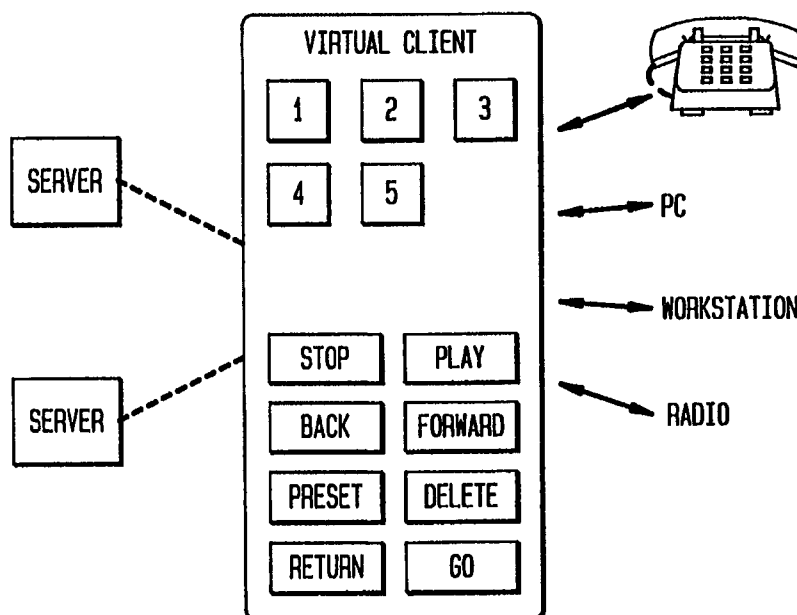




INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G10L 7/08	A1	(11) International Publication Number: WO 98/27542 (43) International Publication Date: 25 June 1998 (25.06.98)
(21) International Application Number: PCT/US97/23524 (22) International Filing Date: 16 December 1997 (16.12.97) (30) Priority Data: 08/770,025 19 December 1996 (19.12.96) US (71) Applicant: BELL COMMUNICATIONS RESEARCH, INC. [US/US]; 445 South Street, Morristown, NJ 07960-6438 (US). (72) Inventors: BARBARA, Daniel; 342 Herrontown Road, Princeton, NJ 08540 (US). NAQVI, Shamin, A.; 19 Spring Valley Road, Morristown, NJ 07960 (US). (74) Agents: YEADON, Loria, B. et al.; c/o International Coordinator, Room 1G112R, 445 South Street, Morristown, NJ 07960-6438 (US).		(81) Designated States: CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

(54) Title: AUDIO-BASED WIDE AREA INFORMATION SYSTEM



(57) Abstract

A wide area information system includes a client and at least one server on which audio data is stored. Audio data is selectively transferred to the client from the server by at least occasionally establishing a transfer connection between the server and the client. The transfer of audio data is controlled by user-selected commands and inputs that are entered via an input interface provided on the client. These commands and inputs are also transferred to the server via the at least occasionally established transfer connection between the client and the server. The audio data on the server is organized in page units that are cross-linked by way of hyperaudio links. The page units are characteristically searchable with the use of cyclically repeating directory pages of audible directory information. A data protocol is defined for enabling such an audio-based system.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

5 **AUDIO-BASED WIDE AREA INFORMATION SYSTEM****BACKGROUND OF THE INVENTION****FIELD OF INVENTION**

The present invention relates to a wide area information system for organizing and accessing information, as well as methods for implementation thereof.

10

Description of related art:

The popularity of the World Wide Web (WWW) segment of the Internet has risen dramatically in the recent past as the general public has become more aware of its existence. However, despite its popularity and diversity, audio information plays a noticeably small role on the WWW.

15

Current uses of audio on the WWW include the ability to download audio segments from certain WWW sites. Typically, a user must wait, sometimes for extended periods of time, until the entire audio file is transmitted before the audio file can be heard.

20

A system developed at Texas Instruments, Inc. allows a user to access the WWW through a speech interface, thereby obviating conventional point and click access to WWW hyperlinks. However, this system still requires a visual display terminal to show the data being browsed and to guide the input of commands. According to the TI system, a user can "jump" to different pages by speaking the name of a hyperlink.

25

It can be seen in the foregoing conventional systems that audio is still substantially a secondary object. First, the bulk of the information on a WWW page is simply not audio information. Second, the user still requires a visual interface to enable navigation, even with, for example, the TI speech interface.

30

This latter issue is a problem when a user does not have access to a computer or the like, or is engaged in activities that require the user's visual attention, such as driving an automobile.

In view of the foregoing, it is felt that audio remains an important medium for information, and can satisfy heretofore unmet demands in wide area information systems.

35

However, in applications where attempts have been made to use audio as the primary operational medium, such as phone menu systems (such as voice mail), such systems frequently overwhelm a user by requiring the user to memorize menu lists, yet

- 5 presenting so much information in such lists that recollection is made difficult or impossible.

SUMMARY OF THE INVENTION

10 It is therefore an object of the invention to provide client-server based wide area information system using audio as a primary medium for the information.

The client, sometimes called the virtual client, according to the present invention can play audio files, respond to commands input via an interface, and communicate with servers to retrieve "pages" of audio information. The client may be embodied by any number of devices, such as a telephone, a personal computer, a workstation, or a radio.

15 Servers provide audio information in units generally called audio pages, similar to the WWW. (Later in this discussion, a distinction is made between Audio Web pages and directory pages.) The audio pages are indexed via a directory service. Typically, according to the present invention, a searchable directory page is cyclically presented to the user. Generally each directory level contains audio links to a sub-directory page. The
20 last level of the directory service points to individual audio pages (i.e. Audio Web pages) that are registered in the directory service.

These and other objects and novel features of the present invention will become more apparent from the following detailed description, taken in conjunction with the accompanying drawings.

25

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates an audio client according to the present invention;

Figure 2 schematically illustrates the hierarchy of the directory service levels according to the present invention;

30 Figure 3 illustrates an option menu "window" from a cycling directory page;

Figure 4 illustrates a binary searching method usable with the present invention;

Figure 5 conceptually illustrates the process of navigating a directory page ; and

Figure 6 illustrates a Digital HyperRadio system according to the present invention.

35

5 DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is initially described herein in terms of using audio information as the basis of the system. However, this is done by way of example only, in order to fully illustrate the concepts related to the present invention. It is explicitly contemplated, as discussed later, that these concepts can be applied in other forms or combinations, such as multimedia environments.

Figure 1 illustrates a virtual audio client, and its role in the client-server system according to the present invention.

The virtual audio client (interchangeably referred to herein as client") is the means by which a user navigates the Audio Web. As illustrated in Figure 1, the client can recognize several inputs and commands. For example, the client may recognize the input of the first five cardinal numbers (one through five), and the commands STOP, PLAY, FORWARD, BACK, PRESET, DELETE, RETURN, and GO.

The client is shown in Figure 1 with an appearance similar to a video cassette recorder remote control, with buttons corresponding to the aforementioned inputs. However, this is only illustrative, and the actual manner in which the client is implemented depends on the tangible device in which the client is embodied. That is, the manner in which a user "presses" one of the conceptual input buttons shown in Figure 1 depends on the device being used to access the client. For example, a user might use a telephone, a personal computer, a workstation, a radio, etc.

If a telephone is used, for example, the client may be implemented by establishing a connection between the telephone and a computing device. The connection may be a local (i.e., hardwire) one, or it may be established by dialing to a computing facility that communicates with the phone via a standard phone connection.

A user issues commands using the phone by pressing respective keys on the telephone, with each key on the telephone being mapped to correspond to a certain command. The standard tones generated by the telephone in response to pressing respective number keys are transmitted to the computer, which is suitably able (via conventional sound recognition software) to recognize the tones so as to trigger corresponding programs to implement the desired command(s).

5 Alternatively, a user verbalizes commands into the phone. The computer, equipped with suitable voice recognition software (for example, speaker independent with a limited vocabulary), would map the recognized spoken command to the corresponding command program.

10 If a PC or workstation were used to implement the client, the implementing input device and the computing device would come in a single "package". In this instance, a user might provide input using a graphical interface combined with a mouse-type control device.

15 It is emphasized that the foregoing examples are by way of illustration, and are not meant to limit the means by which the client according to the present invention may be implemented. For example, a more extensive discussion of using radio transceivers appears later in this disclosure.

20 To give an overview of the Audio Web as contemplated, the following hypothetical scenario is offered in a dialogue format, illustrating the input provided by a user and the action of the client and server(s). In the scenario, a user is interested in finding information about the Beatles Anthology (from Capitol Records).

Client: "Welcome to the Audio Web directory server. We have lots of interesting subjects for you: Martinis...Mice...Music!"

25 User: STOP (by "pressing" the STOP key on the virtual client or verbalizing "Stop").

Client: "Please select one of the following choices: ONE: Music, TWO: Marine Life, THREE: Libraries, FOUR: Land Irrigation, FIVE: Kites"

30 User: ONE (by "pressing" or verbalizing the command).

Client: <contacts the appropriate server for the requested audio page>

35 Server: <sends the audio page over to the client>

Client: "You selected the music directory. All kinds of music can be reached from here!

40 *Baroque* < a sample of a baroque piece is played >

 ...*Rock* < a sample of rock music track is played >..."

5 User: STOP

 Client: “Your choices are: ONE: Samba, TWO: Rock, THREE: Jazz,
 FOUR: Mambo, FIVE: Polkas”

10 User: TWO

 Client: <contacts the appropriate server for the requested page>

 Server: <sends the page over to the client>

15 Client: “This is the Rock page: hear more about the greatest groups of all time:
 Beatles,...”

 User: STOP

20 Client: “You want to go to ONE: Beatles, TWO...”

 User: ONE

25 Client: “Welcome to the Audio Web Magical Mystery Tour!!

 These pages contain a collection of Beatles sound files.

 Check out the new Anthology page for the latest

30 Beatles release!

 <a sample of the new Beatles track is played>...”

 User: STOP

35 Client: “Want to visit the ONE: Anthology page, TWO: John Lennon
 page, THREE: Paul McCartney page, FOUR: George Harrison page,
 FIVE: Ringo Starr page”

40 User: ONE

 Client: <contacts the appropriate server for the requested page>

 Server: <sends the page over to the client>

45 Client: “<The initial part of the new Beatles track ‘Free as a Bird’ starts
 playing>...”

5 Thus, the system according to the present invention requires the following functionality.

 A user is provided with a way to search the Audio Web to locate audio pages of interest using a directory service. As seen in the foregoing hypothetical example, such a navigation means through the Audio Web is vital. In the example, the first three pages
10 retrieved were directory pages that were “stacked” according to the hierarchical structure shown in Figure 2.

 Once in a page, “hyperaudio” links are provided to other pages, (frequently, but not always, associated with related subjects). As seen above, for example, the user jumped from the Something New page to the Anthology page using the appropriate
15 hyperaudio link. Hyperaudio links are generally analogous to hypertext links in the WWW (and are interchangeably referred to herein as “audio links”).

 The user can input commands and other information (via the client) to control navigation through the Audio Web. The user could, for example, jump back to the Something New page after finishing with the Anthology page, then back to the Beatles
20 page again, using appropriate navigational commands, as discussed below.

 The user can also place a bookmark to mark a particular audio page, thereby enabling a direct jump to that page instead of navigating through one or more directories. Furthermore, a bookmark list is provided to catalog a user’s selected bookmarks.

 Audio pages can be authored and included in directory page hierarchies. Also,
25 authors can incorporate hyperaudio links in their audio pages.

 Finally, a standard protocol for transmitting audio tracks is used.

 To let users find pages of interest, the following schema of directories is used. When downloaded, each directory is a stream of menu choices that is cyclically presented to the user in an audible manner. The directory format is, of course, use-specific, but is
30 preferably informative, entertaining, and, most importantly, searchable. The user’s first contact with the Audio Web is with a home directory page 10, similar to a “home page” page on the WWW (see Figure 2). This home directory page 10 contains hyperaudio links 16 to one or more sub-directory pages 12. This structure continues for a number of levels, forming a hierarchical tree structure such as the one generally depicted in Figure 2.

- 5 The directory pages 14 (leaf nodes) in a last, or lowermost, level are individual audio pages that have been registered in the directory page.

The stacked "tree" organization shown in Figure 2 is desirable because the fanout of links is kept to a reasonable number while providing manageable indexing for a large number of audio pages. The tree organization used here is comparable to the B-tree
10 organization commonly used in database management systems. Each audio page has a maximum number of audio links and fixed capacity for audio information. This is advantageous because the user is not overwhelmed by extraneous information clutter while listening to a directory page, and, more importantly, the number of choices in any given option menu does not become excessive, such that the user is burdened by having to
15 remember large numbers of choices.

Another way to organize the pages according to the present invention (not illustrated) is to provide a non-stacked plurality of pages in which a given page is linked to at least one other page by a hyperaudio link, without a stacked hierarchy.

Each hyperaudio link is associated with a portion of the audio track of an audio
20 page, in the same manner that a hypertext link in the WWW is associated with a highlighted portion of the text on a page.

As mentioned above, when a user first connects to an Audio Web server, the first audio page the user encounters is a home directory page, which is in the form of a repeating audio track, generally listing a plurality of hyperaudio links. In order to use this
25 directory, the user issues a STOP command from the client, (for example, when an audio link of interest is heard). The directory then provides an option menu consisting of the last k (for example, five) hyperaudio links located before the STOP command was issued. Figure 3 illustrates the concept of the repeating directory page, and, more specifically, the set of k hyperaudio links defined in a window according to a STOP command issued by
30 the user.

At this point, the user hears a new audio track naming the k hyperaudio links. Such an audio track might be as follows:

"These are your choices at this point:

One: <name of the 1st hyperaudio link>;
35 Two: <name of the 2nd hyperaudio link>;

5 Three: <name of the 3rd hyperaudio link>..."

up to the kth hyperaudio link. The user then makes a selection by inputting the corresponding number via the client. The system therefore defines a dynamic option menu based on a time-based window of the past k links. It will be appreciated that this concept of a dynamic window guarantees that the user will be presented each time with a set of k options, regardless of the fanout (i.e. the total number of hyperaudio links defined on the given audio page). This is advantageous because the user need only deal with a relatively small number of options at any one time, keeping the selection thereof manageable.

15 However, even with the dynamic window concept, a user still could be faced with listening to a large list of hyperaudio links before finding a topic of interest.

Therefore, according to the present invention, a method is provided for facilitating searching through directory pages. In particular, a binary search methodology is useful here. Binary searching is known as a recursive way to quickly search through ordered items.

20 Figure 4 is a general illustration of the binary search methodology. In the first "pass", the search starts from the item i_m at the middle of the list. If i_m is "bigger than" (for example, numerically greater than, or alphabetically after) the item i_r being sought, the search is moved to the "left" of the middle of the list, toward the "lower" end of the list. On the other hand, if i_m is "smaller than" (for example, numerically smaller than, or alphabetically preceding) the item being sought, the search is moved to the "right" of the middle of the list, toward to the "higher" end.

30 Generally, each move is to the middle of one of the subportions defined by the middle point at the start of the pass. Therefore, in the first pass illustrated in Figure 4, the move is, in general, made from the middle of the entire list i_m , to the middle of one of the two subportions (in this case, halves) defined by i_m , where each of these subportions extends between i_m and one of the ends of the entire list.

Thus, as illustrated in Figure 4, in the first pass, a decision is made that $i_r < i_m$. Therefore, the first move is made to the left from i_m to i_l , where i_l is the item midway between the end of the list and item i_m .

5 In the second pass, a decision is made that $i_r > i_l$. Therefore, the second move is made to the right from i_l to i_r , since item i_r is midway between i_l and i_m . Since i_r happens to be the item being sought, the search is complete.

 The recursive "narrowing down" process continues as necessary to arrive at the item being sought. It is known that the number of passes necessary is proportional to the
10 logarithm of the size (number of items) of the entire list.

 As applied to the present invention, the binary search is used in a browsing/searching system in the following manner.

 When a user wants to search an audio track (which can be done at any time while the audio track is being played), the user inputs a BACK or FORWARD command, as
15 appropriate. In response, the browser locates the middle time point t_m between the current time position t_l in the audio track and the end or start of the audio track, t_e . The browser then plays the audio track from t_m forward. If the user decides that the item being sought is before (i.e., to the "left" of) t_m , then a BACK command is issued, and the browser plays the audio starting from the midpoint between t_l and t_m . On the other hand, if
20 the user decides that the item being sought is after (i.e., to the "right" of) t_m , then a FORWARD command is input, making the audio playback advance to the midpoint between t_m and t_e .

 As was explained relative to Figure 4, when this approach is repeated, the search arrives at the target item relatively quickly. When the user finally arrives at the target
25 item, the user issues a STOP command. Thereafter, the above-described dynamic menu window is built, based on the target item.

 Once the user selects a hyperaudio link from the dynamic menu window, the corresponding audio page is sent to the client from the server, and the above-described process begins over. More specifically, if the newly selected audio page is a directory, the
30 user will hear again repeated audio information, as described above. If the newly selected audio page is an Audio Web page, then the user hears the audio information contained therein, and can use it as desired. Figure 5 illustrates the concept of the navigating the Audio Web according to the present invention.

 Audio Web pages are processed in a manner similar to directory pages. The only
35 substantive difference is that Audio Web pages are not repeated as directory pages are.

5 The user receives the information linearly, as presented in the description of the page. When hearing this audio information, the user can use the STOP command, which prompts a menu of choices formed by the last k hyperaudio links in the page, as illustrated in Figure 3. Thereafter, the user can select a number corresponding to the hyperaudio link of choice, and branch or jump to the new page corresponding to that hyperaudio link
10 (see, again, Figure 5).

It will be appreciated that, over time, the user will probably come across many pages to which he or she will wish to return later, but without having to go through the relatively time-consuming process of searching through the directories. A bookmark, as its name implies, simply "marks" a particular page of interest so that it can be jumped to
15 directly. Bookmarks may be accumulated in a bookmark list that can be organized as a directory page, in the manner of the directory pages discussed above. In this context, a bookmark list has the same structure as a leaf node in the directory structure.

In order to build the bookmark list, the user can use a PRESET command to assign a bookmark to a page of interest and add it to the list by adding the name of the
20 page. The name of the page may be taken from the name of the hyperaudio link pointing to the page or, alternatively, from the name provided by (e.g., spoken by) the user, in combination with the page address. (Page addressing is discussed in more detail below.)

To retrieve a bookmarked page, the user uses a GO command that moves directly to the page in question and retrieves it for play. When the desired page is complete, the
25 user stops the audio and proceeds to branch out to other pages using hyperaudio links on the bookmarked page.

It will be appreciated that, over time, the bookmark list may itself become relatively large, and searching it can become cumbersome. Therefore, binary searching, as described above, is again used here to manage the search of bookmarks. It is assumed
30 that the bookmark list is ordered lexicographically, in order to apply binary searching.

An issue understandably arises as to how the audio tracks are played or the speakable commands recognized.

In the WWW, the Hypertext Transfer Protocol (HTTP) is successful because it is stateless. That is, whenever a WWW page is accessed, the server on which that page is
35 located simply transfers the contents of that page to the client, and does not "remember"

5 anything about the request thereafter. The client, on which the WWW browser runs, handles page layout and user interface issues. Therefore, in effect, when a hypertext link is selected ("clicked on") by the user, a separate network connection is established to bring the newly selected page to the client. This property makes WWW servers "scaleable".

10 This feature is also important in the Audio Web. It is impractical to have a server connected to the client continuously while the user is only manipulating a given page. Continuously remembering the state of connections and processing commands from multiple clients would be a tremendous burden on servers. Moreover, the system would be unscaleable as usage grows. Therefore, a Hyperaudio Transfer Protocol (HATP) is
 15 provided for the Audio Web and acts in a manner similar to the HTTP for the WWW.

In is a particular feature of the present invention that the HATP is a connectionless, or stateless, protocol that delivers audio pages to the client.

In the case of the Audio Web, the equivalent to the Hypertext Markup Language (HTML) will be a Hyperaudio Markup Language (HAML) with the following structure:

20 *pagetype* <audio track>
audiolinks < i : t_i^b , t_i^e , hatp address>

The *pagetype* indicates whether the page is a directory page or an Audio Web page. (A distinction is required here because directory pages are played cyclically whereas Audio Web pages are not.) The description of the hyperaudio links in the page is given after the
 25 entire audio track. This allows the entire audio track to be sent as a unit, preferably in a convenient data-compressed format (which may be conventionally known). Hyperaudio links have three components: link number (subscript i above), a pair of elapsed time values (relative to the beginning of the entire audio track) that represent the beginning t_i^b and end t_i^e of the portion of the audio track that serves as the hyperaudio link name, and
 30 finally, the HATP address of the page to which the hyperaudio link points.

HATP is preferably very similar to HTTP and should be a stateless protocol consisting of four stages:

1. Connection: establishment of a data connection by the client to a respective server, using a Universal Resource Locator (URL).
- 35 2. Request: sending a request message to the server from the client.

- 5 3. Response: sending a response message to the client from the server.
4. Close: closing the connection after the response message is complete.

Of course, some substantive differences will exist between HATP and HTTP, in terms of concepts known in HTTP.

1. The ACCEPT header line specifies HAML as the only format accepted by
10 the client.
2. The ACCEPT-ENCODING header line specifies the encoding format of
the HAML document (e.g., x-compress, x-zip, etc.) and the compressing format for the
audio track (e.g., MPEG).
3. A method called AUDIOSEARCH is implemented instead of
15 TEXTSEARCH, as in HTTP. This permits an object to be queried by an audio stream.
The search form of the GET method is used to query an object.

Hereinafter, the methods by which the virtual client implements the above-described functionality, according to the present invention, are discussed.

Generally, the client runs the HATP protocol to request and receive HAML page
20 descriptions. It also recognizes and implements the above-described set of commands
(e.g., STOP, FORWARD, BACK, PLAY, PRESET, etc.). Finally, the client plays the
audio tracks from Audio Web pages.

1. The STOP command

While a page is playing, the user may issue the STOP command. At that point the
25 browser stops playing the audio track, and builds up the dynamic menu window including
the last k links in the audio track prior to the point at which the STOP command was
issued.

In order to do this, the browser tracks the elapsed time from the beginning of the
audio track and also records the time at which the STOP command was issued. This time
30 is represented as t_s . Thereafter, the Audio Web browser in the client follows the following
algorithm.

Find the first link j in the link list, such that $t_j^b > t_s$;

From link j, and while searching the link list backwards, do:

 menu= \emptyset (to make the menu window initially empty);

35 m=1;

5 GL (Get Link): next link i , searching backwards
 if item i is in *menu* yet
 goto ST (if the page has fewer than, for example, 5 links);
 else,
 insert item $\langle m, \text{link name}, \text{link address}, t_i^b, t_i^e \rangle$ into *menu*.
 10 $m=m+1$
 if $m < 5$ (for example) goto GL

ST: play the menu items in the menu list and wait for *choice* (a number
 from 1 to 5)

Retrieve the page pointed to by the link address in entry $m=\text{choice}$.

15 2. As mentioned above, the FORWARD and BACK commands are useful in
 connection with binary searching of the directory pages. The browser uses the following
 algorithms to implement these commands.

 In terms of elapsed time, t_l and t_r are the beginning and ending endpoints,
 respectively, of an audio track in question, with t_s being the starting point of the audio
 20 track. At the start of playback, $t_l=t_s=0$. As mentioned above, t_r is set equal to the end
 time of the audio track.

a. When the FORWARD command is issued, the following procedure is
 followed:

 Stop playing the audio track
 25 Set $t_l=t_s$
 Set $t_s=\frac{1}{2}(t_l + t_r)$
 Start playing audio from time t_s .

b. If the BACK command is issued, then the following procedure is followed:
 Stop playing the audio
 30 Set $t_s=t_m$
 $t_s=\frac{1}{2}(t_l + t_r)$
 Start playing audio from new time t_s .

3. The RETURN command is used to order the browser to retrieve the page
 that immediately preceded the current page. Preferably, the browser maintains a memory
 35 cache to store the page immediately preceding the current page. More preferably, the

5 browser should remember the state in which the immediately preceding page was left, (by, for example, caching the hyperaudio link number to which the user jumped), so that the audio track in the immediately preceding page can be replayed from the point at which it was left.

4. The PLAY command is used to resume playback after the STOP command
10 is used to stop playback. The PLAY command causes playback to restart at the point it stopped following the STOP command.

5. As mentioned above, the PRESET command is used to add a new entry to the bookmark list. It causes the browser to add a link to the current page in the bookmark list, along with the name of the link. The browser may, for example, prompt
15 the user to verbalize the name of the link in order to add an entry to the bookmark list, or it may take the name of the link from the page in the memory cache. The bookmarked link is entered as an HAML link descriptor, as described above.

6. The GO command prompts the browser to start playing back the bookmark list as a repeating stream, thereby treating the bookmark list as a variant of the
20 directory system described previously. After the GO command is used, the system proceeds until the user uses the STOP command in order to create a dynamic menu window, or the FORWARD or BACK commands to perform a binary search of the bookmark list.

Web pages in the WWW can be queried using text strings. Therefore, a WWW
25 page may contain databases of information, such as city-by-city weather data, catalogs, and the like. A similar feature is provided for the Audio Web.

According to the present invention, a user queries an audio page by sending an audio string and receiving corresponding audio information in response. When presented with such a page, the client waits for the user to verbalize an query and send the request
30 to the server using the AUDIOSEARCH method described above.

Such audio querying requires either accurate, speaker independent speech recognition ability with a sizable vocabulary, or, in the alternative, means to “approximately match” audio streams.

5 The concepts related to the Audio Web, as disclosed herein, are particularly adaptable to radio listening. Therefore, the Digital HyperRadio ("DHR") is also presented here in accordance with the present invention.

 In order to illustrate the concept of DHR, a comparison is made in Figure 6 to the radio transmissions of an international radio broadcast by, for example, the British
10 Broadcasting Corporation ("BBC"). An individual listening to the BBC in New York City, for example, would tune a radio to a frequency used by a BBC repeater station. Using DHR, however, a list of hyperaudio links are broadcast using part of the available frequency band, in a manner similar to the explanation above. In effect, a radio
15 transmission has the format of an Audio Web page using the HAML format. Each hyperaudio link points to a respective Audio Web server which provides more information about the topic named in the link. In this arrangement, therefore, the audible broadcast radio transmission consists only of news headlines, for example, while the respective full news stories are obtained in the Audio Web pages pointed to by the radio transmission
20 "page". Using a specially equipped radio transceiver as the client, a listener stops the transmission of the news headlines, get a menu of links, select a topic of interest, and download the corresponding full news story from an Audio Web server.

 The links might also point to other radio broadcast frequencies to which the radio transceiver is automatically tuned after the selecting the corresponding link. These transmissions, for example, provide expanded information on the topic carried in the
25 initial broadcast, or offers a more elaborate list of choices.

 It will be appreciated that the concepts described herein may be usefully applied to other forms of media.

 For video information, directory pages might consist of cyclical streams of "trailers" (i.e., short clips from programs used to entice a viewer to view the entire
30 program). The user stops the directory stream using the STOP command, and a dynamic menu window is created. The menu window displays five (for example) video still frames, one from each of the five trailers shown immediately prior to the use of the STOP command. The user then chooses one of the programs by inputting ONE, TWO, ... , FIVE. Preferably, each trailer shown in the initial directory page would be overviews of a

5 common topic, such as "Music". Thus, having selected on the initial trailers, the user sees another cyclical stream of trailers related to the common topic of the initial trailer.

Once the right video page is found, the user receives the page as a video program. The page also includes hypervideo links that each point to a different video page. As with the Audio Web, the user stops the video playback at any time during the program and
10 builds the dynamic menu window of the last, for example, five hypervideo links identified in the program up to that point. Selecting one of the hypervideo links from the menu causes the client to download the corresponding video program from a server and start playing it.

Video pages according to the present invention are formatted in a manner similar
15 to audio pages, with hypervideo links having the format $\langle i: t_i^b, t_i^e, \text{address} \rangle$, where t_i^b , t_i^e represent the beginning and end, respectively, of the frame that serves as the hypervideo link name.

Preferably, the client for this format is implemented with an interactive television unit, with a remote control unit with appropriate push-button inputs for the above-
20 described commands. Advantageously, the data bandwidth needed to send commands uplink (to a proxy client running the software) is minimal.

Text information is also managed according to the present invention. For example, a ticker-type display of text can be used to show a cycling directory page. As before, the user stops the display at a desired location, builds a dynamic menu window,
25 and selects an item of interest. When the desired text page is reached, it is displayed in a limited state, or elsewhere, if an extra real state is available.

Hypertext links are encoded in a manner similar to hyperaudio and hypervideo links, using tuples in the form $\langle i: b, e, \text{address} \rangle$. The only difference is that b and e represent the displacement relative to the beginning of the text stream in bytes, instead of
30 elapsed time.

Using TextWeb, a user uses a hand-held computer unit (sometimes known as Personal Digital Assistants, or PDAs) as the client. A PDA typically has limited computing ability and a small visual display. Nonetheless, the user can satisfactorily use a PDA to display directory pages, dynamic menu windows, and text pages. Moreover,

- 5 the user may request full text pages to be sent to other devices where they can be more easily read, such as a fax machine in a fax-on-demand mode.

Finally, a multimedia web is implemented by combining audio, video, and text into displays. In this case, all three media would be first class objects. At each stage of the web concept, the user receives multimedia information in accordance with the concepts
10 describe above, relative to audio, video, and text.

It will be appreciated that this invention encompasses new browsers and navigators for the WWW as well. In particular, cyclical directories may be incorporated into the currently known methods of WWW navigation.

5 WHAT IS CLAIMED IS:

1. A wide area information system comprising:
a client; and
at least one server having audio data stored thereon and being constructed and arranged to at least occasionally have an audio data transfer connection with said client
10 whereby said audio data is selectively transferred to said client.
2. A wide area information system as claimed in claim 1, wherein said client includes an input interface and an output interface.
3. A wide area information system as claimed in claim 2, wherein said input interface includes means for controlling said transfer of audio data from said at least one
15 server to said client.
4. A wide area information system as claimed in claim 3, wherein said means for controlling said transfer of audio data comprises means for inputting a plurality of cardinal numbers and at least one command action.
5. A wide area information system as claimed in claim 3, wherein said output
20 interface comprises means for making said audio data transferred to said client audible.
6. A wide area information system as claimed in claim 4, wherein said output interface comprises means for making said audio data transferred to said client audible.
7. A wide area information system as claimed in claim 6, wherein said client comprises a telephone having a keypad and an earpiece,
25 wherein said means for inputting a plurality of cardinal numbers and at least one command action comprises respective keys or combinations of keys of said keypad which are mapped to respective ones of said plurality of cardinal numbers and to said at least one command action, wherein one of said client and said at least one server is provided with means for recognizing tones generated by said telephone when said keys or
30 combinations of keys of said keypad are pressed and for converting said recognized tones into a form recognized by said server,
wherein said means for making said audio data audible comprises means for converting said transferred audio data into a sound signal audible through said earpiece of said telephone.

5 8. A wide area information system as claimed in claim 6, wherein said client comprises a telephone having a mouthpiece and an earpiece,

 wherein said means for inputting a plurality of cardinal numbers and at least one command action comprises means for recognizing spoken words corresponding to said plurality of cardinal numbers and at least one command action received through said
10 mouthpiece and for converting said recognized spoken words into a form recognizable by said at least one server,

 wherein said means for making said audio data audible comprises means for converting said transferred audio data into a sound signal audible through said earpiece of said telephone.

15 9. A wide area information system as claimed in claim 6, wherein said client comprises a computer including an interconnected keyboard, a central processing unit, and a visual display,

 wherein said means for inputting said plurality of cardinal numbers and at least one command action comprises said keyboard,

20 wherein said computer further includes a sound generating device, said means for making said audio data audible comprising said sound generating device.

 10. A wide area information system as claimed in claim 5, wherein said audio information stored on said at least one server comprises a plurality of pages organized in a stacked hierarchy, each said page being linked by cross-reference to at least one other said
25 page.

 11. A wide area information system as claimed in claim 10, wherein said stacked hierarchy includes a home page at a topmost level thereof, said home page being the first audio data transferred to said client.

 12. A wide area information system as claimed in claim 11, wherein said home
30 page is a directory page containing audio data corresponding to a repeating audio track, said repeating audio track being made audible by said means for making said audio data audible.

 13. A wide area information system as claimed in claim 12, wherein said repeating audio track contains a plurality of cross references to pages in a level of said
35 hierarchy below said home page.

- 5 14. A wide area information system as claimed in claim 13, wherein said hierarchy includes at least one intermediate level, said at least one intermediate level containing at least one page with at least one cross-reference link to a page in a level of said hierarchy above said at least one intermediate level, and at least one cross reference to a page in a level below said at least one intermediate level.
- 10 15. A wide area information system as claimed in claim 14, wherein said hierarchy includes a bottom-most level containing at least one page with at least one cross-reference link to at least one page in a level of said hierarchy thereabove.
16. A wide area information system as claimed in claim 13, wherein said hierarchy includes a bottom-most level containing at least one page with at least one
- 15 cross-reference link to at least one page in a level of said hierarchy thereabove.
17. A wide area information system as claimed in claim 1, wherein said audio data stored on said at least one server comprises a plurality of pages, each page of said plurality of pages being linked by cross-reference to at least one other page of said plurality of pages.
- 20 18. A wide area information system as claimed in claim 10, comprising a first server and a second server, each having audio data stored thereon, said audio data on said first and second servers comprising respective first and second pluralities of pages, each page of said first and second pluralities of pages being linked by cross-reference to at least one other page of said first and second plurality of pages, respectively,
- 25 wherein at least one page in said first plurality of pages is linked by cross-reference to a page in said second plurality of pages, said first and second servers being constructed and arranged to at least occasionally have an audio data transfer connection therebetween, whereby audio data is transferable between said first and second servers.
19. A wide area information system as claimed in claim 13, wherein said means
- 30 for controlling said transfer of said audio data includes means for inputting commands for controlling said repeating audio track of said directory page.
20. A wide area information system as claimed in claim 19, wherein said means for inputting commands includes means for inputting a command to stop playback of said repeating audio track.

- 5 21. A wide area information system as claimed in claim 20, wherein said client includes means for identifying a subset of said plurality of cross references in said repeating audio track immediately prior to receiving said command to stop playback of said repeating audio track.
- 10 22. A wide area information system as claimed in claim 21, wherein said client includes means for selecting one of said cross references in said subset.
23. A wide area information system as claimed in claim 22, wherein said means for identifying a subset of said plurality of cross references generates an audible list of said cross references in said subset.
- 15 24. A wide area information system as claimed in claim 23, wherein said cross references in said subset are associated corresponding cardinal numbers, wherein said means for inputting commands for controlling said repeating audio track include means for inputting a cardinal number for selecting one of said cross references in said subset, whereby a page corresponding to said selected cross reference is transferred to said client via said audio data transfer connection.
- 20 25. A wide area information system as claimed in claim 19, wherein said means for inputting commands includes means for inputting search commands whereby said repeating audio track can be searched to find a desired cross reference.
- 25 26. A wide area information system as claimed in claim 25, wherein said search commands include a FORWARD command for advancing playback of said repeating audio track, relative to a given instant, by an interval of time, and a BACK command for moving back playback of said repeating audio track, relative to a given instant, by an interval of time.
- 30 27. A wide area information system as claimed in claim 26, wherein said intervals of time by which playback of said repeating audio track is advanced or moved back are variable in accordance with a binary search algorithm.
28. A wide area information system as claimed in claim 27, wherein said intervals of time by which playback is advanced or moved back progressively decreases in correspondence to repeated input of said FORWARD and BACK commands.
- 35 29. A wide area information system as claimed in claim 10, wherein said client includes means for identifying a page with a user-created cross-referencing bookmark,

5 whereby said page is directly selectable for transfer from said at least one server to said client by selecting said bookmark, wherein said client includes means for storing a plurality of said user-created bookmarks.

30. A wide area information system as claimed in claim 29, wherein said means for storing plurality of bookmarks is a directory page, said directory page containing
10 audio data corresponding to a repeating audio track audibly listing said plurality of bookmarks.

31. A wide area information system as claimed in claim 30, wherein said means for controlling said transfer of audio data includes means for inputting a command to retrieve a bookmarked page.

15 32. A wide area information system as claimed in claim 10, wherein said client includes a memory cache constructed and arranged to hold therein at least one page immediately previous to a most current page transmitted to said client.

33. A wide area information system as claimed in claim 32, wherein said means for controlling said transfer of audio data includes means for inputting commands to move
20 between said most current page and said at least one page held in said memory cache.

34. A wide area information system as claimed in claim 20, wherein said means for inputting commands includes means for inputting a command to resume playback of said audio track after said playback is stopped after inputting said command to stop playback.

25 35. A wide area information system as claimed in claim 29, wherein said client includes means for deleting a bookmark associated with a page.

36. In a wide area information system comprising a client and a server having audio data stored thereon in the form of a plurality of pages, the client and the server having an at least occasional audio data transfer connection therebetween, an audio data
30 transfer protocol comprising:

 a connection step, during which the audio data transfer connection is established between the client and the server;

 a request step, during which the client transmits a request for audio data to the server;

5 a response step, during which the server transmits the audio data requested by the client in said request step; and

 a close step, during which the audio data transfer connection between the client and the server is broken.

37. The audio data transfer protocol as claimed in claim 36, wherein, during
10 said request step, the client only accepts audio data, the audio data being in the form of a page and including:

 a page type identifier, whereby a page of audio data is identified as being a
 directory page, containing audio data corresponding to a repeating audio track, or a non-
 directory page, containing audio data corresponding to a linearly played audio track; and
15 the audio data transmitted from the server and corresponding to an audio track.

38. The audio data transfer protocol as claimed in claim 37, wherein said page
 further includes at least one data cluster corresponding to at least one cross reference
 linked to another said page, the data cluster including:

 an index number for identifying said at least one cross reference;
20 a time index indicating the location in said audio data, in terms of time elapsed in
 playing said audio track, of the beginning of said cross-reference;

 a time index indicating the location in said audio data, in terms of time elapsed in
 playing said audio track, of the end of said cross-reference; and

 an address of the page to which the cross-reference points.

39. The wide area information system as claimed in claim 1, wherein said client
25 is a radio transceiver, said at least one server is a radio transmitting-receiving station, and
 said audio data is selectively transmitted to said radio transceiver from said radio
 transmitting-receiving station via radio transmission.

40. The wide area information system as claimed in claim 39, wherein said
30 radio transceiver is constructed and arranged to send commands to said radio
 transmitting-receiving station via radio transmission for controlling said transfer of audio
 data.

41. A method of providing a selection menu for selecting an item from among
 a plurality of items, comprising the steps of:

35 presenting a first plurality of items in a sequence;

5 stopping the presentation of the plurality of items;
 presenting a second plurality of the items selected from the first plurality of items,
according to a point in the sequence where the presentation of the first plurality items was
stopped in said stopping step.

 42. A method of selectively transferring audio data stored on a server from the
10 server to a client, comprising the steps of:

 establishing a data transfer connection between the server and the client;
 selectively transferring audio data from the server to the client; and
 after said step of transferring, breaking the data transfer connection between the
server and the client.

15 43. The method according to claim 42, further comprising a step of specifying
audio data to be transferred in said transferring step.

 44. The method according to claim 43, wherein said step of specifying
comprises:

 providing a selection menu at the client identifying a plurality of choices, each
20 choice corresponding to a respective audio data portion;
 selecting one of the plurality of choices from the selection menu; and
 transferring the selected choice from the client to the server, such that the audio
data corresponding to the selected choice is transferred to the client from the server.

 45. The method according to claim 44, wherein said step of providing a
25 selection menu comprises:

 presenting a first plurality of the choices in a sequence;
 stopping the presentation of the first plurality of choices;
 presenting a second plurality of the choices from among the sequence of choices
of the first plurality, the second plurality being selected in accordance with the point in the
30 first plurality at which the presentation of the first plurality is stopped, whereby the
second plurality makes up the selection menu.

 46. The method according to claim 42, further comprising the step of
converting the selectively transferred audio data into an audible sound signal.

1/3

FIG. 1

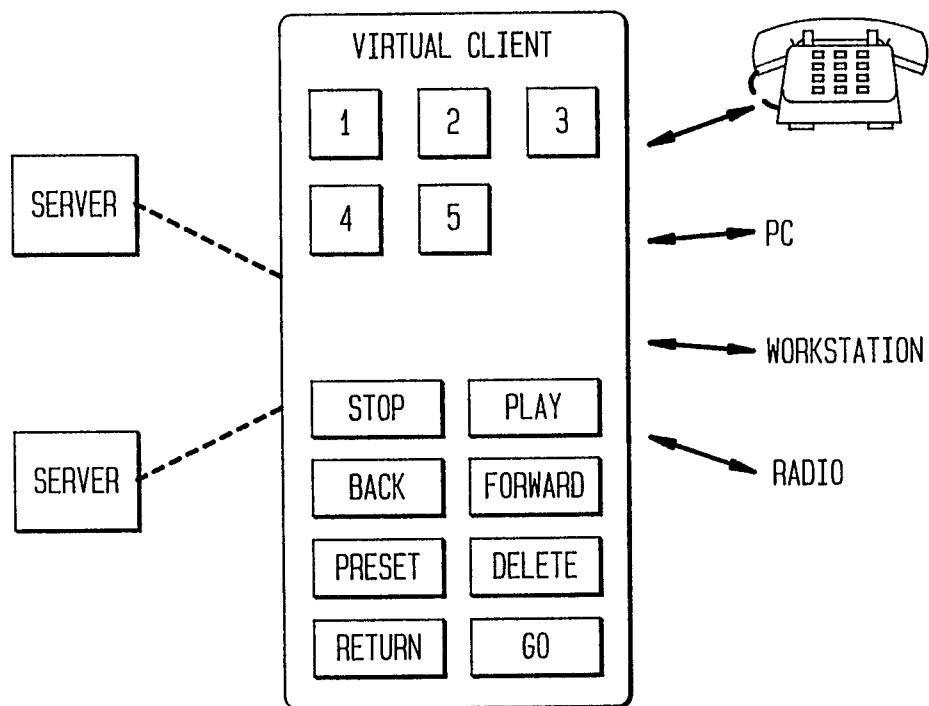
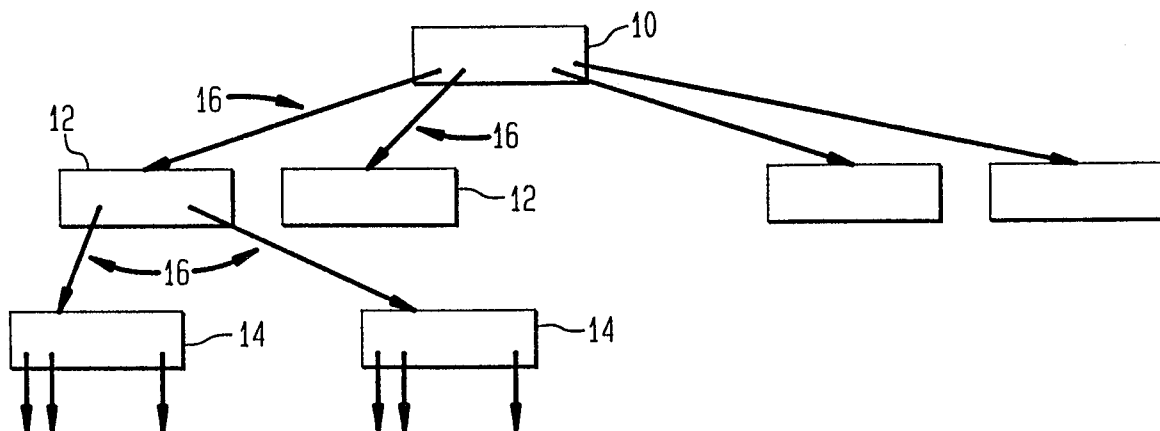


FIG. 2



2/3

FIG. 3

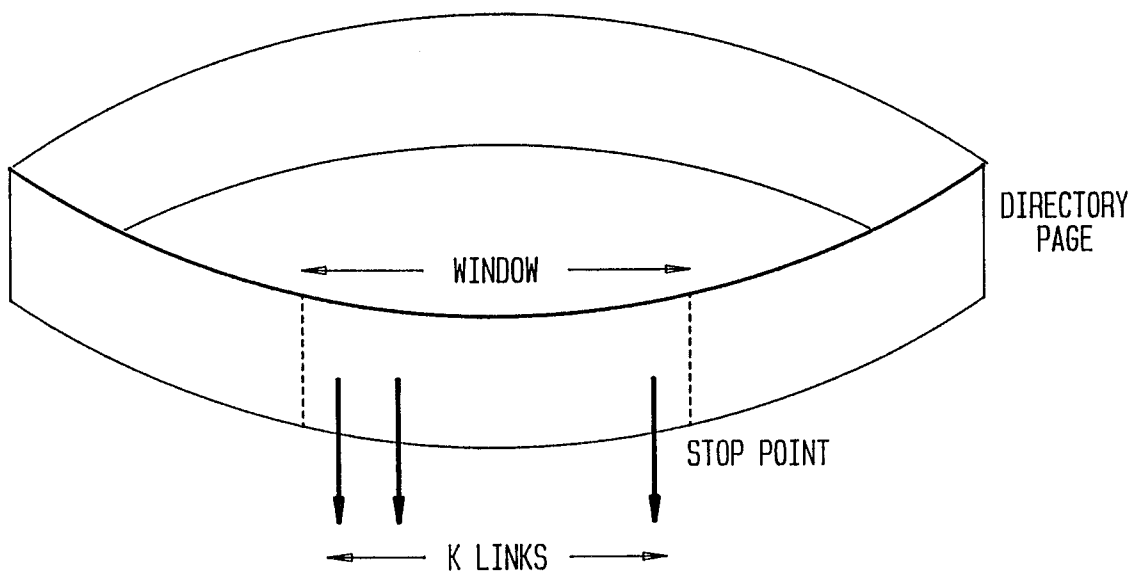
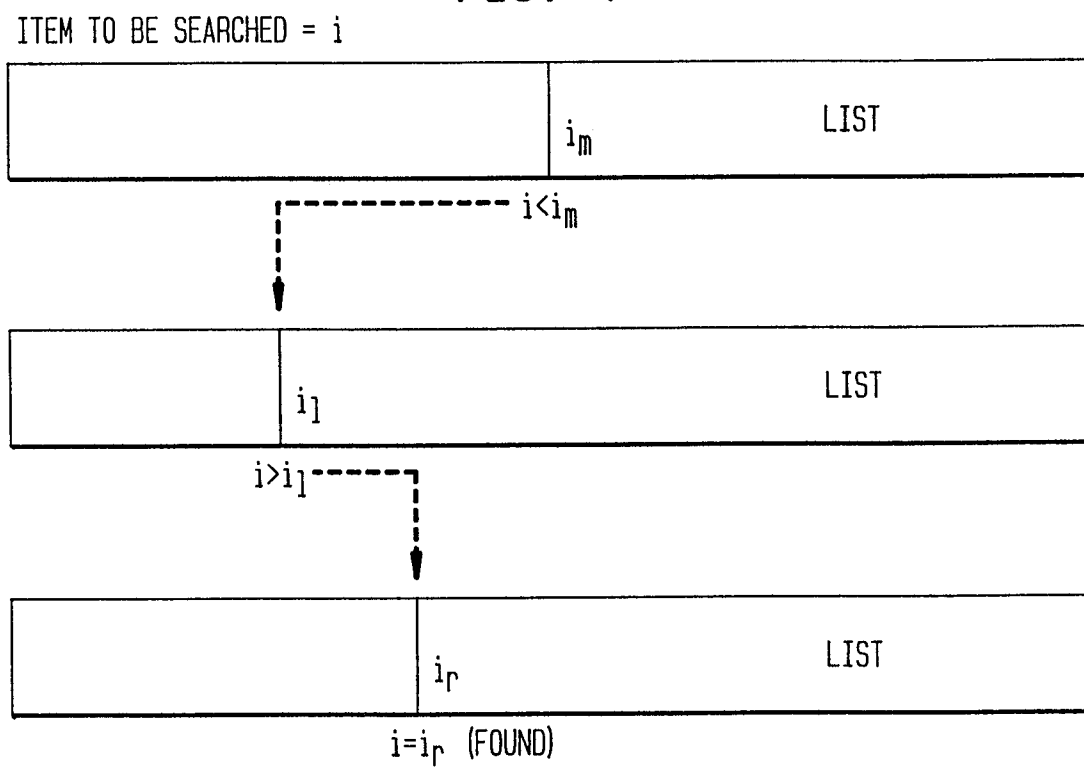


FIG. 4



3/3

FIG. 5

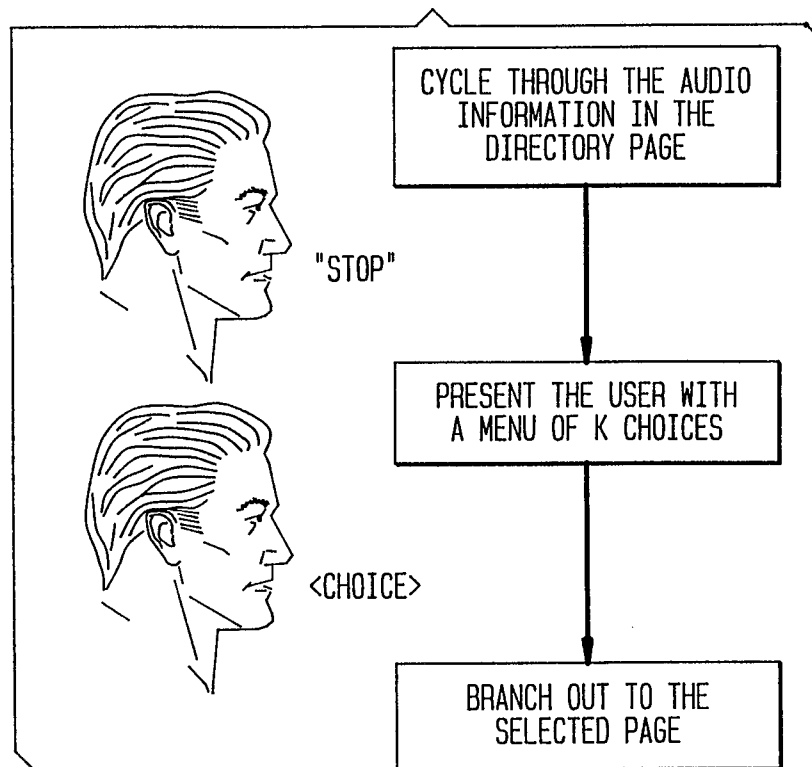
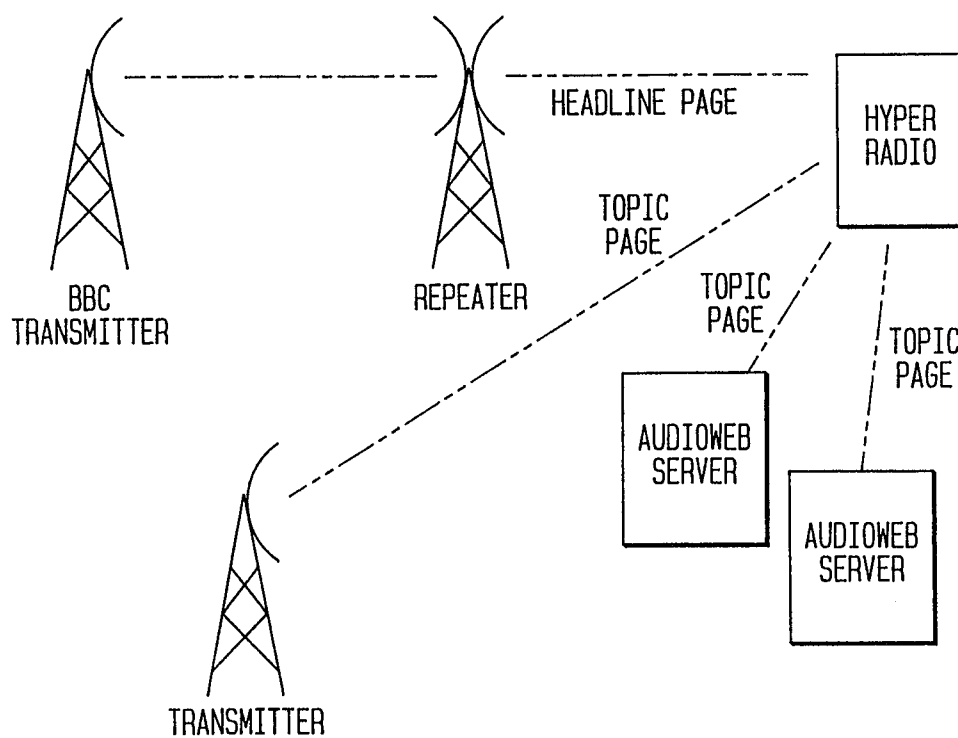


FIG. 6



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/23524

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G10L 7/08

US CL :704/251

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 704/251, 270, 272, 275, 278; 379/67, 88, 89

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

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

APS, STN (WPIDS, INSPEC.)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,524,141 A (BRAUN ET AL) 04 JUNE 1996, see Fig. 1.	1-46
Y, P	US 5,671,354 A (ITO ET AL) 23 SEPTEMBER 1997, see abstract.	1-46
A	US 5,586,235 A (KAUFFMAN) 17 DECEMBER 1996, see abstract	1-46
Y	US 5,572,643 A (JUDSON) 05 NOVEMBER 1996, see abstract.	1-46
A, P	US 5,681,108 A (MILLER) 28 OCTOBER 1997, see abstract.	1-46



Further documents are listed in the continuation of Box C.



See patent family annex.

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

Date of the actual completion of the international search

26 MARCH 1998

Date of mailing of the international search report

21 APR 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

RICHEMOND DORVIL

Telephone No. (703) 305-9600