paramètres associés à un parleur comprennent la hauteur à laquelle ladite parole est synthétisée.

37. Procédé selon l'une quelconque des revendications 34 à 36, dans lequel lesdits paramètres associés à un parleur comprennent le fait de déterminer si la parole est synthétisée sous la forme d'une voix masculine ou d'une voix féminine.

38. Procédé selon l'une quelconque des revendications 28 à 37, dans lequel ledit fichier de données vocales comprend une ou plusieurs liaisons vers d'autres données et dans lequel ledit procédé comprend l'étape de récupération desdites autres données en réponse à un ordre appliqué en entrée par un utilisateur.

39. Procédé selon la revendication 38, dans lequel ladite une ou lesdites plusieurs liaisons comprennent une partie d'adresse qui identifie l'emplacement de stockage desdites autres données et desdites données vocales associées à ladite liaison.

40. Procédé selon l'une quelconque des revendications 28 à 39, dans lequel ladite étape de récupération récupère un synthétiseur vocal à partir d'un terminal éloigné pour synthétiser des signaux vocaux correspondant auxdites données vocales.

41. Procédé selon l'une quelconque des revendications 28 à 39, comprenant en outre les étapes consistant

à synthétiser et délivrer en sortie un signal vocal correspondant aux données de phonèmes sous la commande desdites données de commande.

42. Procédé selon la revendication 41, comprenant en outre l'étape consistant à délivrer en sortie à un haut-parleur ledit signal vocal synthétisé.

43. Procédé selon la revendication 41 ou 42, dans lequel ladite étape de récupération récupère des données d'établissement pour commander une procédure d'établissement pour un synthétiseur vocal avant la génération dudit signal vocal correspondant aux données vocales.

44. Procédé selon l'une quelconque des revendications 41 à 43, comprenant en outre l'étape de réception d'un ordre d'entrée provenant d'un utilisateur pour se déplacer vers différentes parties d'un fichier de données vocales devant être synthétisées.

45. Procédé selon la revendication 44 lorsqu'elle dépend de la revendication 37, dans lequel ledit ordre d'entrée est un ordre pour passer à une liaison suivante ou à une liaison précédente dans ledit fichier de données vocales.

46. Procédé selon l'une quelconque des revendications 28 à 45, comprenant en outre l'étape consistant à délivrer en sortie à un utilisateur les données de texte reçues.

47. Procédé selon la revendication 46, dans lequel ladite étape de sortie délivre en sortie à un visuel lesdites données de texte.

48. Procédé selon la revendication 46 ou 47, dans lequel ladite étape de traitement est exécutée après que ladite étape de sortie a délivré en sortie lesdites données de texte audit utilisateur.

49. Procédé selon l'une quelconque des revendications 28 à 48, dans lequel lesdites données de texte font partie d'un fichier de données hypertexte.

50. Procédé selon l'une quelconque des revendications 28 à 49, dans lequel lesdites données de texte font partie d'une page web.

51. Procédé selon l'une quelconque des revendications 28 à 50, dans lequel lesdites données de texte comprennent des données identifiant un emplacement de stockage dudit fichier de données vocales et dans lequel ladite étape de récupération récupère ledit fichier de données vocales à partir dudit emplacement de stockage.

52. Procédé de récupération de données comprenant les étapes qui consistent :

à un premier terminal informatique :

à recevoir des données de texte ;
à traiter les données de texte reçues pour déterminer si elles comprennent ou non des données identifiant un fichier de données vocales qui est associé aux données de texte reçues et qui comprend des données de phonèmes et des données de commande destinées à être utilisées par un synthétiseur vocal dans la synthèse d'un signal vocal correspondant aux données de phonèmes sous la commande des données de commande ; et
à demander à un second terminal informatique éloigné d'envoyer ledit fichier de données vocales audit premier terminal informatique ; et
audit second terminal informatique éloigné :

à recevoir ladite demande pour ledit fichier de données vocales ;
à récupérer ledit fichier de données vocales conformément à ladite demande ; et
à transmettre ledit fichier de données vocales récupéré audit premier terminal informatique.

53. Procédé selon la revendication 52, exécuté sur l'Internet.

54. Support de stockage stockant des instructions exécutables par un processeur pour commander un processeur afin qu'il exécute le procédé selon l'une quelconque des revendications 28 à 53.

55. Instruction exécutable par un processeur pour commander un processeur afin qu'il exécute le procédé selon l'une quelconque des revendications 28 à 53.

FIG.1

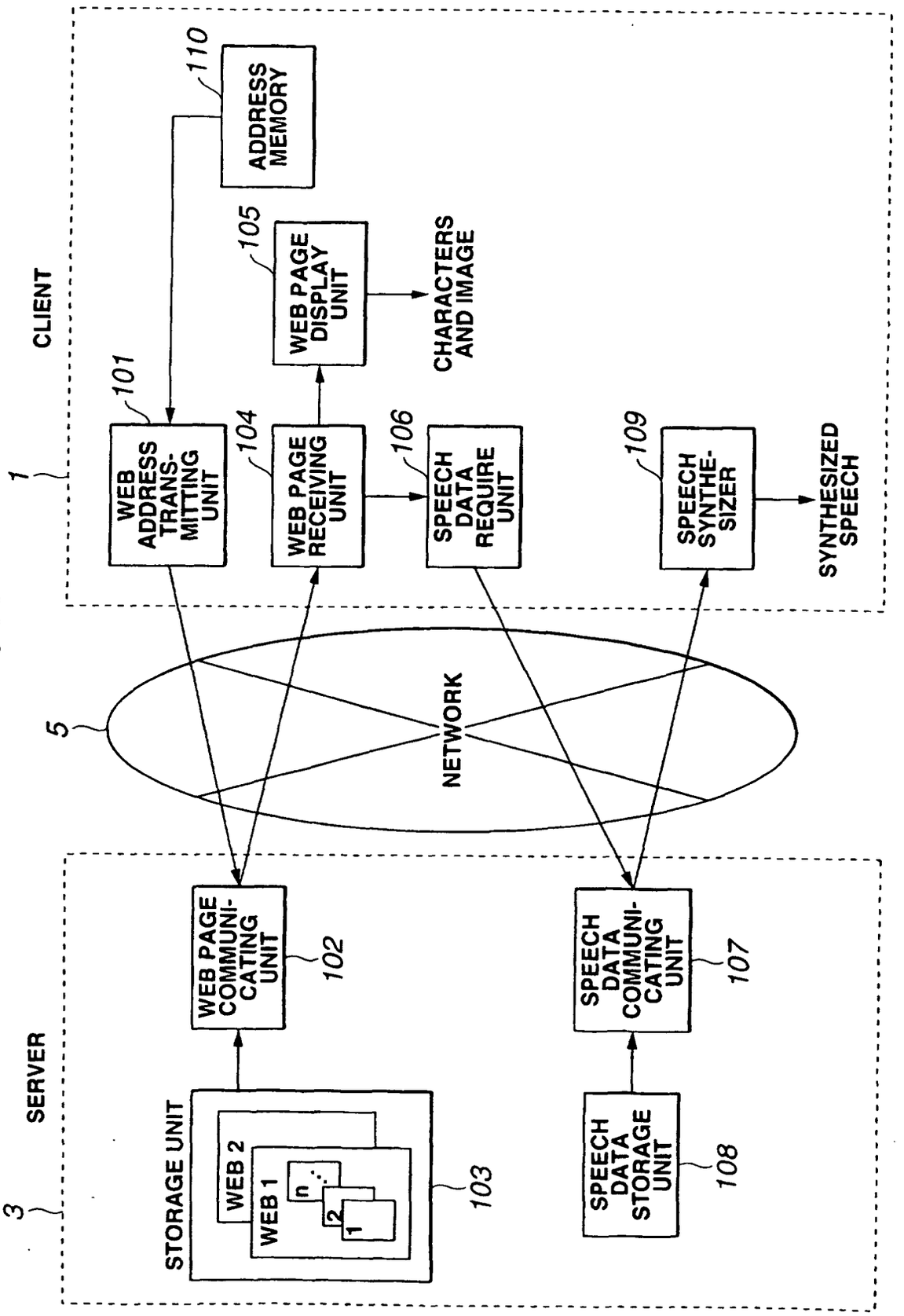

CLIENT
1
ADDRESS MEMORY
110
WEB PAGE DISPLAY UNIT
105
CHARACTERS AND IMAGE
WEB ADDRESS TRANS-MITTING UNIT
101
WEB PAGE RECEIVING UNIT
104
SPEECH DATA REQUIRE UNIT
106
SPEECH SYNTHE-SIZER
109
SYNTHESIZED SPEECH
5
NETWORK
SERVER
3
WEB PAGE COMMUNI-CATING UNIT
102
SPEECH DATA COMMUNI-CATING UNIT
107
STORAGE UNIT
WEB 2
WEB 1
n
2
1
103
SPEECH DATA STORAGE UNIT
108

FIG.2a

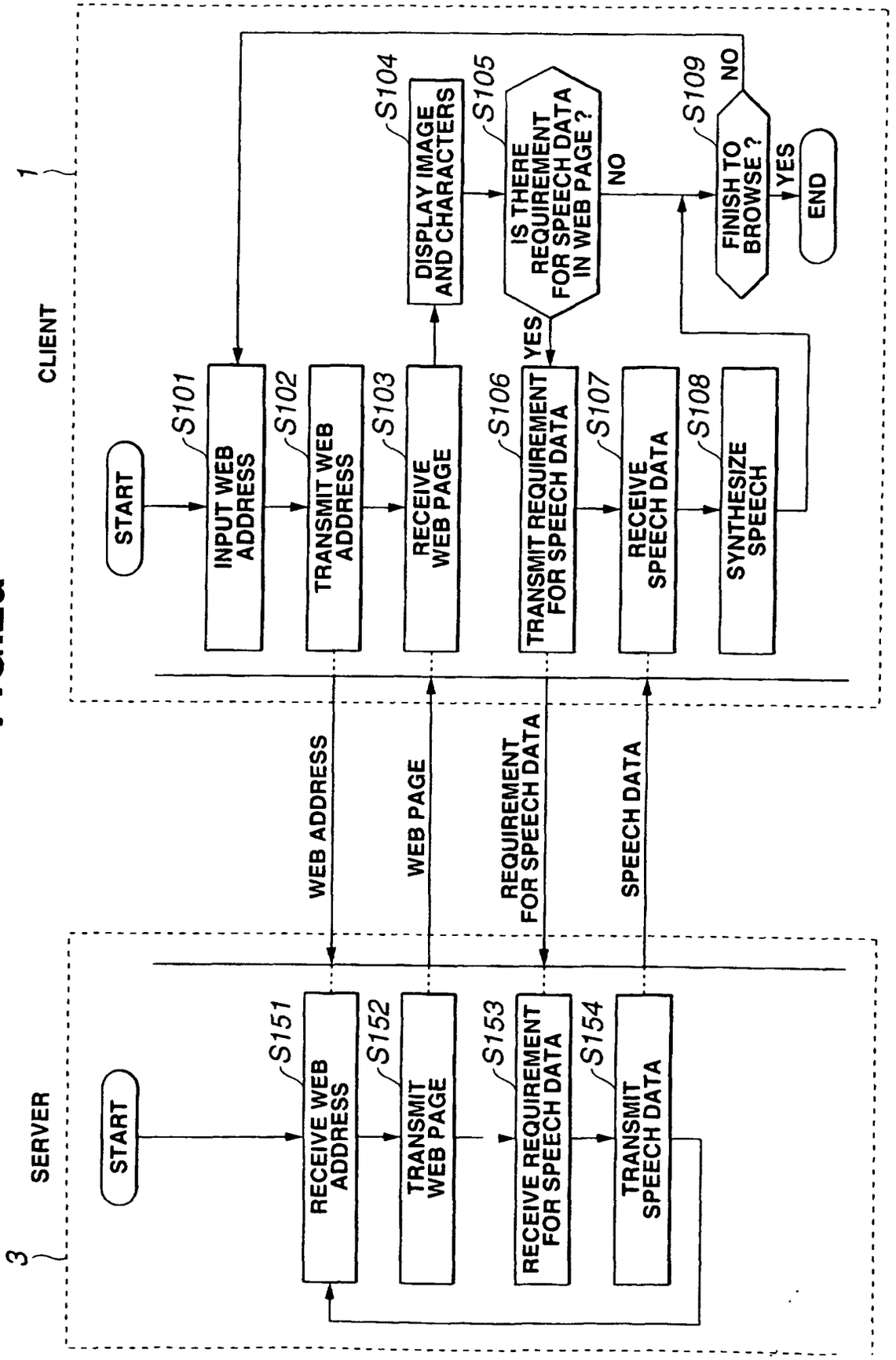

CLIENT
1
START
S101
INPUT WEB ADDRESS
S102
TRANSMIT WEB ADDRESS
S103
RECEIVE WEB PAGE
S104
DISPLAY IMAGE AND CHARACTERS
S105
IS THERE REQUIREMENT FOR SPEECH DATA IN WEB PAGE ?
YES
S106
TRANSMIT REQUIREMENT FOR SPEECH DATA
S107
RECEIVE SPEECH DATA
S108
SYNTHESIZE SPEECH
NO
S109
FINISH TO BROWSE ?
YES
END
NO
WEB ADDRESS
WEB PAGE
REQUIREMENT FOR SPEECH DATA
SPEECH DATA
SERVER
3
START
S151
RECEIVE WEB ADDRESS
S152
TRANSMIT WEB PAGE
S153
RECEIVE REQUIREMENT FOR SPEECH DATA
S154
TRANSMIT SPEECH DATA

FIG.2b

<HEADER> 11
<Position><Size of Window>
<Font Size>
<Text for Display> 13
<Speech Command> 15
<File Name><Address>
17 19

FIG.2c

<HEADER> 11
<Position><Size of Window>
<Speech Command> 15
17 <File Name><Address>
<Font Size> 19
<Text for Display> 13

FIG.2d

<HEADER> 11
<Position><Size of Window>
<Font Size>
<Text for Display> 13
<Speech Command> 15
17a <File Name 1><Address 1> 19a
17b <File Name 2><Address 2> 19b
<Font Size>
<Text for Display> 13

FIG.3

Machine Readable Phonemic Alphabet

Simple Vowels

25

Symbol	Use	Symbol	Use
ih	bid	er	bird
iy	bead	ax	'schwa'
eh	bed	ao	pot
ae	bad	oo	port
aa	bard	uh	put
ah	bud	uw	boot

Diphthongs

27

Symbol	Use	Symbol	Use
ey	day	oy	boy
ow	go	iax	beer
aw	cow	eax	bare
uax	tour	ay	eye

Consonants

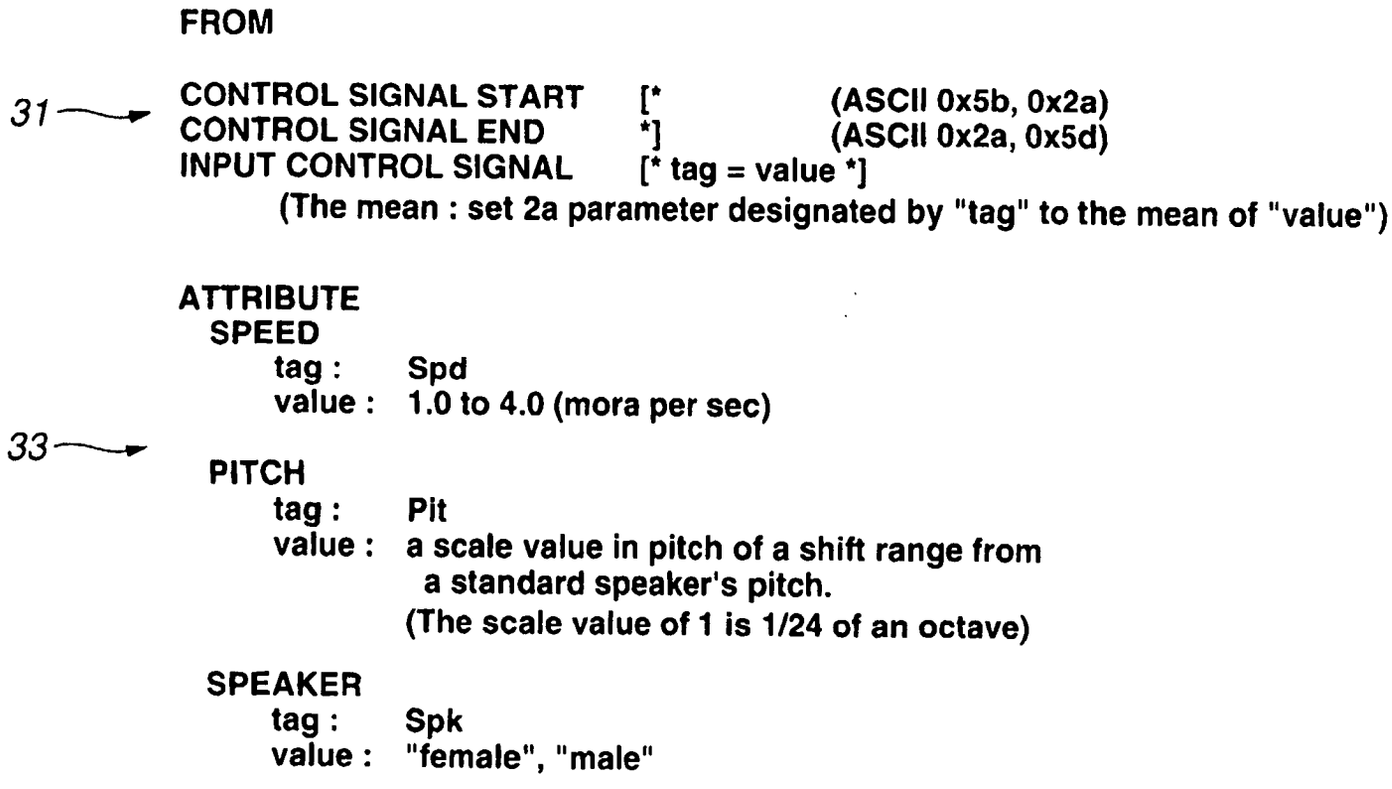
29

Symbol	Use	Symbol	Use
p	pea	v	van
t	tea	s	sea
k	key	z	buzz
b	bee	sh	she
d	dye	zh	beige
g	guy	ch	each
m	me	jh	edge
n	knee	h	hat
ng	sing	y	yes
th	thin	w	way
dh	then	r	ray
f	fan	l	lay

FIG.4

Description		Unicode name	Unicode	Usage/Effect
syllable boundary	-	HYPHEN-MINUS	U+002D	separates syllables
word boundary	#	NUMBER SIGN	U+0023	separates words
word boundary	(space)	SPACE	U+0020	separates words
silence	-	UNDERLINE	U+005f	indicates silent period
primary stress	1	MODIFIER LETTER VERTICAL LINE	U+02C8	precedes affected vowel
secondary stress	2	MODIFIER LETTER LOW VERTICAL LINE	U+02CC	precedes affected vowel
word boundary	(blank)	SPACE	U+0020	separates words
period	.	FULL STOP	U+002E	pitch fall, pause
question mark	?	QUESTION MARK	U+003F	pitch rise, pause
exclamation	!	EXCLAMATION MARK	U+0021	raised pitch range, pause
comma	,	COMMA	U+002C	continuation rise, pause

FIG.5



EP 1 073 036 B1

FIG.6

New York City.
n uw 1 y ao r 1 k s ih 1 d iy .

This is a Text-to-Speech system.
dh ih s ih z ax t eh 1 k s t t uh s p iy 1 ch s ih 1 s t ax m .

FIG.7

(Text For Display)
Social and cultural support program. 39
Under the Canon's philosophy of "Kyosei" living and working together for the common good of mankind, diversified activities are promoted by Canon-group companies all over the world in the field of social contribution and cultural support.

(Speech Data) 41
[*Spd=7*][*Pit=10*][*Spk=female*]s ow 1 sh ax l ax n k ah 1 l ch er r ax l s ax p ao r 1 t p r ow 1 g r ae m .
[*Spd=10*][*Pit=10*][*Spk=male*]ah 1 n d er d h ax k ae 1 n ax n z f ih l aa 1 s ih f iy ah v k ay 1 y ih z iy l ih 1 v ih nx ax n w er 1 k ih nx t ih g eh 1 dh er f ao r dh ax k aa 1 m ax n g uh 1 d ax v m ae 1 n k ay n d d ih v er 1 s ih f ay d ae k t ih 1 v ih d iy z aa r p r ax m ow 1 d ih d b ay k ae 1 n ax n g r uw 1 p k ah 1 m p ih n iy z ao 1 l ow 1 v er dh ax w er 1 l d ih n dh ax f iy 1 l d ax v s ow 1 sh ax l k aa 2 n ch r ih b y uw 1 sh ax n ax n k ah 1 l ch er r ax l s ax p ao r 1 t .

FIG.8

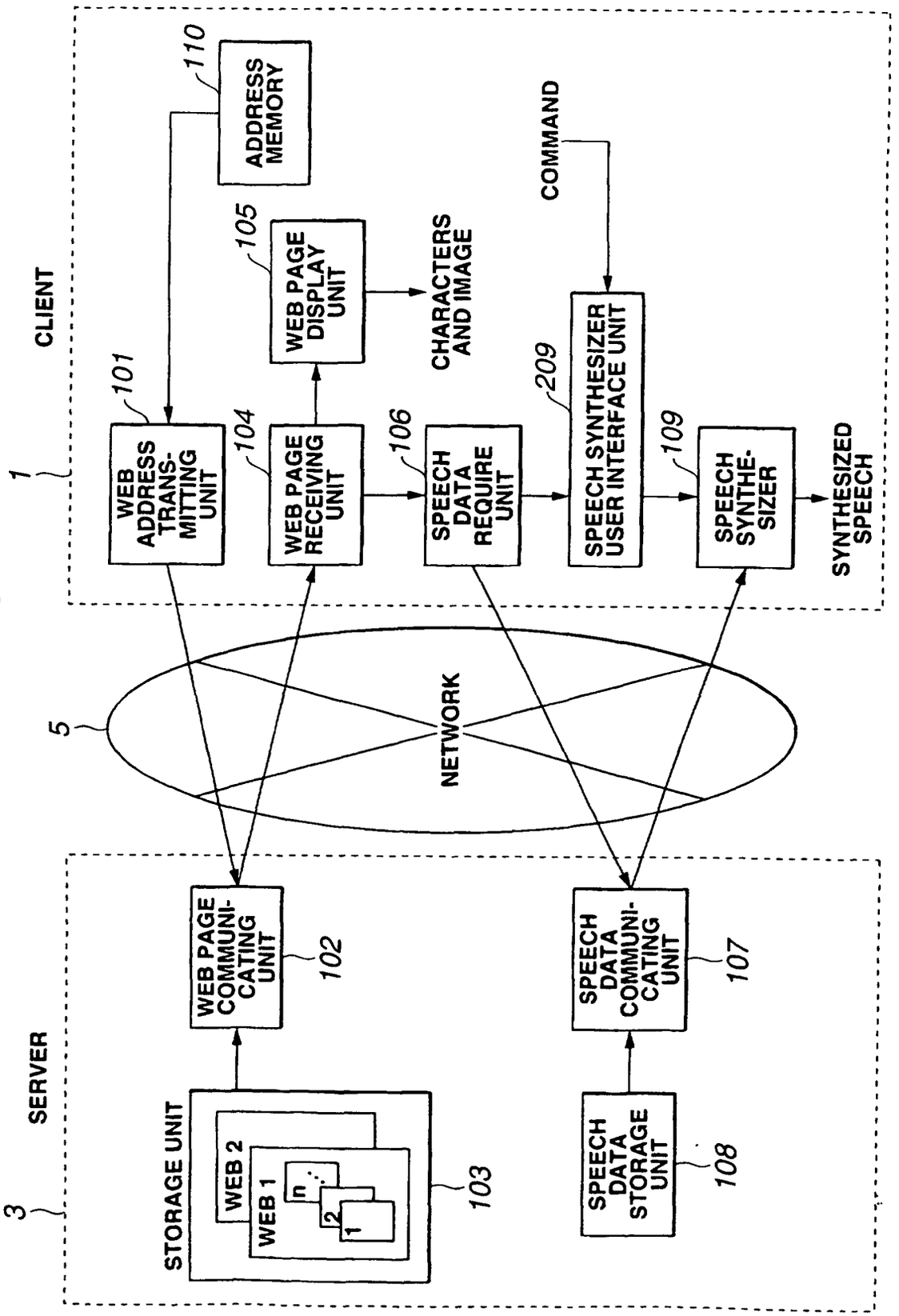

CLIENT
1
ADDRESS MEMORY
110
COMMAND
WEB PAGE DISPLAY UNIT
105
CHARACTERS AND IMAGE
WEB ADDRESS TRANS-MITTING UNIT
101
WEB PAGE RECEIVING UNIT
104
SPEECH DATA REQUIRE UNIT
106
SPEECH SYNTHESIZER USER INTERFACE UNIT
209
SPEECH SYNTHE-SIZER
109
SYNTHESIZED SPEECH
NETWORK
5
SERVER
3
WEB PAGE COMMUNI-CATING UNIT
102
SPEECH DATA COMMUNI-CATING UNIT
107
STORAGE UNIT
WEB 2
WEB 1
n
2
1
103
SPEECH DATA STORAGE UNIT
108

FIG.9

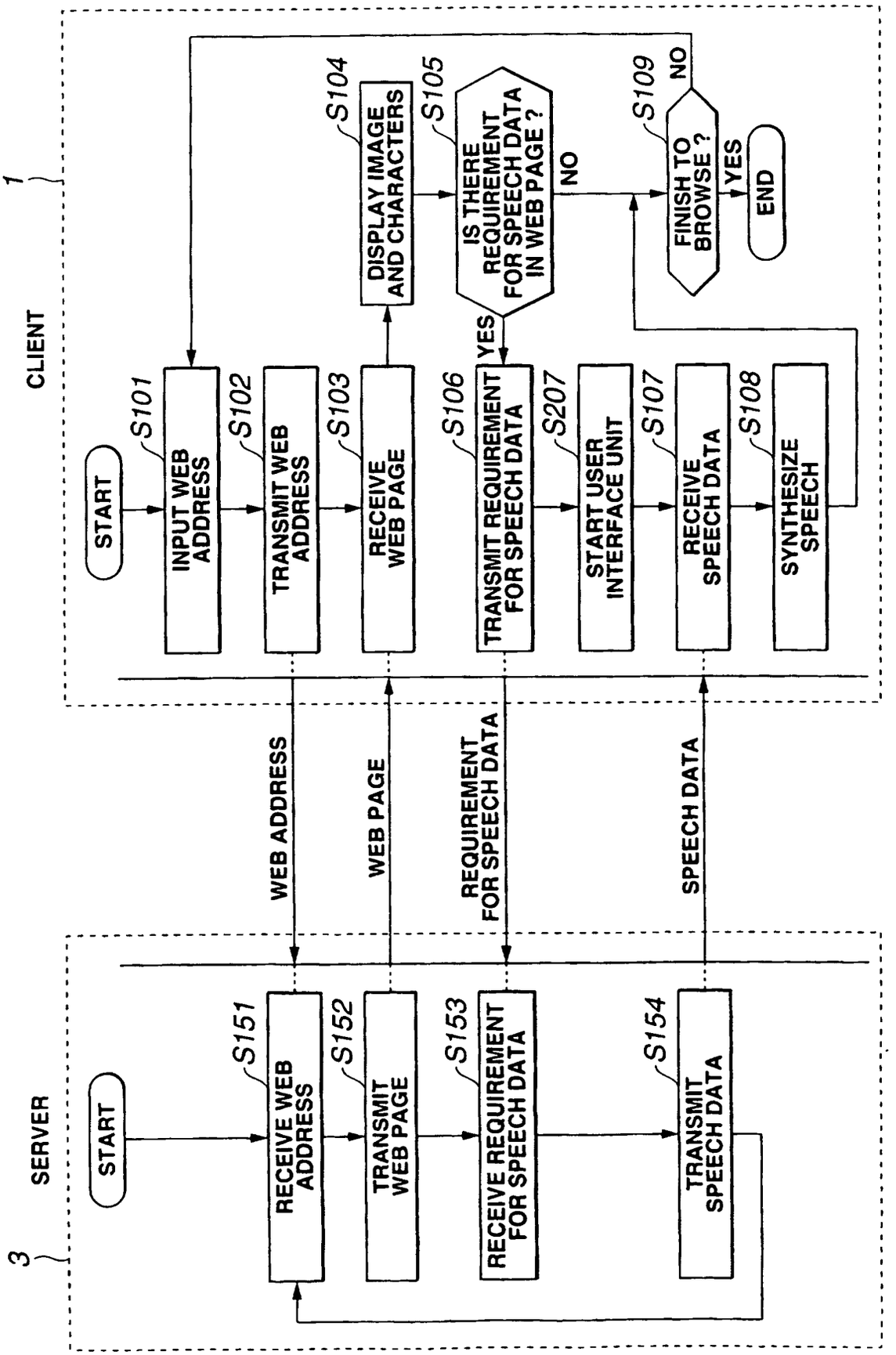

CLIENT
1
START
S101
INPUT WEB ADDRESS
S102
TRANSMIT WEB ADDRESS
S103
RECEIVE WEB PAGE
S104
DISPLAY IMAGE AND CHARACTERS
S105
IS THERE REQUIREMENT FOR SPEECH DATA IN WEB PAGE ?
NO
YES
S106
TRANSMIT REQUIREMENT FOR SPEECH DATA
S207
START USER INTERFACE UNIT
S107
RECEIVE SPEECH DATA
S108
SYNTHESIZE SPEECH
S109
FINISH TO BROWSE ?
NO
YES
END
WEB ADDRESS
WEB PAGE
REQUIREMENT FOR SPEECH DATA
SPEECH DATA
SERVER
3
START
S151
RECEIVE WEB ADDRESS
S152
TRANSMIT WEB PAGE
S153
RECEIVE REQUIREMENT FOR SPEECH DATA
S154
TRANSMIT SPEECH DATA

FIG.10

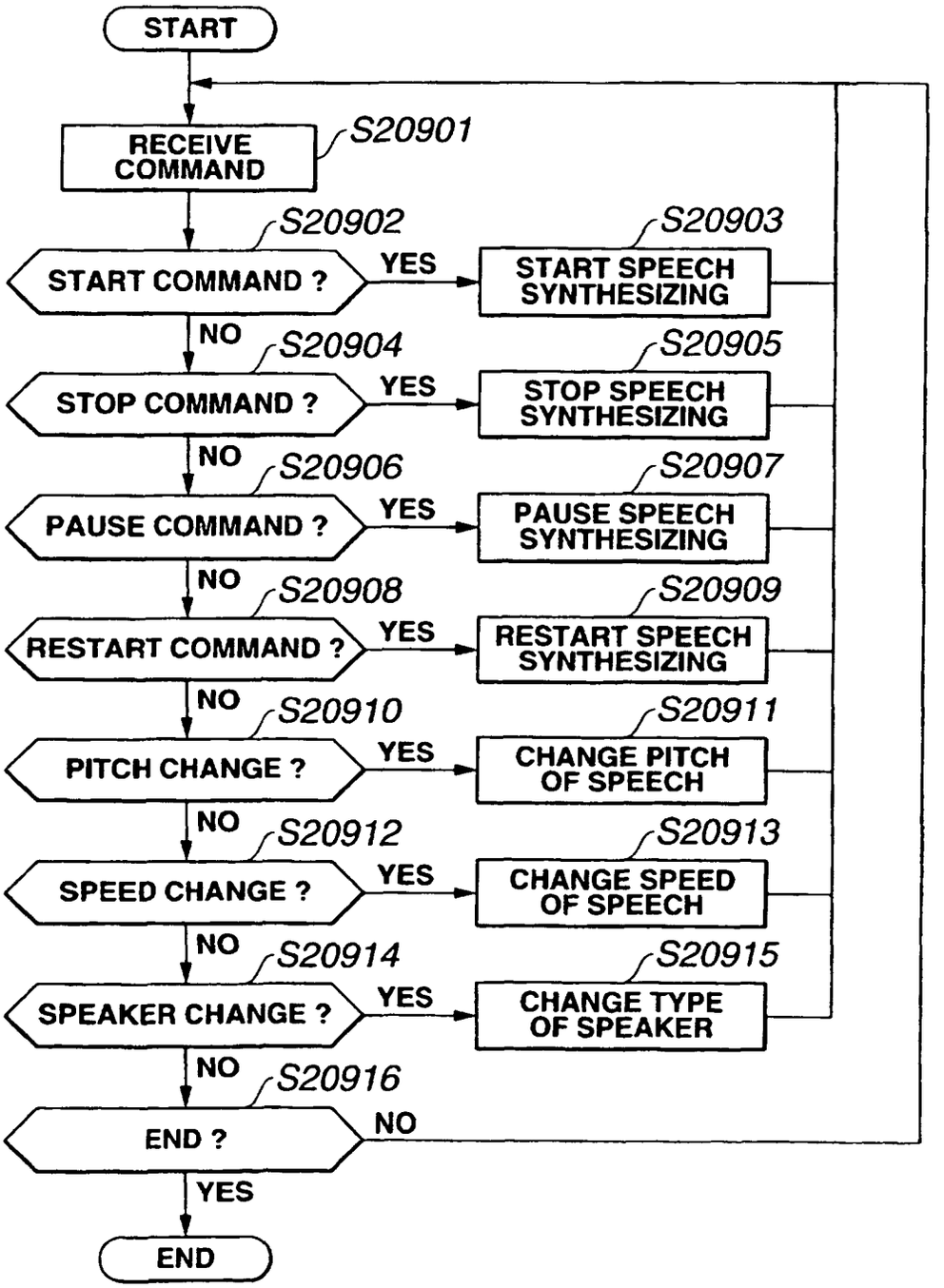

START
RECEIVE COMMAND
S20901
S20902
START COMMAND ?
YES
S20903
START SPEECH SYNTHESIZING
NO
S20904
STOP COMMAND ?
YES
S20905
STOP SPEECH SYNTHESIZING
NO
S20906
PAUSE COMMAND ?
YES
S20907
PAUSE SPEECH SYNTHESIZING
NO
S20908
RESTART COMMAND ?
YES
S20909
RESTART SPEECH SYNTHESIZING
NO
S20910
PITCH CHANGE ?
YES
S20911
CHANGE PITCH OF SPEECH
NO
S20912
SPEED CHANGE ?
YES
S20913
CHANGE SPEED OF SPEECH
NO
S20914
SPEAKER CHANGE ?
YES
S20915
CHANGE TYPE OF SPEAKER
NO
S20916
END ?
NO
YES
END

FIG.11a

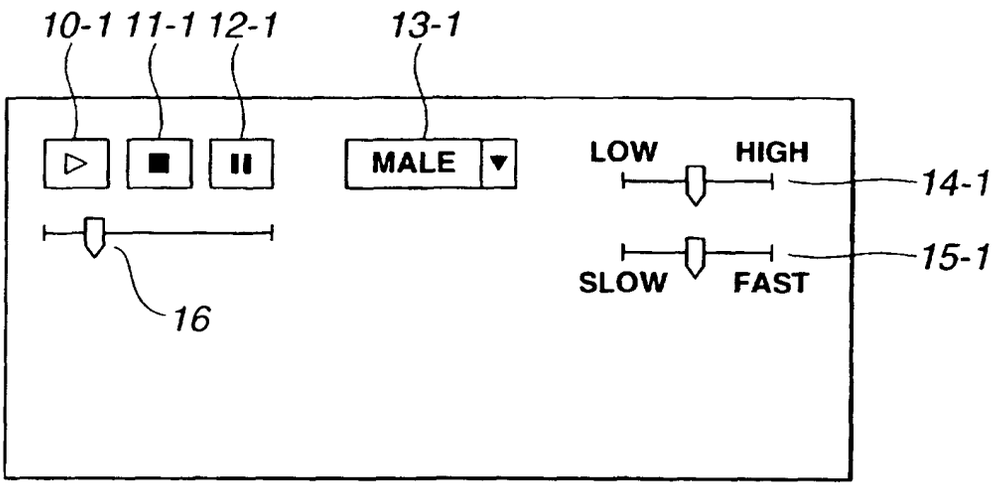


FIG.11b

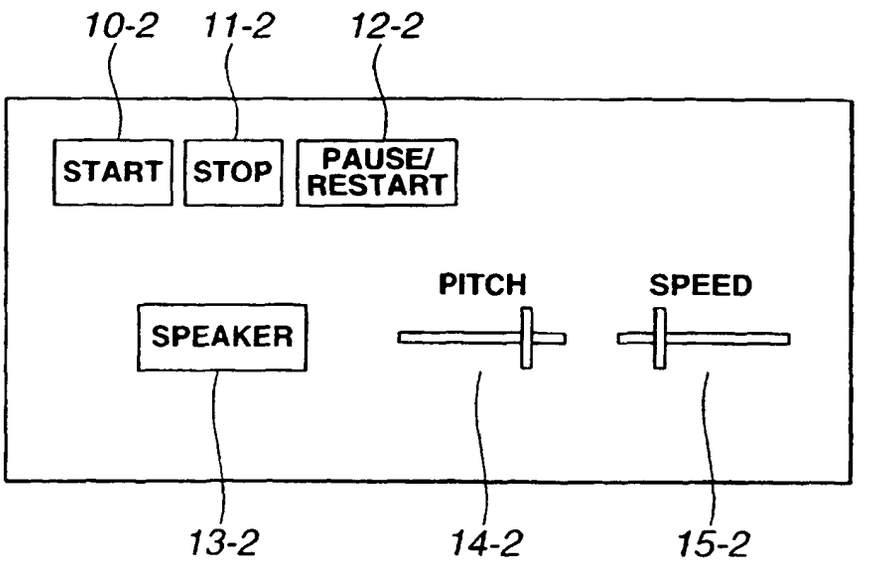


FIG.12

INPUT		NUMERIC KEY	GENERAL KEY
START AND STOP OF SPEECH		"Enter"	"Enter"
PAUSE AND RESTART OF SPEECH		" . "	" . "
SPEED	FASTER	" 6 "	" O "
	SLOWER	" 4 "	" U "
PITCH	HIGHER	" + " AND " 6 "	" ; " AND " O "
	LOWER	" + " AND " 4 "	" ; " AND " U "
SPEAKER CHANGE		" + " AND " 0 "	" ; " AND " M "

FIG.13

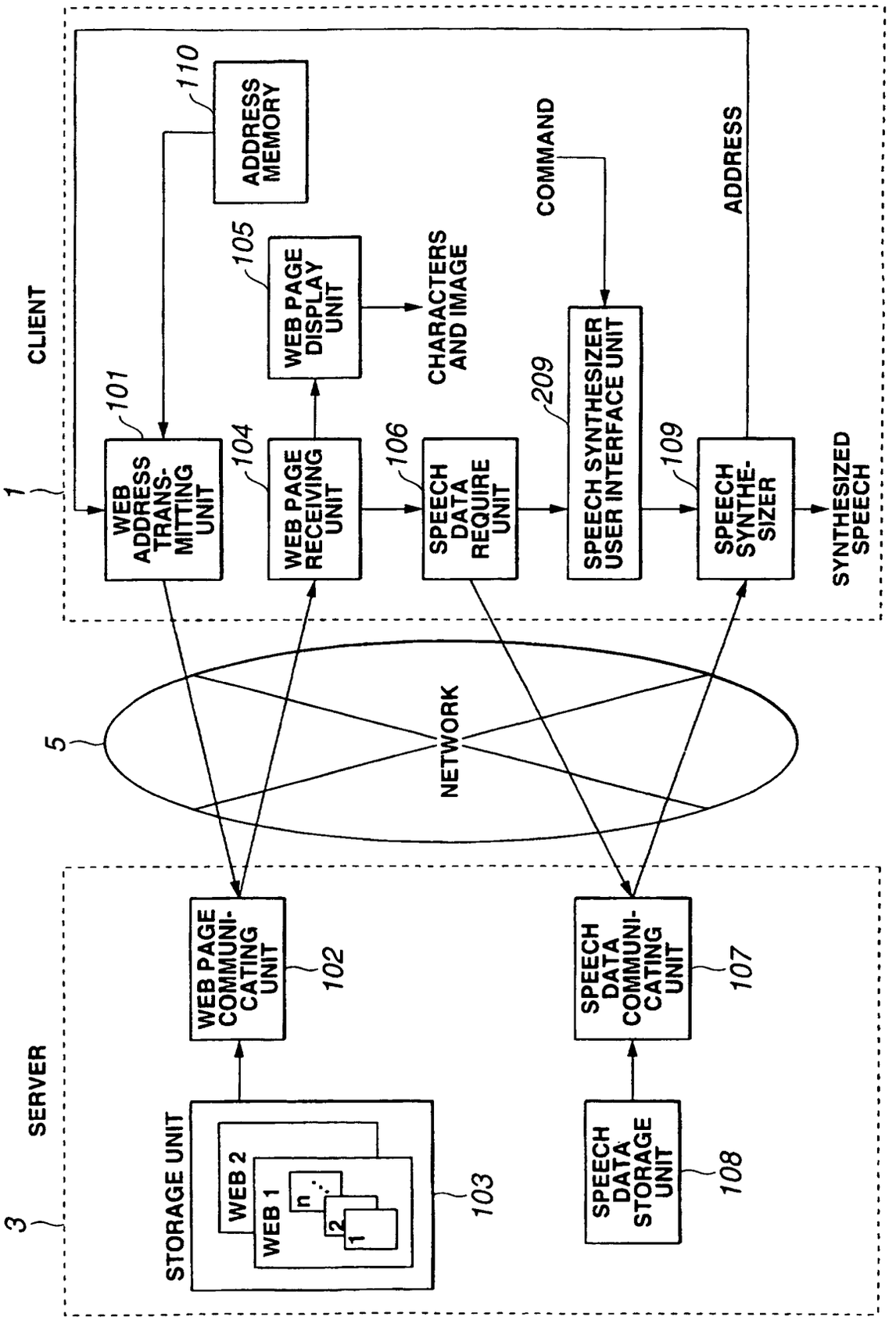

CLIENT
1
WEB ADDRESS TRANS-MITTING UNIT
101
WEB PAGE RECEIVING UNIT
104
WEB PAGE DISPLAY UNIT
105
ADDRESS MEMORY
110
SPEECH DATA REQUIRE UNIT
106
SPEECH SYNTHESIZER USER INTERFACE UNIT
209
SPEECH SYNTHE-SIZER
109
CHARACTERS AND IMAGE
COMMAND
ADDRESS
SYNTHESIZED SPEECH
NETWORK
5
SERVER
3
WEB PAGE COMMUNI-CATING UNIT
102
STORAGE UNIT
WEB 2
WEB 1
n
2
1
103
SPEECH DATA COMMUNI-CATING UNIT
107
SPEECH DATA STORAGE UNIT
108

FIG.14

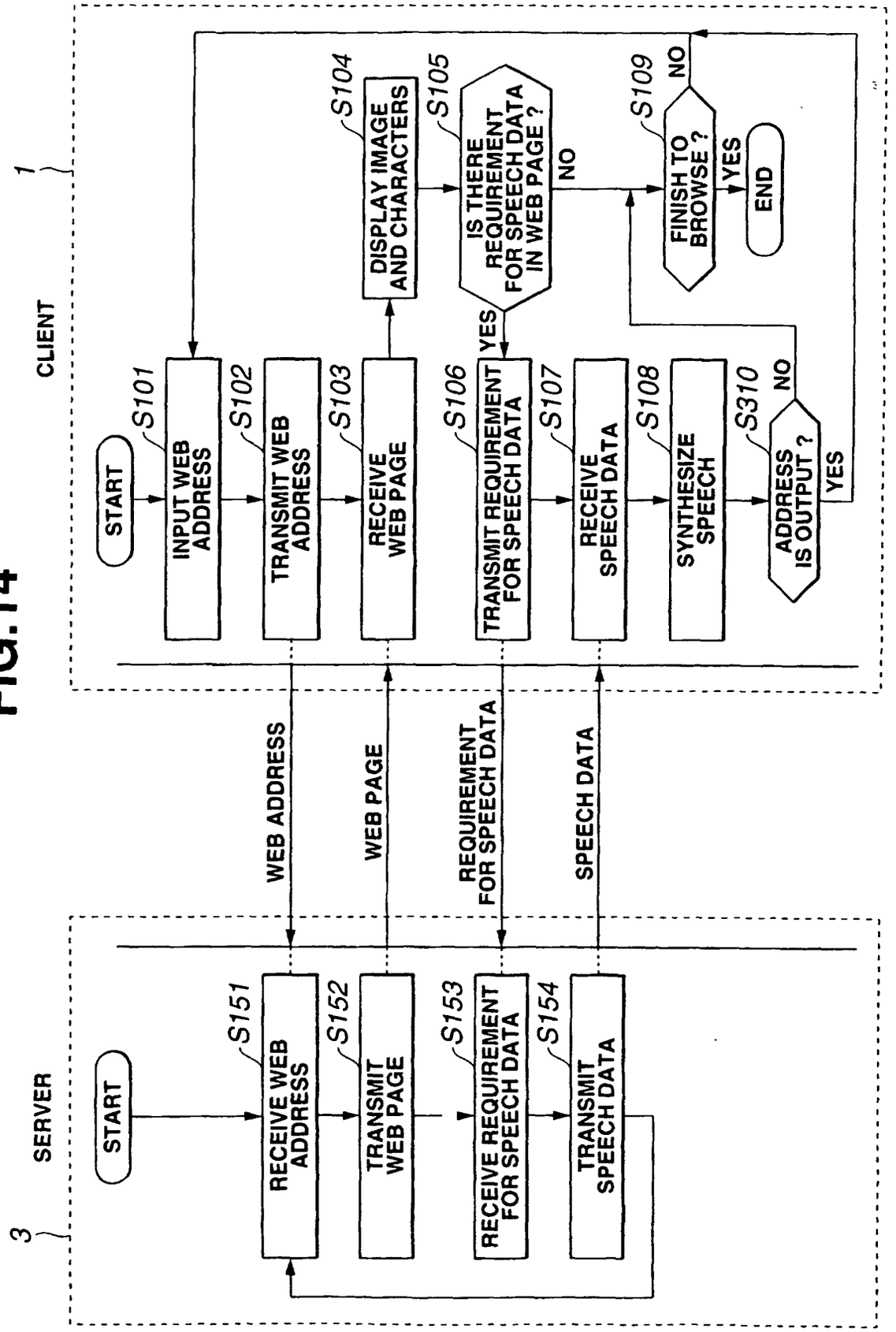

CLIENT
1
START
S101
INPUT WEB ADDRESS
S102
TRANSMIT WEB ADDRESS
S103
RECEIVE WEB PAGE
S104
DISPLAY IMAGE AND CHARACTERS
S105
IS THERE REQUIREMENT FOR SPEECH DATA IN WEB PAGE ?
NO
YES
S106
TRANSMIT REQUIREMENT FOR SPEECH DATA
S107
RECEIVE SPEECH DATA
S108
SYNTHESIZE SPEECH
S310
ADDRESS IS OUTPUT ?
NO
YES
S109
FINISH TO BROWSE ?
NO
YES
END
WEB ADDRESS
WEB PAGE
REQUIREMENT FOR SPEECH DATA
SPEECH DATA
SERVER
3
START
S151
RECEIVE WEB ADDRESS
S152
TRANSMIT WEB PAGE
S153
RECEIVE REQUIREMENT FOR SPEECH DATA
S154
TRANSMIT SPEECH DATA

FIG.15

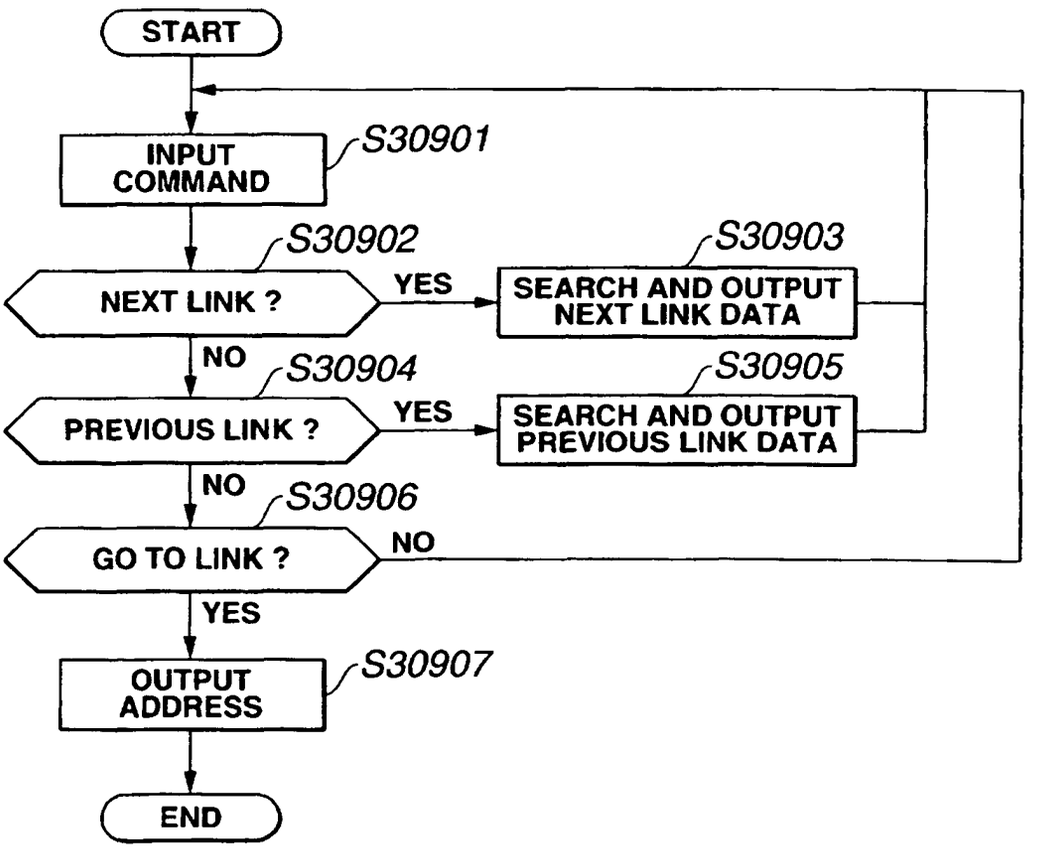

START
INPUT COMMAND
S30901
S30902
NEXT LINK ?
YES
S30903
SEARCH AND OUTPUT NEXT LINK DATA
NO
S30904
PREVIOUS LINK ?
YES
S30905
SEARCH AND OUTPUT PREVIOUS LINK DATA
NO
S30906
GO TO LINK ?
NO
YES
OUTPUT ADDRESS
S30907
END

FIG.16

(Text For Display)
Social and cultural support program.
Under the Canon's philosophy of "Kyosei" living and working together for the common good of mankind, diversified activities are promoted by Canon-group companies all over the world in the field of social contribution and cultural support.

(Speech Data)
[*Spd=7*][*Pit=10*][*Spk=female*]s ow 1 sh ax l ax n k ah 1 l ch er r ax l s ax p ao r 1 t p r ow 1 g r ae m .
[*Spd=10*][*Pit=10*][*Spk=male*]ah 1 n d er d h ax k ae 1 n ax n z f ih l aa 1 s ih f iy ah v k ay 1 y ih z iy l ih 1 v ih nx ax n w er 1 k ih nx t ih g eh 1 dh er f ao r dh ax k aa 1 m ax n g uh 1 d ax v m ae 1 n k ay n d d ih v er 1 s ih f ay d ae k t ih 1 v ih d iy z aa r p r ax m ow 1 d ih d b ay k ae 1 n ax n g r uw 1 p k ah 1 m p ih n iy z
162 [*Link="http://example.canon.co.jp/html" f ao r 1 k ae n o n g r uw p 1 k ah m p ax n ih z p ey jh ih zh , 2 p uh sh ax m k iy , n aw*] (162-1, 162-2)
ao 1 l ow 1 v er dh ax w er 1 l d ih n dh ax f iy 1 l d ax v s ow 1 sh ax l k aa 2 n ch r ih b y uw 1 sh ax n ax n k ah 1 l ch er r ax l s ax p ao r 1 t .

FIG.17

INPUT	NUMERIC KEY	GENERAL KEY
NEXT LINK	" 3 "	" L "
PREVIOUS LINK	" 9 "	" J "
GO TO THE LINK	" 0 "	" M "

FIG.18

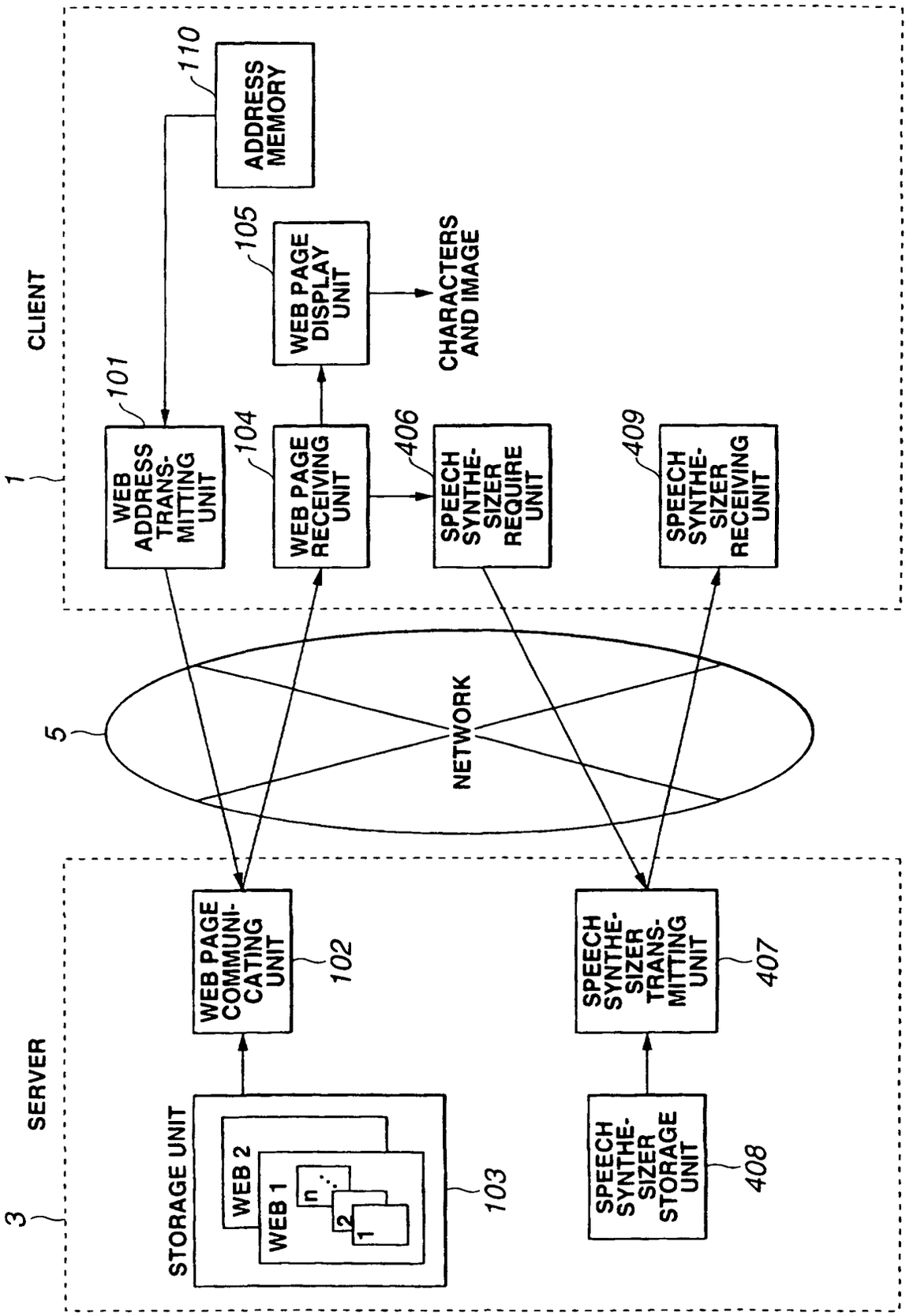
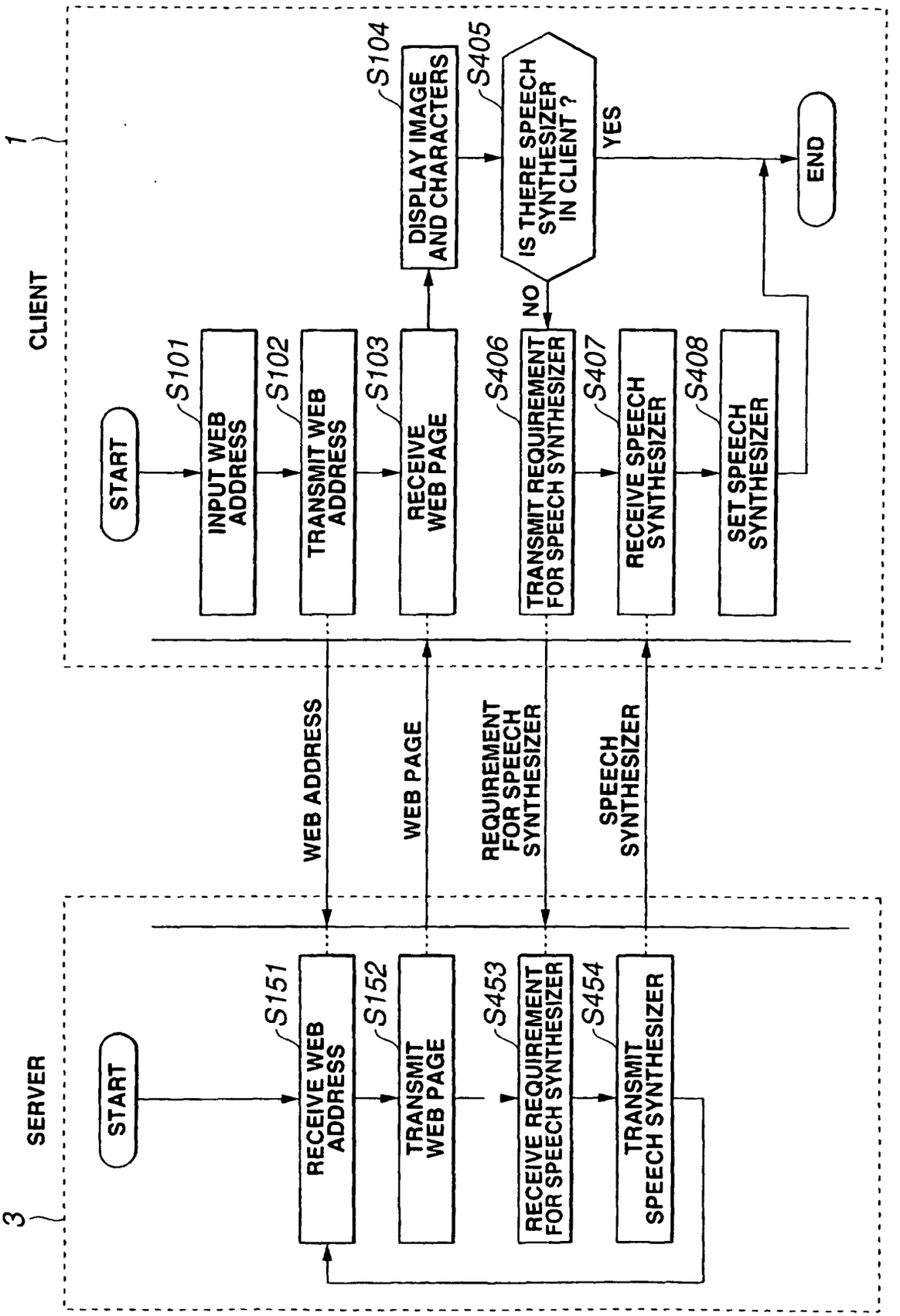

CLIENT
1
ADDRESS MEMORY
110
WEB PAGE DISPLAY UNIT
105
CHARACTERS AND IMAGE
WEB ADDRESS TRANS-MITTING UNIT
101
WEB PAGE RECEIVING UNIT
104
SPEECH SYNTHE-SIZER REQUIRE UNIT
406
SPEECH SYNTHE-SIZER RECEIVING UNIT
409
NETWORK
5
SERVER
3
WEB PAGE COMMUNI-CATING UNIT
102
STORAGE UNIT
WEB 2
WEB 1
n
2
1
103
SPEECH SYNTHE-SIZER TRANS-MITTING UNIT
407
SPEECH SYNTHE-SIZER STORAGE UNIT
408

FIG.19


CLIENT
1
START
S101
INPUT WEB ADDRESS
S102
TRANSMIT WEB ADDRESS
S103
RECEIVE WEB PAGE
S104
DISPLAY IMAGE AND CHARACTERS
S405
IS THERE SPEECH SYNTHESIZER IN CLIENT ?
YES
NO
END
S406
TRANSMIT REQUIREMENT FOR SPEECH SYNTHESIZER
S407
RECEIVE SPEECH SYNTHESIZER
S408
SET SPEECH SYNTHESIZER
WEB ADDRESS
WEB PAGE
REQUIREMENT FOR SPEECH SYNTHESIZER
SPEECH SYNTHESIZER
SERVER
3
START
S151
RECEIVE WEB ADDRESS
S152
TRANSMIT WEB PAGE
S453
RECEIVE REQUIREMENT FOR SPEECH SYNTHESIZER
S454
TRANSMIT SPEECH SYNTHESIZER