

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
1 November 2007 (01.11.2007)

PCT

(10) International Publication Number
WO 2007/123797 A1

(51) International Patent Classification:
G06F 17/00 (2006.01) **H04L 29/00** (2006.01)

(21) International Application Number:
PCT/US2007/008063

(22) International Filing Date: 3 April 2007 (03.04.2007)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/789,176 4 April 2006 (04.04.2006) US

(71) Applicant (for all designated States except US): **JOHNSON CONTROLS TECHNOLOGY COMPANY** [US/US]; 915 East 32nd Street, Holland, Michigan 49424 (US).

(71) Applicants and

(72) Inventors: **DOUTHITT, Brian, L.** [US/US]; 2041 Laurie Lane, Holland, Michigan 49424 (US). **SCHRIPSEMA, Karl, W.** [US/US]; 996 Kenwood Drive, Holland, Michigan 49423 (US). **SIMS, Michael, J.** [US/US]; 10662 Deer Ridge Court, Zeeland, Michigan 49464 (US).

(74) Agent: **ZIEBERT, Joseph, N.**; Foley & Lardner LLP, 777 E. Wisconsin Avenue, Milwaukee, Wisconsin 53202 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

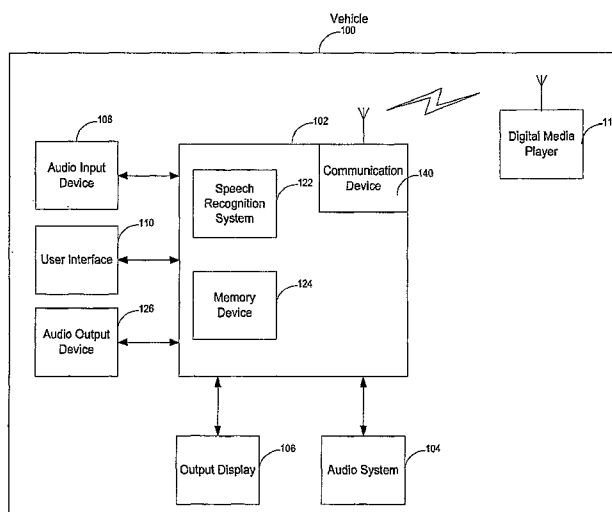
(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT,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

Published:

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SYSTEM AND METHOD FOR EXTRACTION OF META DATA FROM A DIGITAL MEDIA STORAGE DEVICE FOR MEDIA SELECTION IN A VEHICLE



(57) Abstract: A method is provided for extracting meta data from a digital media storage device in a vehicle over a communication link between a control module of the vehicle and the digital media storage device. The method includes establishing a communication link between control module of the vehicle and the digital media storage device, identifying a media file on the digital media storage device, and retrieving meta data from a media file, the meta data including a plurality of entries, wherein at least one of the plurality of entries includes text data. The method further includes identifying the text data in an entry of the media file and storing the plurality of entries in a memory.

WO 2007/123797 A1

SYSTEM AND METHOD FOR EXTRACTION OF META DATA FROM A DIGITAL MEDIA STORAGE DEVICE FOR MEDIA SELECTION IN A VEHICLE

CROSS REFERENCE

[0001] The present application claims the benefit of priority, as available under 35 U.S.C. § 119(e)(1) to U.S. Provisional Patent Application No. 60/789,176 titled "System and Method for Extraction of Meta Data from a Digital Media Storage Device for Media Selection in a Vehicle" filed on April 4, 2006 (which is incorporated by reference in its entirety).

FIELD

[0002] The present application relates generally to the field of data extraction from a digital audio/video source in a vehicle and in particular, a system and method for extracting and processing meta data from media files accessible from a digital media storage device, such as a media player, in the vehicle.

BACKGROUND

[0003] Vehicles typically include a number of electronic systems such as an audio system, a mobile telephone system, a navigation system, video entertainment system, etc. Along with those systems, media players (i.e., digital audio/video players or MP3 players) may be setup in the vehicle for a user to access digital audio or video content stored on the media player. A conventional media player includes a data storage device (i.e., memory) and an embedded software application that allows a user to transfer media files to the player and create custom lists of selected media files, referred to as playlists. Media files may be transferred to the memory of the player from various sources, including disks (e.g., CDs, DVDs, etc.) and Internet sites. Most players are configured to be coupled to a personal computer via a connection port (e.g., USB port, parallel port, serial port, etc.) in order to transfer media files. The media files typically include data, referred to as meta data, which identifies the contents of the file. For example, with digital audio files, such as MP3 files, the meta data comprises a tagging format which may include the title of the song, the artist's

name, the title of the album, track number, genre, etc. The tagging format for MP3 (i.e., MPEG Audio Layer III) files is commonly referred to as ID3.

[0004] A media player is typically configured to include a memory having one or more media files stored therein. A media player in a vehicle may be configured to be coupled to or integrated with other vehicle electronic systems, such as the audio system, for playing the media files from the player. In one known configuration, a media player may be connected via an analog audio output (e.g., a headphone jack) from the media player to the auxiliary input of the vehicle's audio system, such that the playback is controlled by the controls of the media player. In another known configuration, the media player's control interface is coupled to the control interface of the head unit of an audio system (i.e., radio), in addition to connecting the analog audio output of the media player to auxiliary input of the vehicle's audio system. In this configuration, the user can control the playback of the media from the head unit of the audio system. Yet, in another configuration, an audio system may include a database of meta data about certain media files. The audio system is configured to determine the meta data of a media file based on a unique identifier relating to that file.

[0005] It would be advantageous to provide a system for extracting meta data from a media player in a vehicle that is configured to: 1) establish a communication link between a control module and a media player; 2) identify media files stored on a media player; 3) retrieve meta data from a media file of the media player upon receiving a command from a user; 4) automatically retrieve meta data from a media file of the media player upon establishing a wireless communication link with the media player; 5) generate and/or update a database of meta data based on the meta data extracted from the media player, including converting the information to acoustic baseforms so that a user may access and manipulate the media files via voice commands and speech recognition; 6) audibly list the meta data entries of a media file based on an audible command of the user; 7) generate a second database of meta data using the meta data extracted from a second media player, wherein the meta data of the first media player is retained across power cycles, thereby eliminating the time consuming process of extracting meta data each time a media player is connected to the system; and 8) provide for the use of multiple databases concurrently, for example, when two or more media players are linked to the control module concurrently.

SUMMARY

[0006] In accordance with one embodiment, a method is provided for extracting meta data from a digital media storage device in a vehicle over a communication link between a control module of the vehicle and the digital media storage device. The method includes establishing a communication link between control module of the vehicle and the digital media storage device, identifying a media file on the digital media storage device, and retrieving meta data from a media file, the meta data including a plurality of entries, each entry including text data or enumerated code. The method further includes identifying the text data or code in each entry of the media file and storing the plurality of entries in a memory.

[0007] In accordance with another embodiment, a control system in a vehicle is provided for extracting meta data from a digital media storage device over a communication link. The system includes a communication module for establishing a communication link with the digital media storage device, a processing module coupled to the communication module, the processing module being configured to retrieve via communication module meta data associated with a media file from the digital media storage device, the meta data including a plurality of entries, wherein each of the plurality of entries includes text data, and a memory module configured to store the plurality of entries retrieved from the digital media storage device.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] The innovation will be more readily understood by reference to the following description taken with the accompanying drawings, in which:

[0009] Figure 1 is a block diagram of a control system for extracting meta data from a media player in a vehicle in accordance with an exemplary embodiment.

[0010] Figure 2 is a block diagram of a control system for extracting meta data from a media player in a vehicle including the control module of FIG. 1 in accordance with an embodiment.

[0011] Figure 3 is a block diagram of an integrated system for controlling a vehicle's audio and telephony systems.

[0012] Figure 4 illustrates a method for extracting meta data from a media player in response to a user command in accordance with an embodiment.

[0013] Figure 5 illustrates a method for processing media file data from a media player and organizing the media files in accordance with an embodiment.

[0014] Figure 6 illustrates a method for speech-enabled playback of a media file in accordance with an embodiment.

[0015] Figure 7 illustrates a method for audibly listing media files in accordance with an embodiment.

DETAILED DESCRIPTION

[0016] Figure 1 is a block diagram of a control system for extracting meta data from a media player in a vehicle in accordance with an exemplary embodiment. The control system enables a vehicle occupant, such as a driver, to extract the meta data from the occupant's media player 114 and subsequently generates a phonemic representation of the extracted text data. Other vehicle occupants (e.g., passengers) may also use the system to extract the meta data from their personal media player 114. Alternatively, a control system may be provided that extracts meta data from a mass storage device. However, the description herein will primarily refer to an exemplary control system for extracting meta data from a media player 114. It should be understood that although a system for extracting meta data will be described in detail herein with reference to media player 114, one or more of the systems and methods for extracting meta data herein may be applied to, and find utility in, other types of digital storage or audio/video devices as well. For example, one or more of the systems for extracting meta data may be suitable for use with a flash drive, a micro hard drive, a USB memory stick, etc., all of which may require certain device-specific hardware and software to extract the meta data.

[0017] The exemplary system shown in Figure 1 enables a wireless communication link to be established between an audio system 104 of vehicle 100 and a media player 114 (e.g., a media player with a Bluetooth-enabled connection) of a vehicle occupant. Alternatively, a control system may be provided that includes a wired connection to a vehicle occupant's personal media player 114. The wired connection may include a cradle for media player 114 or a similar wired adapter, such as an analog audio output connection (e.g., a headphone jack). The following description will refer to an exemplary control system using a wireless communication link between the vehicle audio system and media player 114 of a vehicle occupant, however, it should be understood that methods and systems for retrieving and processing meta data from a media player described herein may be used in conjunction with

other configurations of control systems, such as a wired connection to a vehicle occupant's media player, etc.

[0018] In Figure 1, the control system in vehicle 100 includes a control module 102 coupled to the audio system 104 and an antenna 112. In an alternative embodiment, control module 102 may be incorporated into audio system 104. Antenna 112, in conjunction with a communication device 140 in control module 102, enables two-way communication with media player 114. Communication device 140 may be configured as a communication circuit including analog and/or digital components, such that communication device 140 is capable of transmitting and receiving short-range radio frequency ("RF") signals in any of a variety of data transmission formats, such as a Bluetooth communications protocol, an IEEE 802.11 communications protocol or other personal area network wireless communications protocols or data formats. Media player 114 may include a plurality of transceiver circuits, for example, a Bluetooth transceiver circuit configured to communicate with communication device 140.

[0019] Control module 102 is also coupled to a user interface 110 and an audio input device 108. User interface 110 may be used to receive input commands from a vehicle occupant via, for example, pushbuttons, switches, a keypad, a touch screen display, etc. Alternatively, input commands to control module 102 may include a set of audio signals from a vehicle occupant. For example, a vehicle occupant may speak directly into the audio input device 108 to provide input commands to the control module 102. Audio input device 108 may include one or more audio input devices configured to receive an oral command from a vehicle occupant. The oral command may be any word or phrase that the occupant may speak, utter, or otherwise provide to cause the control system or another system to perform a function. A speech recognition system (or data processing module) 122 in control module 102 may be used to process various data signals, such as audio signals (e.g., oral input commands) received via audio input device 108 and to recognize words or phrases in the oral command. Voice recognition technologies known in the art may be implemented in speech recognition system 122. For example, speech recognition system 122 may comprise any speech recognition software or engine such as IBM Embedded ViaVoice®, manufactured by International Business Machines Corporation.

[0020] Control module 102 may also be coupled to a display 106. Display 106 may comprise a small cathode ray tube (CRT), liquid crystal display (LCD), or various other types of visual displays which are easily visible in various lighting conditions. Control module 102 may comprise one or more analog and/or digital electrical or electronic components, and may include a microprocessor, a microcontroller, application-specific integrated circuit (ASIC), programmable logic and/or other analog and/or digital circuit elements configured to perform various input/output, control, analysis and other functions described herein. Control module 102 may also include memory 124, including volatile and non-volatile memory in order to, for example, store a computer program, other software, or a data file to perform the functions described herein. Each element in vehicle 100 shown in Figure 1 may be mounted in or coupled to the same or different interior vehicle elements such as a dashboard, an overhead console, an instrument panel, a visor, a rearview mirror, or other appropriate locations within the vehicle. Preferably, user interface 110, audio input device 108 and output display 106 are mounted in a location convenient to the vehicle occupants, in particular, the driver.

[0021] Audio system 104 may include, for example, a radio, an amplifier and at least one speaker configured to provide audio signals to one or more vehicle occupants. The speakers (not shown) are configured to receive audio output data from control module 102 and/or other systems (e.g., a digital audio file from a media player, information prompts or other messages provided by control module 102, etc.). The speakers may be a part of the vehicle audio system 104 or may be a dedicated audio output device 126 serving only control module 102.

[0022] As mentioned above, in the exemplary system shown in Figure 1, control module 102 is linked via a wireless communications link with a media player 114 of a vehicle occupant. A communication link may be established with different media players as different occupants are enabled to use the control system. A communication link may be established, for example, in response to a command received from the user via user interface 110 or an oral command received via audio input device 108. The wireless communication link operates in accordance with a wireless communication standard, such as Bluetooth communication protocol, IEEE 802.11 communication protocol, or other appropriate wireless communication standards. Accordingly, media player 114 is enabled

to operate in accordance with or to communicate with the wireless communications standard used for the wireless communication link.

[0023] The communication link between control module 102 and media player 114 enables extraction of meta data from the media files stored in media player 114 to control module 102. Subsequently, the playback of the media files is controlled from user interface 110 or audio input device 108. For example, after a communication link is established between a media player 114 and control module 102, the process of extracting the meta data may begin upon establishing the communication link or as a result of a command by a user. The user may provide the command via inputting the command into the user interface 110 or spoken commands into audio input device 108. For example, a song may be played through control module 102 and audio system 104 by receiving an oral command from a vehicle occupant and sending the playback command to media player 114 via the wireless communication link. Audio information received by control module 102 may be provided to the user via audio system 104 (e.g., via speakers, not shown) and data retrieved from the media player 114 may be displayed on display 106. Media player 114 may be located anywhere within the proximity of vehicle 100, such as in an occupant's pocket or briefcase, in the trunk or within a range of communication with communication device 140.

[0024] Figure 2 is a block diagram of a control system for extracting meta data from a media player in a vehicle including the control module 102 of Figure 1. Memory 124 includes volatile memory 136 and non-volatile memory 138 for storing computer or software programs to perform functions of control module 102. For example, memory 124 includes a predetermined vocabulary of command words that may be recognized by speech recognition engine 130 of speech recognition system 122. The predetermined vocabulary may be used to manipulate systems and applications in the vehicle, in order to perform a function. Speech recognition engine 130 is configured to recognize words or phrases in an oral command and to compare the recognized word to the predetermined vocabulary of command words, in order to determine an appropriate action and/or the appropriate system or application to which an instruction should be delivered. Speech recognition system 122 also includes a text-to-speech engine 132 and a text-to-grammar engine 134. Text-to-speech engine 132 is configured to convert text (e.g., ASCII text) to an audio representation (e.g., PCM) that may be played through audio output devices 126 or audio system 104. Text-to-grammar engine 134 is configured to convert text (e.g., ASCII text) to a phonemic

representation that may be used by speech recognition engine 130 to recognize words or text in an oral command from a user. Text-to-speech engine 132 and text-to-grammar engine 134 may be part of the same application or different applications. Text-to-speech and text-to-grammar technologies known in the art may be used such as Embedded ViaVoice® Text-To-Speech engine and Text-To-Grammar engine, manufactured by International Business Machines Corporation.

[0025] Memory 124 includes a database 158, which may also be used to store meta data for the media files of each unique media player or other mass storage device for use by a user during operation of the control system with a particular media player. Memory 124 may include meta data for each media file extracted from a media player. The meta data of each media file includes a plurality of entries. Each media file comprises audio data and meta data, wherein the meta data may include a plurality of entries, representing a context of each media file in memory 124, such as song title (e.g., "We Belong Together"), album title (e.g., "Breakthrough"), artist (e.g., John Doe), genre (rhythm and blues), time length (e.g., 3:10), and track number (Track 5), etc. In Figure 2, media player 114 include media file data 128, such as song title, album title, artist, genre, time length, track number, etc., stored in, for example, memory 148 of media player 114. A user may utilize known methods to add, edit, and other wise manipulate media files in media player 114. Memory 148 may also include meta data for each media file in memory 148 of media player 114. Each media file may include a plurality of entries, representing a context of each media file in memory 148, such as song title, album title, artist, genre, time length, and track number, etc. As mentioned, control module 102 includes database 158 stored in memory 124. Preferably, database 158 is stored in non-volatile memory 138 (e.g., flash EPROM) and is, therefore, maintained across power cycles of the control system.

[0026] In one embodiment, the media file data 128 are retrieved automatically by control module 102 when a wireless communications link is established between control module 102 and media player 114. In one embodiment, media file data 128 retrieved from media player 114 may be stored in volatile memory 136 (e.g., DRAM). Accordingly, the media file data stored in volatile memory 136 is not preserved across power cycles of the control system. In an alternative embodiment, the media file data may be stored in non-volatile memory 138 (database 158 shown in Figure 2) and may be maintained across power cycles and available to a user after power-on initialization. When a second wireless

communications link is established with a different media player (e.g., the media player of another vehicle occupant), the data in volatile memory 136 may be overwritten with the media file data from the second media player. Alternatively, each user may have a media files with meta data from their associated media player stored in non-volatile memory 138. In this embodiment, the media file data for a particular media player 114 would only be accessible when that particular user's media player 114 is connected to the control system via the wireless communications link. Each media player 114 may have a unique identifier that identifies one user's media player 114 from another media player.

[0027] Figure 3 is a block diagram of an integrated system for controlling a vehicle's audio and telephony systems. The control system of Figure 3 enables a user to connect a digital media storage device, such as an iPod® media player, in order to extract meta data from the digital media storage device. Figure 3 illustrates a preferred embodiment of the software architecture employed to support a control system of Figure 1.

[0028] Music manager 242 includes a media cataloger 248 which is configured to extract the meta data from the connected media device (e.g., mass storage device, media player, mobile phone, etc.). In the case of mass storage device 278, it may be connected via a universal serial bus (USB) connection. Media cataloger 248 searches through the file structure of mass storage device 278, in order to identify audio media stored in memory of the device. Media cataloger 248 is configured to parse the media file to locate the embedded metadata (ID3 tags) and extract them. USB Playback 244 decodes the audio media during file reading and playback of the audio media files. Media cataloger 248 also searches through the audio media files for various file parameters, including an ID3 tag, which is a tagging format for MP3 files. Media cataloger 248 may be configured to perform the meta data extraction on a file by file basis, such that media cataloger 248 creates a database, which is shown as music database 252. Media cataloger is configured to request that the user re-catalog the media files on the device, when it detects that a modification or addition has been made to the media files. It should be understood that mass storage device 278 may include media storage devices, such as a thumb drive, a USB stick, etc.

[0029] In the case of media player 274 (such as an iPod® media player), it may connect to music manager 242 through a wireless connection with a Bluetooth dongle. An after-market dongle, called naviPlay™, is currently manufactured by TEN Technology.

Impulsesoft Corporation also created a proprietary protocol that enables the system to extract the meta data from an iPod® media player. Upon connecting the Blue tooth dongle with an iPod® media player, the iPod control 264 recognizes that an iPod® media player is connected. Media cataloger 248 interfaces with Bluetooth control module 260, in order to request the meta data over the Bluetooth connection. Media streaming 266 includes a streaming profile such as advanced audio distribution profile (A2DP), which is used for controlling the streaming of audio data from media player 274 or mobile telephone 276 to the system. It should be further understood that the scope of the present innovation is not limited to an iPod® media player. Media player 274 may include any of several types of media storage devices, digital audio/video players, etc., capable of storing media files in memory.

[0030] Additionally, mobile telephone 276 is shown coupled to Bluetooth control module 260 through Bluetooth host stack 262. Bluetooth control module 260 preferably includes a hands-free profile, in order to inter-operate with the Bluetooth HFP-enabled mobile phone. Alternatively, mobile phone 276 may be a Bluetooth-enabled phone with a hands-free profile. In an alternative embodiment, mobile telephone 276 may also include a media player for streaming-in media files over a Bluetooth connection. Mobile telephone 276 is configured to stream-in audio via media streaming 266. Media cataloger 248 is further configured to extract the meta data from mobile telephone 276, wherein the meta data are stored in music database 252.

[0031] Music database 252 preferably includes a client application programming interface (API) 254 for media cataloger 248 to invoke for requests for cataloging media files. Music database 252 may also include SQL database engine 256 and music database schema 258, serving as a relational database management system. A memory module 268 is coupled to music database 252. Music database 252 includes a volatile memory 270 (e.g., RAM) and a non-volatile memory 272 (e.g., flash). Accordingly, a meta data catalog stored in volatile memory 270 is not preserved across power cycles of the system, whereas the media file may be stored in non-volatile memory 272 and may be maintained across power cycles and available after a power-on initialization.

[0032] Once the media device (e.g., mass storage device 278, media player 274, mobile phone 276, etc.) has connected, music manager 242 informs state manager 222 that a media

device has been located (e.g., a Bluetooth-enabled device is found or a mass storage device has been connected). If cataloging for the media device is necessary, music manager 242 informs state manager 222 that cataloging has been completed. State manager 222 uses these data to update graphical user interface (GUI) 212 and voice user interface (VUI) 232 that cataloging is occurring and/or has been completed via UI Context Manager 224 and integration module 225. The interface between state manager 222, GUI 212, and VUI 232 demonstrate how a synchronous multimodal user interface is achieved. Existing focus management 226 controls various vehicle systems when multiple systems are in use at the same time. Existing hands-free application 228 controls the operation of the connected mobile phone and other vehicle systems when the mobile phone receives a call. For example, if the user is playing back music when the mobile phone receives a call, the system may be configured to pause music playback at that time. Subsequently, the user can disconnect the call and have the music playback automatically resume.

[0033] VUI 232 includes speech recognition engine 234, text-to-speech engine 236, and text-to-grammar engine 238. Microphone 240 and speakers 250 are preferably coupled to VUI 232 for processing various data signals, such as audio signals (e.g., oral input commands, audio output data, etc.). When the cataloging process is complete, state manager 222 notifies VUI 232 that the media device has been cataloged. Upon receiving the update from state manager 222 that cataloging is complete, VUI 232 is configured to parse through music database 252 for each of the entries of the meta data and convert the entries to an acoustic baseform (i.e., phonemic representation of the entry). The acoustic baseforms are stored in a baseform list in volatile memory 270. Therefore, the baseform list is not preserved across power cycles of the system. In an alternative embodiment,

[0034] music manager 242 is notified by a power module (not shown) that the system is shutting down. Music manager 242 writes the RAM-based database(s) to non-volatile memory 272, in order to be maintained across power cycles.

[0035] Additionally, VUI 232 includes software logic for converting text data, including acronyms, punctuation, and unique sounding names, into an acoustic baseform. Text-to-speech engine 236 is configured to convert several such meta data entries. For example, text-to-speech engine 236 is configured to recognize and convert the artist's name "INXS," which phonetically sounds like "in excess." Another example would be the artist's name

“.38 Special,” which phonetically sounds like “thirty-eight special” instead of “point three eight special” In one embodiment, the software logic is embedded directly into VUI 232, wherein VUI 232 best converts the text data according to its own rules. In an alternative embodiment, the system may use a connected mobile phone (e.g., mobile phone 276) to create a dial-up connection to the Internet to access an online music database. Alternatively, the system can access a shared network drive or a database that includes a phonetic representation in speech synthesis mark-up language (SSML). Text-to-speech engine 236 may use the SSML to pronounce the entry correctly and to create the correct acoustic baseform for the entry of the meta data.

[0036] Music manager 242 further includes a USB playback 244 and iPod® media player playback 246. Using either in GUI 212 or VUI 232, the user may select a mode to playback media files with USB playback 244 or iPod® media player playback 246. When the user selects a mode of playback, GUI 212 and/or VUI 232 determine how playback will happen. Upon selecting the artist to be played (e.g., “play the Beatles”), state manager 222 notifies music manager 242, and music manager 242, through the two playback sub-components, accesses music database 252 through the client API 254 to find all the tracks associated with the specified artist. The tracks may then be played back in a pre-determined order or at random. For example, a user may decide to playback the Beatles’ albums in alphabetical order and then within the albums, in the individual track order. Alternatively, a user may decide that to play the albums in the chronological order, starting with the first Beatles’ album.

[0037] Tactile user interface (TUI) 202 includes rotary input 204, general purpose input/output (GPIO) 206, and radio input/output (“radio IO”) 208. TUI 202 is coupled to GPIO 210 and I2C unit 214 (or I²C), which is a serial bus. I2C unit 214 is coupled to GPIO Extender 216 and radio unit 218. Radio unit 218 preferably includes an AM/FM tuner, a power amplifier, equalizer, a satellite radio, and an auxiliary input.

[0038] As the system is playing back the selected track, music manager 242 preferably provides playback status of the track. Media cataloger 248 is configured to perform the meta data extraction, which includes extracting data such as song title, album title, artist, genre, time length, track number, the elapsed time, the recording label, featured artist(s), etc.

[0039] State manager 222 is configured to provide the playback status to GUI 212 and VUI 232. GUI 212 can show the playback status of the current track in the list of the selected tracks. VUI 232 also allows the user to “bargue in,” to the playing track and input a desired command. For example, a user may press a button, in order to activate speech recognition engine 234, and, then, the user may utter the appropriate command to begin playback of another media file (e.g., “next,” “pause,” “play ‘Yellow Submarine’”).

[0040] Figure 4 illustrates a method for retrieving media file data from a media player in response to a user command in accordance with an embodiment. Media file data 128 from media player 114 is retrieved in response to a command from a user in the vehicle. The command may be received via a user interface 110 (e.g., a button or switch) or as an oral command via audio input device 108. After a wireless communications link is established between control module 102 and media player 114, a user may provide a command to retrieve media file data from the media player. At step 302, control module 102 (shown in Figure 2) receives a command from a user to retrieve media file data from the connected media player. As mentioned, a user may provide a request via user interface 110 (Figure 2), for example, using a button or switch. Alternatively, the user may provide an oral command to request retrieval of the meta data. At step 304, control module 102 sends a request to the media player 114 via the wireless communications link and retrieves the meta data from the media player 114. Once the media file data are transferred to control module 102, the data are captured in a file at step 306. The retrieved media files are then processed at step 308 before being stored in memory 124. After processing, the media files are stored in memory 124 (shown in Figure 2). As mentioned previously, the media file data may be stored in volatile memory 136 (shown in Figure 2). Accordingly, the media file data are not preserved across power cycles of the control system. Preferably, the media file data may be stored in non-volatile memory 138 (shown in Figure 2) and may be maintained across power cycles and available to a user after power-on initialization. As mentioned above, when a wireless communications link is established with a different media player (e.g., the media player of another vehicle occupant), the data in volatile memory 136 is overwritten with the media file data from the new media player.

[0041] Figure 5 illustrates a method for processing media file data from a media player. At step 402, control module 102 (shown in Figure 2) accesses the media files and identifies the individual entries of the meta data such as song title, album title, artist, genre, time

length, track number, the elapsed time, recording label, featured artist, etc. An acoustic baseform may be generated for the text data of each entry of a media file using text-to-grammar engine 134 (shown in Figure 2). The acoustic baseform is a phonemic representation that may be used by the speech recognition engine 130 (shown in Figure 2) to recognize a user's utterance and, therefore, allow speech recognition to be performed against each entry of a media file. At step 404, an acoustic baseform is generated based on the text data of the media file. At step 406, the acoustic baseforms are stored in a baseform list that is associated with the media files. Accordingly, speech recognition may be performed against the text data of a media file. Returning to Figure 4, once the media file data are stored in memory in step 310, the media files are available to a user for speech-enabled manipulation of a media file.

[0042] As mentioned, a user may manipulate (e.g., play, pause, stop, next, previous, fast-forward, rewind, etc.) a media file using an entry of the meta data stored in memory. Figure 6 illustrates a method for speech-enabled playback using memory 124 in accordance with an embodiment. At step 502, a control module 102 (shown in Figure 2) receives a command to initiate playback of a media file from a user via audio input device 108. In one embodiment, control module 102 prompts the user to indicate which media file to playback, at step 504. The prompt may be played via audio output devices 126 or the audio system 104. Alternatively, the user may provide (i.e., utter) the entry name with an oral command to play a song (e.g., "play <track title>"). After the system receives the track title of the desired media file at step 506, the system identifies the desired entry at step 508. If there are similar entries in the memory 124, the speech recognition engine 130 selects the best match. At step 510, the speech recognition engine determines if there is more than one entry associated with an acoustic baseform. For example, if the user commands the system to "play The White Album," the system may be configured to recognize that more than one media file is associated with that album title. The control module 102 may be configured to list (via an audible prompt or via display 106) the associated media files (e.g., songs or tracks) for that album name. Alternatively, control module 102 can also be configured to just begin playback of the media files (i.e., tracks) of the album from the beginning of album or from a pre-selected track number. The user may then select a media file at step 512. Once a media file has been determined, control module 102 prompts the user to confirm the track title of the media file at step 514. After confirmation is received at step 516, control module 102 begins playback of the media file. Alternatively, the system may

be configured to determine that there is more than one album with a particular phrase in the album title (e.g., "Greatest Hits"). In this case, control module 102 is configured to recognize that more than one album title includes the uttered phrase, based on a confidence scoring algorithm. Speech recognition engine 130 provides the best matches, and the user is then prompted to select the desired album from among the matches.

[0043] Control module 102 (Figure 2) is also configured to provide a user with a listing of the meta data of the available media files upon request (e.g., "list albums," "list artists," etc.). Figure 7 illustrates a method for audibly listing media files, in accordance with an embodiment. At step 602, control module 102 receives a command from the user via a audio input device 108 (Figure 2) to list the contents of database 158 by meta data entry type. In response to the command, control module 102 audibly list the meta data entries of each media file at step 604. For example, if the user's command is "list albums," each album title would be listed. The text (e.g., album title) of each media file is converted to an audio representation using text-to-speech engine 132 and the audio may be played in the vehicle using audio output devices 126 or audio system 104. Preferably, the text of the entries is displayed in alphanumeric order sorted character by character on display 106. A predetermined time period elapses between the playing of each entry. In addition, control module 102 is configured to allow the user to interrupt the audible listing of the selected meta data by permitting the user to "barge-in." A user may barge in during the playing of an entry text or during a predetermined time after the playing of the entry in order to execute a command. If the user provides a command (i.e., barges in) at step 606, the control module 102 recognizes and performs the requested action at step 608. For example, a user may provide a command to skip to the next entry in the list of selected meta data or to skip to the entry beginning at the next alphanumeric character. A user may also provide commands to select a desired media file once it is heard or to repeat a previous entry. If a media file is selected, a user may then initiate an action (e.g., playback of a media file). When the last media file is reached (or when a file is selected) at step 610, the process ends. If the last entry has not been reached, the process returns to step 604 and continues until either an entry is selected or the last entry in the database 158 is reached.

[0044] While the exemplary embodiments illustrated in the figures and described above are presently preferred, it should be understood that these embodiments are offered by way of example only. Accordingly, the present invention is not limited to a particular

embodiment, but extends to various modifications that nevertheless fall within the scope of the appended claims. The order or sequence of any processes or method steps may be varied or re-sequenced according to alternative embodiments.

[0045] Describing the invention with figures should not be construed as imposing on the invention any limitations that may be present in the figures. The present invention contemplates methods, systems and program products on any machine-readable media for accomplishing its operations. The embodiments of the present invention may be implemented using an existing computer processors, or by a special purpose computer processor for an appropriate vehicle system, incorporated for this or another purpose or by a hardwired system.

[0046] As noted above, embodiments within the scope of the present invention include program products comprising machine-readable media for carrying or having machine-executable instructions or data structures stored thereon. Such machine-readable media can be any available media which can be accessed by a general purpose or special purpose computer or other machine with a processor. By way of example, such machine-readable media can comprise RAM, ROM, EPROM, EEPROM, CD-ROM or other optical disk storage, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to carry or store desired program code in the form of machine-executable instructions or data structures and which can be accessed by a general purpose or special purpose computer or other machine with a processor. When information is transferred or provided over a network or another communications connection (either hardwired, wireless, or a combination of hardwired or wireless) to a machine, the machine properly views the connection as a machine-readable medium. Thus, any such connection is properly termed a machine-readable medium. Combinations of the above are also included within the scope of machine-readable media. Machine-executable instructions comprise, for example, instructions and data which cause a general purpose computer, special purpose computer, or special purpose processing machines to perform a certain function or group of functions.

[0047] It should be noted that although the diagrams herein may show a specific order of method steps, it is understood that the order of these steps may differ from what is depicted. Also two or more steps may be performed concurrently or with partial concurrence. Such variation will depend on the software and hardware systems chosen and on designer choice.

It is understood that all such variations are within the scope of the invention. Likewise, software implementations of the present invention could be accomplished with standard programming techniques with rule based logic and other logic to accomplish the various connection steps, processing steps, comparison steps and decision steps.

[0048] The foregoing description of embodiments of the invention has been presented for purposes of illustration and description. It is not intended to be exhaustive or to limit the invention to the precise form disclosed, and modifications and variations are possible in light of the above teachings or may be acquired from practice of the invention. The embodiments were chosen and described in order to explain the principals of the invention and its practical application to enable one skilled in the art to utilize the invention in various embodiments and with various modifications as are suited to the particular use contemplated.

WHAT IS CLAIMED IS:

- 1 1. A method for extracting meta data from a digital media storage device in a
2 vehicle over a communication link between a control module of the vehicle and the digital
3 media storage device, the method comprising:
4 establishing a communication link between the control module of the vehicle
5 and the digital media storage device;
6 identifying a media file on the digital media storage device;
7 retrieving meta data from a media file, the meta data including a plurality of
8 entries, wherein at least one of the plurality of entries includes text data;
9 identifying the text data in an entry of the media file; and
10 storing the plurality of entries in a memory.
- 1 2. A method according to claim 1, wherein the meta data are retrieved in
2 response to establishing the communication link between the control module and the digital
3 media storage device.
- 1 3. A method according to claim 1, wherein the meta data are retrieved in
2 response to a command by a user.
- 1 4. A method according to claim 3, wherein the command from the user is
2 received via a user interface.
- 1 5. A method according to claim 3, wherein the command is a spoken command.
- 1 6. A method according to claim 1, wherein the communication link between the
2 control module and the digital media storage device is wireless.
- 1 7. A method according to claim 1, wherein the plurality of entries includes any
2 of the following: song title, album title, artist, genre, time length, track number, elapsed
3 time, recording label, recording date, composer, producer, and featured artist.
- 1 8. A method according to claim 1, further comprising generating an acoustic
2 baseform of the text data for an entry; and
3 storing the acoustic baseform for the entry in a baseform list.

1 9. A method according to claim 1, further comprising identifying a playlist on
2 the digital media storage device; and

3 retrieving the playlist from the digital media storage device, wherein the
4 playlist includes a custom arrangement of media files.

1 10. A method according to claim 1, wherein the acoustic baseforms are
2 configured to allow each entry to be speech recognized.

1 11. A control system in a vehicle for extracting meta data from a digital media
2 storage device over a communication link, the system comprising:

3 a communication module for establishing a communication link with the
4 digital media storage device;

5 a processing module coupled to the communication module, the processing
6 module being configured to retrieve via the communication module meta data associated
7 with a media file from the digital media storage device, the meta data including a plurality
8 of entries, wherein at least one of the plurality of entries includes text data; and

9 a memory module configured to store the plurality of entries retrieved from
10 the digital media storage device.

1 12. A system according to claim 11, wherein the processing module is
2 configured to identify the text data of an entry and generate a phonemic representation of
3 the text data of the entry.

1 13. A system according to claim 11, further comprising a user interface
2 configured to receive input, to select at least one media file based on the media file's
3 associated meta data, and to play the media file's audio data via an audio system.

1 14. A system according to claim 11, wherein the user interface comprises a
2 display and the processing module is configured to provide meta data to a user via the
3 display.

1 15. A system according to claim 11, further comprising an audio input device
2 configured to receive oral input commands from a user.

1 16. A system according to claim 11, further comprising a speech recognition
2 module configured to compare oral input commands received from a user to a
3 predetermined set of input commands, wherein the speech recognition module determines
4 an appropriate response based on the oral input command.

1 17. A system according to claim 11, further comprising a text conversion engine
2 configured to convert the text data of an entry to an audible speech representation of the text
3 data.

1 18. A system according to claim 11, wherein the memory module comprises a
2 volatile memory.

1 19. A system according to claim 11, wherein the memory module comprises a
2 non-volatile memory.

1 20. A system according to claim 11, wherein the phonemic representation of the
2 text data of an entry is configured to facilitate speech recognition of the entry of the media
3 file.

1 21. A system according to claim 11, wherein the system is configured to support
2 media encoding/decoding algorithms, including any of the following: MP3, MP3 Pro, Ogg
3 Vorbis, Advanced Audio Coding (AAC), Windows Media Audio (WMA), and Waveform
4 Audio (WAV).

1 / 5

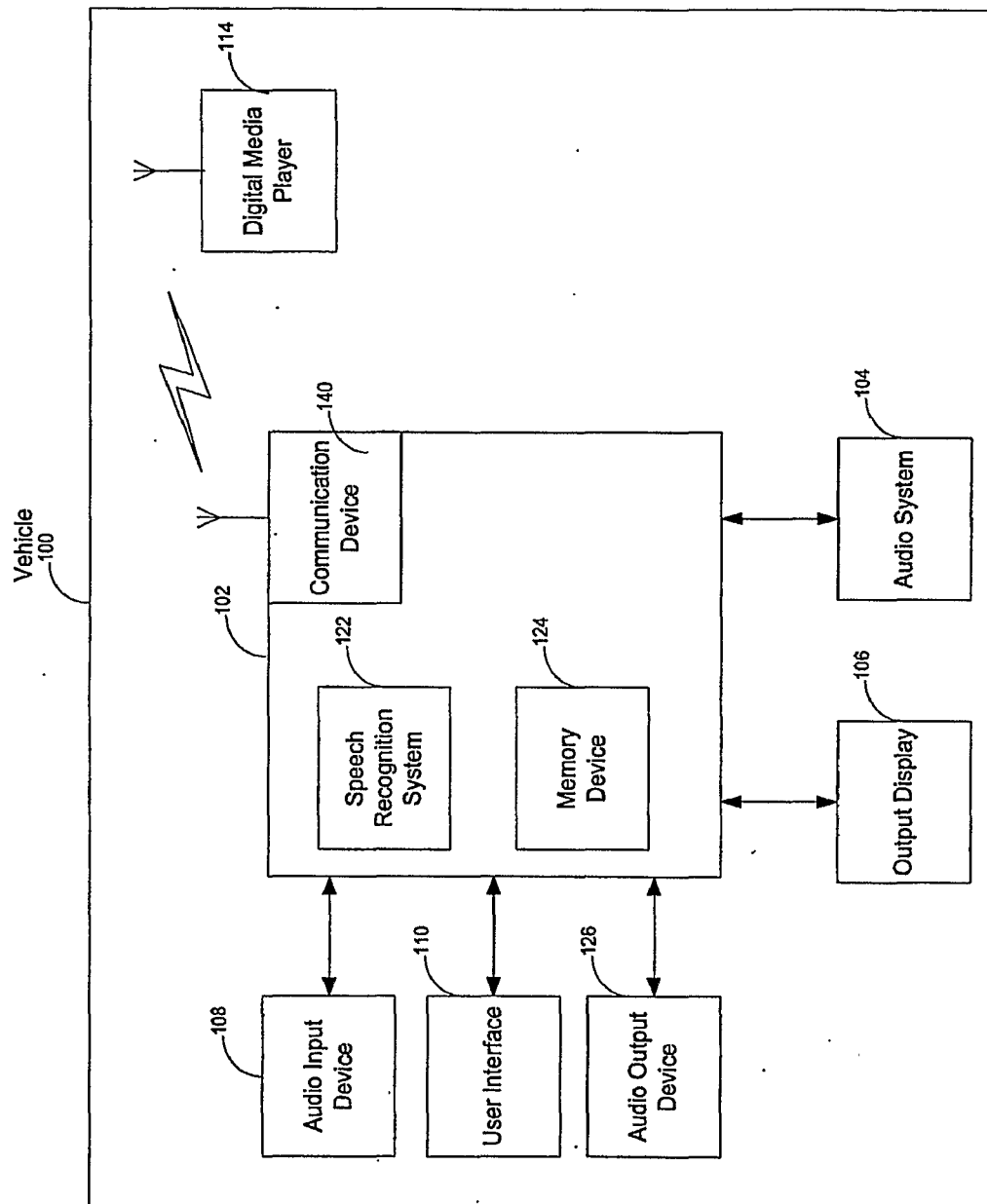


FIG. 1

2 / 5

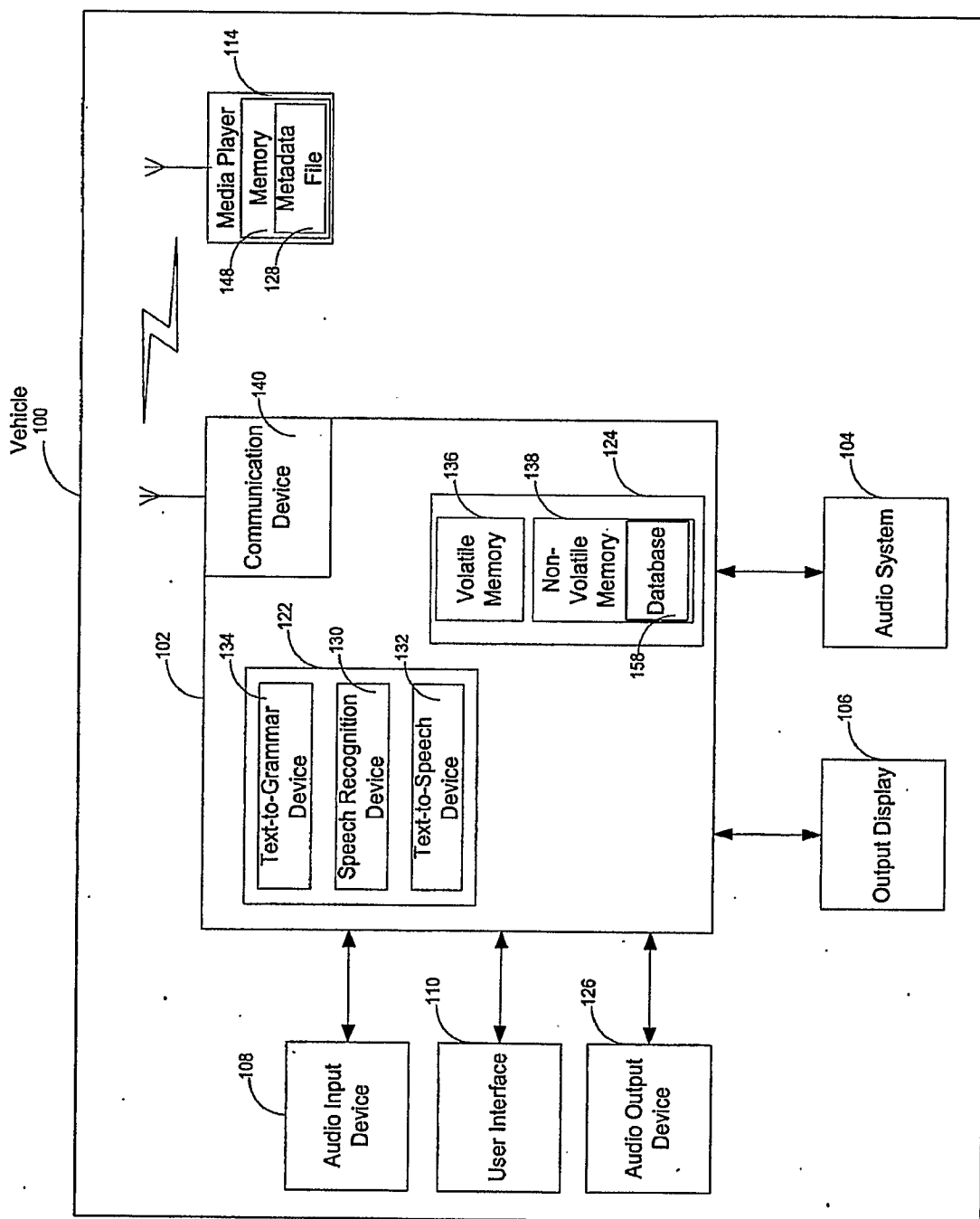


FIG. 2

3 / 5

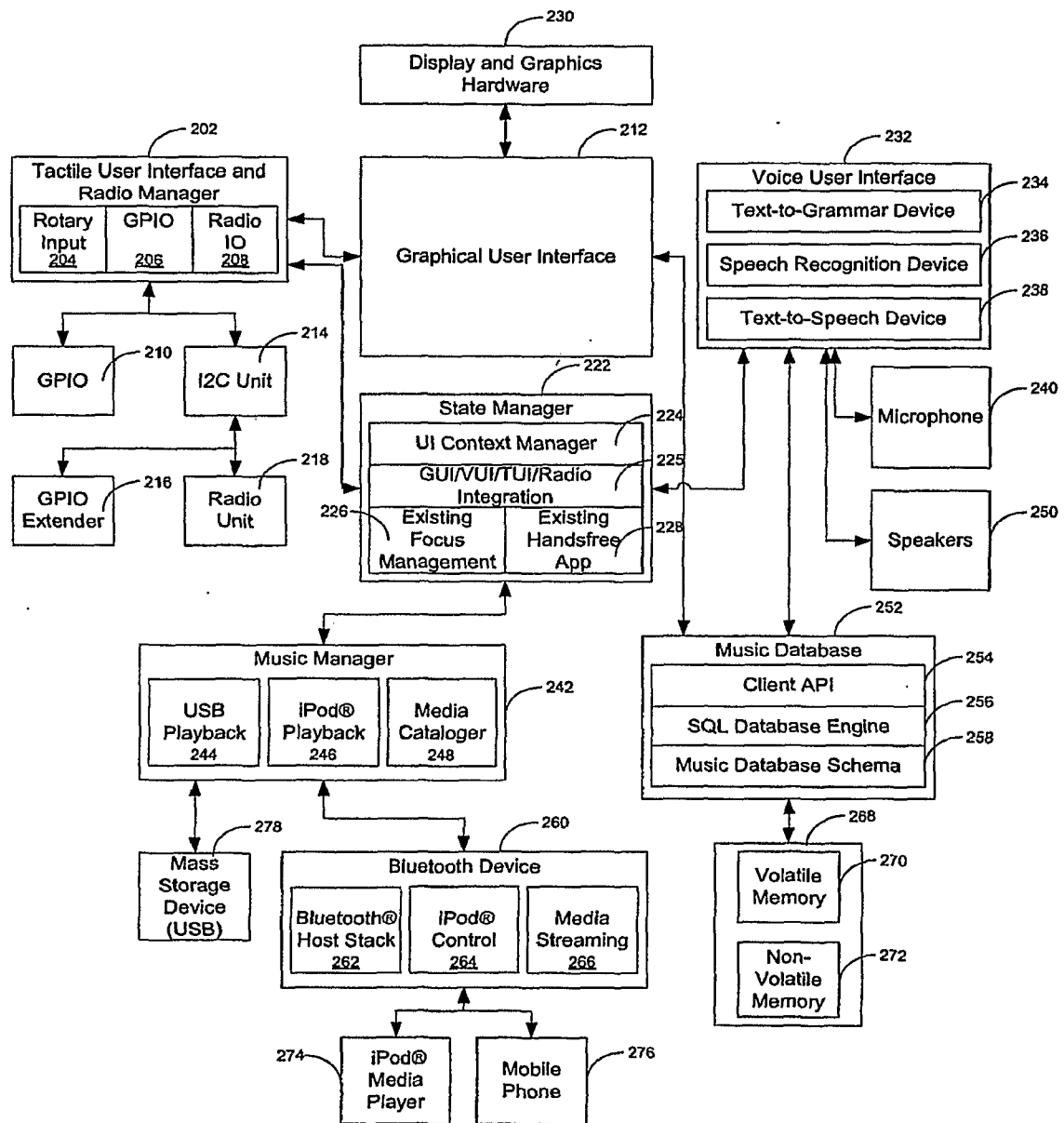


FIG. 3

4 / 5

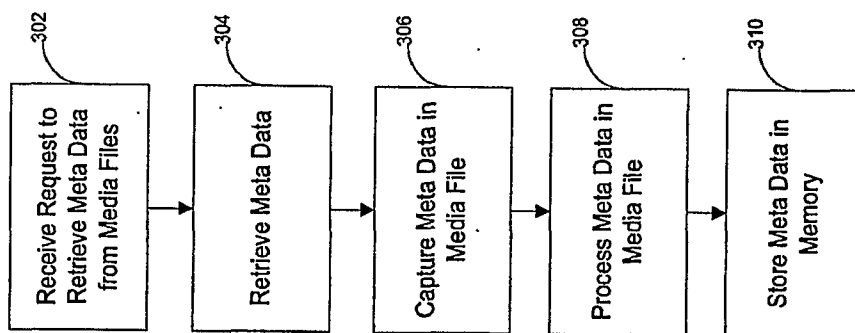


FIG. 4

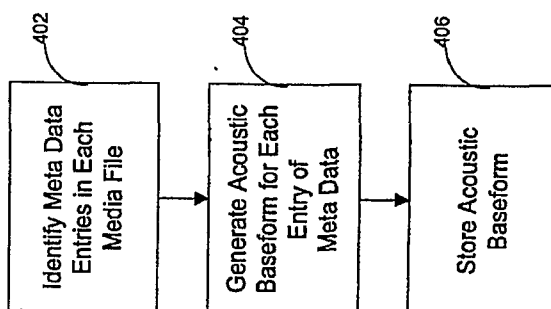
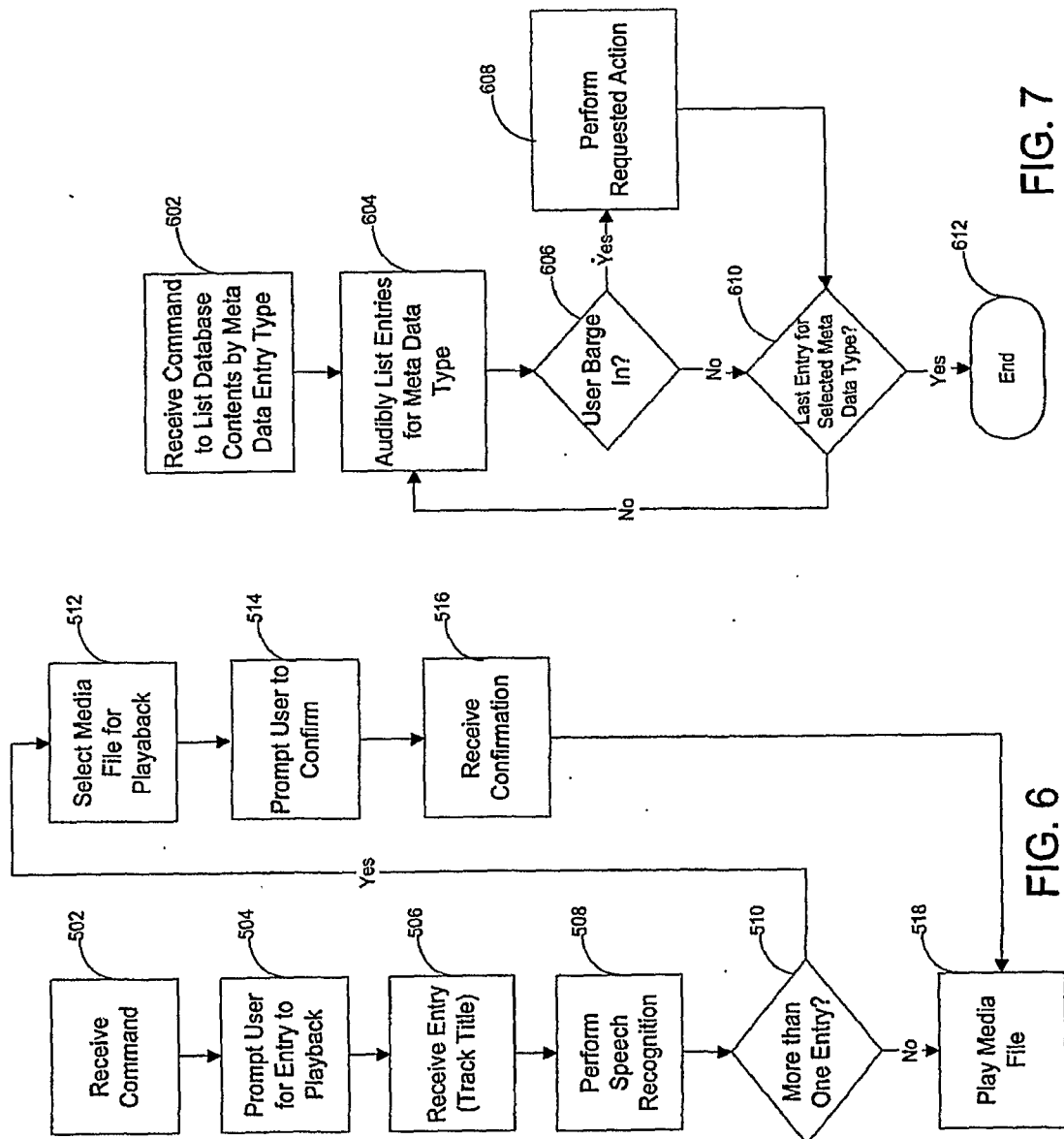


FIG. 5

5 / 5



INTERNATIONAL SEARCH REPORT

International application No

PCT/US2007/008063

A. CLASSIFICATION OF SUBJECT MATTER
 INV. G06F17/00 H04L29/00

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)
 H04L G06F

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

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

EPO-Internal, PAJ, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/194611 A1 (KAWANA YUTA [JP]) 7 October 2004 (2004-10-07) paragraphs [0004], [0005] paragraphs [0041] - [0047] paragraph [0013] paragraphs [0025] - [0039] paragraphs [0051], [0052] paragraphs [0057], [0058] -----	1-21
A	US 2005/038813 A1 (APPARAO VIDUR [US] ET AL) 17 February 2005 (2005-02-17) the whole document -----	1-21

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

17 September 2007

Date of mailing of the international search report

25/09/2007

Name and mailing address of the ISA/

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

Authorized officer

Pereira, Mafalda

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2007/008063

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004194611 A1	07-10-2004	DE 102004016319 A1 JP 2004309795 A	04-11-2004 04-11-2004
US 2005038813 A1	17-02-2005	US 2007198563 A1 WO 2005019985 A2	23-08-2007 03-03-2005