



## (51) International Patent Classification:

H03F 1/26 (2006.01)

## (21) International Application Number:

PCT/US20 12/07 1966

## (22) International Filing Date:

28 December 2012 (28. 12.2012)

## (25) Filing Language:

English

## (26) Publication Language:

English

## (30) Priority Data:

13/338,588 28 December 2011 (28. 12.2011) U S

## (71) Applicant: ARBITRON, INC. [US/US]; 9705 Patuxent

Woods Drive, Columbia, Maryland 21046 (US).

## (72) Inventors: LYNCH, Wendell; 907 Southlawn Avenue, East Lansing, Michigan 48823 (US). TENBROCK, Michael; 7244 Dockside Lane, Columbia, Maryland 21045 (US). NEUHAUSER, Alan; 1512 Flora Court, Silver Spring, Maryland 20910 (US). JAIN, Anand; 3253 Halcyon Court, Ellicott City, Maryland 21143 (US). KRUG,

William K.; 145 N. Hallifax Avenue 702, Daytona Beach, Florida 32118 (US).

## (74) Agents: ZURA, Peter et al; Barnes &amp; Thornburg LLP, P.O. Box 2786, Chicago, Illinois 60690-2786 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, 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, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK,

[Continued on nextpage]

(54) Title: SYSTEM AND METHOD FOR UTILIZING AUDIO ENCODING FOR MEASURING MEDIA EXPOSURE WITH ENVIRONMENTAL MASKING

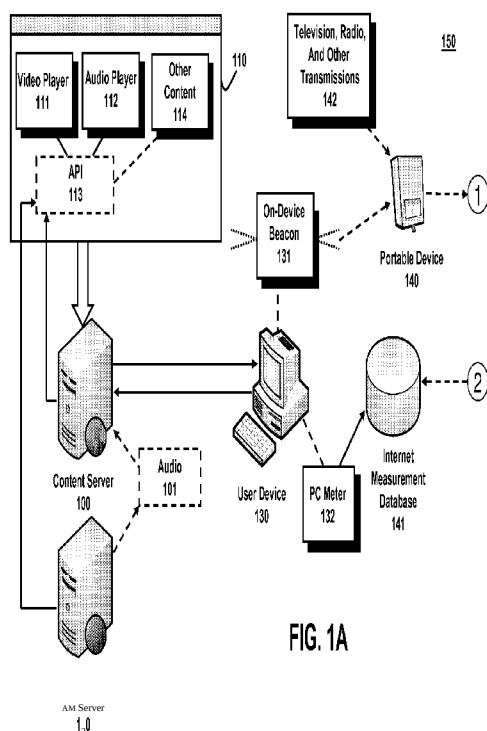


FIG. 1A

(57) Abstract: An audio beacon system, apparatus and method for collecting information on a panelist's exposure to media. An audio beacon is configured as on-device encoding technology that is operative in a processing device (e.g., cell phone, PDA, PC) to enable the device to encode an environmental sound and transmit it for a predetermined period of time. The acoustically transmitted data is received and processed by a portable audience measurement device, such as Arbitron's Personal People Meter™ ("PPM"), or other specially equipped portable device to enable audience measurement systems to achieve higher levels of detail on panel member activity and greater association of measurement devices to their respective users.

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**  
— with international search report (Art. 21(3))

- 1 -

**SYSTEM AND METHOD FOR UTILIZING AUDIO ENCODING FOR  
MEASURING MEDIA EXPOSURE WITH ENVIRONMENTAL MASKING**

**RELATED APPLICATIONS**

[0001] The present application is a continuation-in-part of U.S. Patent Application No. 12/425,464, titled "System and Method for Utilizing Audio Beacons in Audience Measurement" filed April 17, 2009, and U.S. Patent Application No. 12/425,556, titled "System and Method for Utilizing Supplemental Audio Beacons in Audience Measurement," also filed on April 17, 2009. Both applications are assigned to the assignee of the present application and are incorporated by reference in their entireties herein

**TECHNICAL FIELD**

[0002] The present disclosure relates to systems and processes for communicating and processing data, and, more specifically, to communicate media data exposure that may include coding that provides media and/or market research.

**BACKGROUND INFORMATION**

[0003] The use of global distribution systems such as the Internet for distribution of digital assets such as music, film, computer programs, pictures, games and other content continues to grow. In many instances, media offered via traditional broadcast mediums is supplemented through similar media offerings through computer networks and the Internet. It is estimated that Internet-related media offerings will rival and even surpass traditional broadcast offerings in the coming years.

[0004] Techniques such as "watermarking" have been known in the art for incorporating information signals into media signals or executable code. Typical watermarks may include encoded indications of authorship, content, lineage, existence of copyright, or the like. Alternatively, other information may be

- 2 -

incorporated into audio signals, either concerning the signal itself, or unrelated to it. The information may be incorporated in an audio signal for various purposes, such as identification or as an address or command, whether or not related to the signal itself.

[0005] There is considerable interest in encoding audio signals with information to produce encoded audio signals having substantially the same perceptible characteristics as the original unencoded audio signals. Recent successful techniques exploit the psychoacoustic masking effect of the human auditory system whereby certain sounds are humanly imperceptible when received along with other sounds.

[0006] Arbitron has developed a new and innovative technology called Critical Band Encoding Technology (CBET) that encompasses all forms of audio and video broadcasts in the measurement of audience participation. This technology dramatically increases the both the accuracy of the measurement and the quantity of useable and effective data across all types of signal broadcasts. CBET is an encoding technique that Arbitron developed and that embeds identifying information (ID code) or other information within the audio portion of a broadcast. An audio signal is broadcast within the actual audio signal of the program, in a manner that makes the ID code inaudible, to all locations the program is broadcast, for example, a car radio, home stereo, computer network, television, etc. This embedded audio signal or ID code is then picked up by small (pager-size) specially designed receiving stations called Portable People Meters (PPM), which capture the encoded identifying signal, and store the information along with a time stamp in memory for retrieval at a later time. A microphone contained within the PPM receives the audio signal, which contains within it the ID code.

[0007] Further disclosures related to CBET encoding may be found in U.S. Pat. No. 5,450,490 and U.S. Pat. No. 5,764,763 {*Jensen et al.*} in which information is represented by a multiple-frequency code signal which is incorporated into an audio signal based upon the masking ability of the audio signal. Additional examples include U.S. Patent No. 6,871,180 (*Neuhauser et al.*) and U.S. patent 6,845,360 {*Jensen et al.*}, where numerous messages represented by multiple

- 3 -

frequency code signals are incorporated to produce and encoded audio signal. Other examples include U.S. patent no. 7,239,981 (Kolessar et al.). Each of the above-mentioned patents is incorporated by reference in its entirety herein.

[0008] The encoded audio signal described above is suitable for broadcast transmission and reception and may be adapted for Internet transmission, reception, recording and reproduction. When received, the audio signal is processed to detect the presence of the multiple-frequency code signal. Sometimes, only a portion of the multiple-frequency code signal, e.g., a number of single frequency code components, inserted into the original audio signal, is detected in the received audio signal. However, if a sufficient quantity of code components is detected, the information signal itself may be recovered.

[0009] Other means of watermarking have been used in various forms to track multimedia over computer networks and to detect if a user is authorized to access and play the multimedia. For certain digital media, metadata is transmitted along with media signals. This metadata can be used to carry one or more identifiers that are mapped to metadata or actions. The metadata can be encoded at the time of broadcast or prior to broadcasting. Decoding of the identifier may be performed at a digital receiver. Other means of watermarking include the combination of digital watermarking with various encryption techniques known in the art.

[00010] While various encoding and watermarking techniques have been used to track and protect digital data, there have been insufficient advances in the fields of cross-platform digital media monitoring. Specifically, in cases where a person's exposure to Internet digital media is monitored in addition to exposure to other forms of digital media (e.g., radio, television, etc.), conventional watermarking systems have shown themselves unable to effectively monitor and track media exposure. Furthermore, there is a need to integrate exposure to digital media across platforms where the digital media includes formats that are not traditionally subject to audio encoding. Moreover, there is a need in the art to properly "mask" such signals using environmental sounds and/or sounds native to a device that is conducting beaconing processes.

- 4 -

### **SUMMARY**

[0001 1] Accordingly, an audio beacon system, apparatus and method is disclosed for collecting information on a panelist's exposure to media. Under a preferred embodiment, the audio beacon is configured as on-device encoding technology that is operative in a panelist's processing device (e.g., cell phone, PDA, PC) to enable the device to acoustically transmit user/panelist data for a predetermined period of time. The acoustically transmitted data is received and processed by a portable audience measurement device, such as Arbitron's Personal People Meter™ ("PPM") or specially equipped cell phone, laptop etc., to enable audience measurement systems to achieve higher levels of detail on panel member activity and greater association of measurement devices to their respective panelists. Additionally, the acoustic transmissions are configured to utilize environmental sounds that are advantageous in being less obtrusive to users.

[00012] Additional features and advantages of the various aspects of the present disclosure will become apparent from the following description of the preferred embodiments, which description should be taken in conjunction with the accompanying drawings.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

[00013] FIG. 1A is a block diagram illustrating a portion of an audio beaconing system under one exemplary embodiment;

[00014] FIG. 1B is a block diagram illustrating another portion of an audio beaconing system under the embodiment illustrated in Figure 1A;

[00015] FIG. 2 is a tabular illustration of an audio beaconing and audio matching process under another exemplary embodiment;

[00016] FIG. 3 illustrates a block diagram of a server-side encoding process under yet another exemplary embodiment;

- 5 -

[00017] FIG. 4 illustrates an exemplary watermarking process for a digital media file suitable for use in the embodiment of FIGs. 1A-B;

[00018] FIG. 5 illustrates a block diagram of a client-side encoding process under yet another exemplary embodiment;

[00019] FIG. 6 illustrates an exemplary audio waveform having encoded data therein; and

[00020] FIG. 7 illustrates an exemplary device configured to select environmental sounds for audio transmission.

### **DETAILED DESCRIPTION**

[00021] FIG. 1A is an exemplary block diagram illustrating a portion of an audio beaconing system 150 under one embodiment, where a web page 110 is provided by a page developer and published on content server 100. The web page preferably contains an embedded video player 111 and audio player 112 (that is preferably not visible), together with an application programming interface (API) 113. Other content 114 (e.g., HTML, text, etc.) is also provided on web page 110, which may or may not be coupled through API 113. API 113 is preferably embodied as a set of routines, data structures, object classes and/or protocols provided by libraries and/or operating system services in order to support the video player 111 and audio player 112. Additionally, the API 113 may be language-dependent (i.e. available only in a particular programming language) or language-independent (i.e., can be called from several programming languages, preferably an assembly/C-level interface). Examples of suitable API's include Windows API, Java Platform API, OpenGL, DirectX, Simple DirectMedia Layer (SDL), YouTube API, Facebook API and iPhone API, among others.

[00022] In one preferred embodiment, API 113 is configured as a beaconing API object. Depending on the features desired, the API object may reside on an Audience Measurement (AM) server 120, so that the object may be remotely

- 6 -

initialized, thus minimizing the objects software's exposure to possible tampering and to maintain security. Alternately, the API object can reside on the content server 100, where the API object may be initialized under increased performance conditions.

[00023] When initialized, API 113 can communicate the following properties: (1) the URL of the page playing the media, (2) URL of the media being served on the page, (3) any statically available media metadata, and (3) a timestamp. It is understood that additional properties may be communicated in API 113 as well. In one configuration of FIG. 1A, an initialization request is received by API 113, to create a code tone that is preferably unique for each website and encode it on a small inaudible audio stream. Alternatively, the AM server 120 could generate a pre-encoded audio clip 101, with a code tone, for each site and forward it on the content server 100 in advance.

[00024] The encoded audio stream would then travel from content server 100 to the web page 110 holding audio player 110. In a preferred embodiment, audio player 110 may be set by the page developer as an object instance, where the visible property of player 110 is oriented as "false" or set to a one-by-one dimension in order to minimize the visual interference of the audio player with the web page. The encoded audio stream may then be played out in parallel with the media content being received from the web page 110. The encoded audio stream would preferably repeat at predetermined time periods through an on-device beacon 131 resident on a user device 130 as long as the user is on the same website. The beacon 131, would enable device 130 to acoustically transmit the encoded audio stream so that a suitably configured portable device 140 (e.g., PPM) can receive and process the encoded information. Beacon 131 could be embedded into an audio player resident on a web page being viewed inside the browser on user device 130, or may be a stand-alone application on user device 130.

[00025] A simplified example further illustrates the operation of the system 150 of FIGs. 1A-B under an alternate embodiment. User device 130 requests content (e.g., <http://www.hulu.com/>) from server 100. When the content is received in user device 130, PC meter software 132 collects and transmits web measurement data to



- 7 -

Internet measurement database 141. One example of a PC meter is comScore's Media Metrix™ software; further exemplary processes of web metering may be found in U.S. Patent 7,493,655, titled "Systems for and methods of placing user identification in the header of data packets usable in user demographic reporting and collecting usage data" and U.S. Patent No. 7,260,837, titled "Systems and methods for user identification, user demographic reporting and collecting usage data usage biometrics", both of which are incorporated by reference in their entirety herein.

[00026] As web measurement data is collected by PC meter 132, beacon 131 acoustically transmits encoded audio, which is received by portable device 140. In the exemplary embodiment, the encoding for the beacon transmission may include data such as a timestamp, portable device ID, user device ID, household ID, or any similar information. In addition to the beacon data, portable device 140 additionally receives multimedia data such as television and radio transmissions 142, which may or may not be encoded, at different times. If encoded (e.g., CBET encoding), portable device can forward transmissions 142 to audio matching server 160 (FIG. IB) for decoding and matching with audio matching database 161. If transmissions 142 are not encoded, portable device 140 may employ sampling techniques for creating audio patterns or signatures, which may also be transmitted to audio matching server 160 for pattern matching using techniques known in the art.

[00027] Audio beacon server 150, shown in FIG. IB, receives and processes/decodes beacon data from portable device 140. Under an alternate embodiment, it is possible to combine audio matching server 160 and audio beacon server 150 to collectively process both types of data. Data from Audio beacon server 150 and audio matching server 160 is transmitted to Internet measurement database 141, where the web measurement data could be combined with audio beacon data and data from the audio matching server to provide a comprehensive collection of panelist media exposure data.

[00028] Under another exemplary embodiment, the video and audio players of webpage 110 are configured to operate as Flash Video, which is a file format used to deliver video over the Internet using Adobe™ Flash Player. The Flash Player

- 8 -

typically executes Shockwave Flash "SWF" files and has support for a scripting language called ActionScript, which can be used to display Flash Video from an SWF file. Because the Flash Player runs as a browser plug-in, it is possible to embed Flash Video in web pages and view the video within a web browser. Commonly, Flash Video files contain video bit streams which are a variant of the H.263 video standard, and include support for H.264 video standard (i.e., "MPEG-4 part 10", or "AVC"). Audio in Flash Video files ("FLV") is usually encoded as MP3, but can also accommodate uncompressed audio or ADPCM format audio.

[00029] Continuing with the embodiment, video beacons can be embedded within an action script that will be running within the video Flash Player's run time environment on web page 110. When an action script associated with web page 110 gets loaded as a result of the access to the page, the script gets activated and triggers a "video beacon", which extracts and store URL information on a server (e.g., content server 100), and launches the video Flash Player. By inserting an audio beacon in the same action script, the audio beacon will be triggered by the video player. Once triggered, the audio beacon may access AM server 120 to load a pre-recorded audio file containing a special embedded compatible code (e.g., CBET). This pre-recorded audio file would be utilized for beacon 131 to transmit for a given period of time (e.g., every  $x$  seconds).

[00030] As a result, the beacon 131 audio player runs as a "shadow player" in parallel to the video Flash Player. If a portable device 140 is in proximity to user device 130, portable device 140 will detect the code and reports it to audio beacon server 150. Depending on the level of cooperation between the audio and video beacon, the URL information can also be deposited onto beacon server 150 along with codes that would allow an audience measurement entity to correlate and/or calibrate various measurements with demographic data.

[00031] Under the present disclosure, media data may be processed in a myriad of ways for conducting customized panel research. As an example, each user device 130 may install on-device measurement software (PC meter 132) which includes one or more web activity monitoring applications, as well as beacon software

- 9 -

131. It is understood that the web activity monitoring application and the beacon software may be individual applications, or may be merged into a single application.

[00032] The web activity monitoring application collects web activities data from the user device 130 (e.g., site ID, video page URL, video file URL, start and end timestamp and any additional metadata about videosite information, URL information, time, etc.) and additionally assigns a unique ID, such as a globally unique identifier or "GUID", to each device. For the beacon 131, a unique composite ID may be assigned including a household ID ("HHID") and a unique user device ID for each device in the household (e.g., up to 10 devices for a family), as well as a portable device ID (PPMID). Panelist demographic data may be included for each web activity on the device.

[00033] Continuing with the example, beacon 131 emits an audio beacon code (ABC) for device in the household by encoding an assigned device ID number and acoustically sending it to portable device 140 to identify the device. Further details on the encoding is provided below. Portable device 140 collects the device ID and sends it to a database along with HHID and/or PPM ID and the timestamp. Preferably, a PPMID is always mapped to a HHID in the backend; alternately an HHID can be set within each PPMID.

[00034] The web activity monitoring and beacon applications may pass information to each other as needed. Both can upload information to a designated server for additional processing. A directory of panelists' devices is built to contain the GUID, HHID, and device ID for panel, and the directory could be used to correlate panelist demographic data and web measurement data.

[00035] Turning to FIG. 2, a tabular illustration of an audio beaconing and audio matching process under another exemplary embodiment is provided. Specifically, the table illustrates a combination of audio beaconing and audio matching and its application to track a video on a content site, such as Hulu.com. FIG. 2. Timeline 200 shows in sections a scenario where a user/panelist plays a ten minute video on Hulu.com. Activities 201 shows actions taken in user system 150

- 10 -

where a video is loaded in the user device 130, and played. At the 5 minute mark (301 sec.), a 15 second advertisement is served. At the conclusion of the advertisement (316 sec.), the video continues to play until its conclusion (600 sec).

[00036] During this time, audio beacon activities 202 are illustrated, where, under one embodiment, on-device beacon 131 transmits continuous audio representing the website (Hulu.com). In addition, beacon also transmits a timestamp, portable device ID, user device ID, household ID and/or any other data in accordance with the techniques described above. Under an alternate embodiment shown in 203, additional data may be transmitted in the beacon to include URLs and video ID's when a video is loaded and played. As the advertisement is served, an event beacon, which may include advertisement URL data, is transmitted. At the conclusion of the video, a video end beacon is transmitted to indicate the user/panelist is no longer viewing specific media.

[00037] When the video and advertisement is loaded and played, additional audio matching may occur in the portable device 140, in addition with audio matching processes explained above in relation to FIGs. 1A-B. Referring to audio matching events 204, portable device data 205 and end-user experience 206 of FIG. 2, portable device data (e.g., demographic ID data) is overlayed along with site information (URL, video ID, etc) when a video is loaded. When the video is played, audio signatures may be sampled periodically by portable device 140, until a content match is achieved. The audio signatures may be obtained through encoding, pattern matching, or any other suitable technique. When a match is found, portable device data is overlayed to indicate that a content match exists. Further signature samples are taken to ensure that the same content is being viewed. When an advertisement is served, the sampled signature will indicate that different content is being viewed, at which point the portable device data is overlayed in the system. When the video resumes, the audio signature indicates the same video is played, and portable device data is overlayed through the end of the video as shown in FIG. 2.

[00038] As explained above, signature sampling/audio matching allows the system 150 to identify and incorporate additional data on the users/panelists and the

- 11 -

content being viewed. Under a typical configuration, the content provider media (e.g., Hulu, Facebook, etc.) may be sampled in advance to establish respective signatures for content and stored in a matching database (e.g., audio matching server 160). The portable device 140 would be equipped with audio matching software, so that, when a panelist is in the vicinity of user device 130, audio matching techniques are used to collect the signature, or "audio fingerprint" for the incoming stream. The signatures would then be matched against the signatures in the matching database to identify the content.

[00039] It is understood by those skilled in the art however, that encoding techniques may also be employed to identify content data. Under such a configuration, content is encoded prior to transmission to include data relating to the content itself and the originating content site. Additionally, data relating to possible referral sites (e.g., Facebook, MySpace, etc.) may be included. Under one embodiment, a content management system may be arranged for content distributors to choose specific files for a corresponding referral site.

[00040] For the media data encoding, several advantageous and suitable techniques for encoding audience measurement data in audio data are disclosed in U.S. Pat. No. 5,764,763 to James M. Jensen, et al., which is assigned to the assignee of the present application, and which is incorporated by reference herein. Other appropriate encoding techniques are disclosed in U.S. Pat. No. 5,579,124 to Aijala, et al., U.S. Pat. Nos. 5,574,962, 5,581,800 and 5,787,334 to Fardeau, et al., U.S. Pat. No. 5,450,490 to Jensen, et al., and U.S. patent application Ser. No. 09/318,045, in the names of Neuhauser, et al., each of which is assigned to the assignee of the present application and all of which are incorporated by reference in their entirety herein.

[00041] Still other suitable encoding techniques are the subject of PCT Publication WO 00/04662 to Srinivasan, U.S. Pat. No. 5,319,735 to Preuss, et al., U.S. Pat. No. 6,175,627 to Petrovich, et al., U.S. Pat. No. 5,828,325 to Wolosewicz, et al., U.S. Pat. No. 6,154,484 to Lee, et al., U.S. Pat. No. 5,945,932 to Smith, et al., PCT Publication WO 99/59275 to Lu, et al., PCT Publication WO 98/26529 to Lu, et

- 12 -

al., and PCT Publication WO 96/27264 to Lu, et al, all of which are incorporated by reference in their entirety herein.

[00042] Variations on the encoding techniques described above are also possible. Under one embodiment, the encoder may be based on a Streaming Audio Encoding System (SAES) that operates under a set of sample rates and is integrated with media transcoding automation technology, such as Telestream's FlipFactory™ software. Also, the encoder may be embodied as a console mode application, written in a general-purpose computer programming language such as "C". Alternately, the encoder may be implemented as a Java Native Interface (JNI) to allow code running in a virtual machine to call and be called by native applications, where the JNI would include a JNI shared library for control using Java classes. The encoder payloads would be configured using specially written Java classes. Under this embodiment, the encoder would use the information hiding abstractions of an encoder payload which defines a single message. Under a preferred embodiment, the JNI encoder would operate using a 44.1kHz sample rate.

[00043] Examples of symbol configurations and message structures are provided below. One exemplary symbol configuration uses four data symbols and one end symbol defined for a total of five symbols. Each symbol may comprise five tones, with one tone coming from each of five standard Barks. One exemplary illustration of Bark scale edges (in Hertz), would be { 920, 1080, 1270, 1480, 1720, 2000 }. The bins are preferably spaced on a 4 X 3.90625 grid in order to provide lighter processing demands, particularly in cases using decoders based on 512 point fast Fourier transform (FFT). an exemplary bin structure is provided below:

**Symbol 0:** { 248, 292, 344, 400, 468 }  
**Symbol 1:** { 252, 296, 348, 404, 472 }  
**Symbol 2:** { 256, 300, 352, 408, 476 }  
**Symbol 3:** { 260, 304, 356, 412, 480 }  
**End Marker Symbol:** { 264, 308, 360, 416, 484 }

[00044] Regarding message structure, an exemplary message would comprise 20 symbols, each being 400 milliseconds in duration, for a total duration of 8 seconds. Under this embodiment, the first 3 symbols could be designated as

- 13 -

match/check criteria symbols, which could be the simple sum of the data symbols or could be derived from an error correction or cyclical redundancy check algorithm. The following 16 symbols would then be designated as data symbols, leaving the last symbol as an end symbol used for a marker. Under this configuration, the total number of possible symbols would be  $4^{16}$  or 4,294,967,296 symbols.

[00045] Variations in the algorithmic process for encoding are possible as well under the present disclosure. For example, a core sampling rate of 5.5125 kHz may be used instead of 8 kHz to allow down-sampling from 44.1 kHz to be efficiently performed without pre-filter (to eliminate aliasing components) followed by conversion filter to 48 kHz. Such a configuration should have no effect on code tone grid spacing since the output frequency generation is independent of the core sampling rate. Additionally, this configuration would limit the top end of the usable frequency span to about 2 kHz (as opposed to 3 kHz under conventional techniques) since frequency space should be left for filters with practical numbers of taps.

[00046] Under one embodiment, a 16 point overlap of a 256 point large FFT is used, resulting in amplitude updates every 2.9 milliseconds for encoding instead of every 2 milliseconds for standard CBET techniques. Accordingly, fewer large FFTs are calculated under a tighter bin resolution of 21.5 Hz instead of 31.25 Hz.

[00047] The psychoacoustic model calculations used for the encoding algorithm under the present disclosure may vary from traditional techniques as well. In one embodiment, bin spans of the clumps may be set by Bark boundaries instead of being wholly based on Critical Bandwidth criteria. By using Bark boundaries, a specific bin will not contribute to the encoding power level of multiple clumps, which provides less coupling between code amplitudes of adjacent clumps. When producing Equivalent Large FFTs, a comparison may be made of the most recent 16 point Small FFT results to a history of squared sums to simplify calculations.

[00048] For noise power computation, the encoding algorithm under the present disclosure would preferably use 3 bin values over a clump: the minimum bin

- 14 -

power (MIN), the maximum bin power (MAX), and the average bin power (AVG). Under this arrangement, the bin values could be modeled as follows:

IF (MAX > (2 \* MIN))

PWR = MIN

ELSE

PWR = AVG

Here, PWR may be scaled by a predetermined factor to produce masking energy.

[00049] A similar algorithm could also be used to create a 48 kHz native encoder using a core sample rate of 6 kHz and a large FFT bin resolution of 23.4375 Hz calculated every 2.67 milliseconds. Such a configuration would differ slightly in detection efficiency and inaudibility from the embodiments described above, but it is anticipated that the differences would be slight.

[00050] With regards to decoding, an exemplary configuration would include a software decoder based on a JNI shared library, which performs calculations up through the bin signal-to-noise ratios. Such a configuration would allow an external application to define the symbols and perform pattern matching. Such steps would be handled in a Java environment using an information hiding extraction of a decoder payload, where decoder payloads are created using specially written Java classes.

[00051] Turning to FIG. 3, an exemplary server-side encoding embodiment is illustrated. In this example, content server 100 has content 320, which includes a media file 302 configured to be requested and played on media player 301 residing on user device 130. When media file 302 is initialized, audio is extracted from the media file and, if the audio is encoded (e.g., MP3 audio), subjected to audio decoding in 304 to produce raw audio 305. To encode the audio for beaconing, device ID, HHID and/or PPMID data is provided for first encoding 306 the data into the raw audio 305, using any suitable technique described above.



- 15 -

[00052] After the first encoding, the audio data is then subjected to a second encoding to transform the audio into a suitable format (e.g., MP3) to produce fully encoded audio 308, which is subsequently transmitted to media player 301 and beacons to portable device 140. Alternately, encoded audio 308 may be produced in advance and stored as part of media file 302. During the encoding process illustrated in FIG. 3, care must be taken to account for processing delays to ensure that the encoded audio is properly synchronized with any video content in media file 302.

[00053] The server-side encoding may be implemented under a number of different options. A first option would be to implement a pre-encoded beacon, where the encoder (306) would be configured to perform real-time encoding of the audio beacon based on the content being served to the users/panelists. The user device would be equipped with a software decoder as described above which is invoked when media is played. The pre-encoded beacon would establish a message link which could be used, along with an identifier from the capturing portable device 140, in order to assign credit. The encoding shared library would preferably be resident at the content site (100) as part of the encoding engine. Such a configuration would allow the transcoding and encoding to be fit into the content site workflow.

[00054] Another option for server-side encoding could include a pre-encoded data load, where the audio is encoded with a message that is based on the metadata or the assigned URL. This establishes a message link which can be used, along with an identifier from the capturing portable device 140, in order to assign credit. The encoding shared library is preferably resident at the content site (100), as part of the encoding engine. Again, this configuration would allow the transcoding and encoding to be fit into the content site workflow.

[00055] Yet another option for server-side encoding could include "on-the-fly" encoding. If a video is being streamed to a panelist, encoding may be inserted in the stream along with a transcoding object. The encoding may be used to encode the audio with a simple one of  $N$  beacon, and the panelist user device 130 would contain software decoding which is invoked when the video is played. This also establishes a message link which can be used, along with an identifier from the capturing portable

- 16 -

device 140, in order to assign credit. The encoding shared library is preferably resident at the content site (100), as part of the encoding engine. Under a preferred embodiment, an ActionScript would invoke the decoding along with a suitable transcoding object.

[00056] FIG. 4 illustrates one embodiment for encoding media under a Flash Video platform 410, where the content is preferably encoded in advance. As raw audio from a video file or other source 400 is received, the audio is subjected to water mark encoding 401, which may include techniques described above for the encoding. Once encoded, the audio is formatted as a Flash file using Adobe Tools 402 such as FLV Creator and SWF Compiler. Once compiled, the file is further formatted using Flash-supported codecs (e.g., H.264, VP6, MPEG-4 ASP, Sorenson H.263) and compression 403 to produce a watermarked A/V stream or file 404.

[00057] FIG. 5 provides another alternate embodiment that illustrates client-side encoding and processing. In this example, user device 130 requests media data. In response to the request, a media file 531 residing on content server 100 is subsequently streamed to the device's browser 520 arranged on user's workspace 510. Media player 521 plays the streamed content and produces raw audio 511. A client-side ActionScript notifies browser 522 and encoder 522 to capture the raw audio on the device's sound mixer, or microphone (not shown), and to encode data using a suitable encoding technique described above. The encoding constructs the data for an independent audio beacon using the captured audio and other data (e.g., device ID, HHID, etc.) where portable device 140 picks up the beacon and forwards the data to an appropriate server for further processing and panel data evaluation.

[00058] Similar to the server-side embodiment disclosed in FIG. 3, care must be taken in the software to account for processing delays in audio pickup and (CBET) encoding of the audio beacon. Preferably, synchronization between audio beacon playback and audio playback (specifically FLV playback) should be accounted for. In alternate embodiments, communication between media player 521 and encoder 522 could be through Actionscript interface APIs, such as "ExternalInterface", which is an application programming interface that enables

- 17 -

straightforward communication between ActionScript and a Flash Player container; for example, an HTML page with JavaScript, or a desktop application with Flash Player embedded, along with encoder application 522. To get information on the container application, an ActionScript interface could be used to call code in the container application, including a web page or desktop application. Additionally, ActionScript code could be called from code in the container application. Also, a proxy could be created to simplify calling ActionScript code from the container application.

[00059] For the panel-side encoding, a beacon embodiment may be enabled by having an encoding message being one from a relatively small set (e.g., 1 of 12), and where each user device 130 is assigned a different message. When portable device 140 detects the encoded message, it identifies the user device 130. Alternately, the encoding message may be a hash of the site and/or URL information gleaned from the metadata. When a panelist portable device 140 detects and reports the encoded message, a reverse hash can be used to identify the site, where the hash could be resolved on one or more remote server (e.g., sever 160).

[00060] In addition to the encoding techniques described above in connection with media content, a simplified beaoning configuration may be arranged where the beacon operates as a complement to media data, independent of the media data, or providing a beacon where no specially encoded data exists. Referred to herein as a "twinkle," the simplified beaoning comprises a constant amplitude acoustic signal or tone that is generated on user device 130. This acoustic tone is then automatically encoded, preferably with identification data (e.g., device ID, HHID and/or PPM ID) and a timestamp. The encoded acoustic tone would then be forwarded to portable device 140 for processing and identification.

[00061] The acoustic tone used for the twinkle is preferably embodied as a pre-recorded constant amplitude tone that is transmitted at predetermined times. The encoding is preferably performed using any of the techniques described above. Under one embodiment, the simplified beaoning process would only forward the encoded, pre-recorded tone, independently of any audio data being received. Thus, referring

- 18 -

back to FIG. 1, it is possible that user device 130 receives only other content 114 from content server 110 in the form of text-based HTML. As PC meter 132 records browsing information, the encoded tone is transmitted to portable device 140, where after further processing (see FIG. 1B), the user identification data is merged into internet measurement database 141. It is understood, that user device 140 may also receive audio data (encoded or unencoded) separately and in addition to other content. While the techniques described above would encode and forward audio data received, the simplified beacon ("twinkle") would also transmit ID information to portable device 140, which, in conjunction with PC meter 132, would subsequently merge panelist data into a common database.

[00062] In another exemplary embodiment, FIG. 6 illustrates audio signal 600 represented as a spectrum of audio 610 over a period of time (e.g., 0.25 seconds), where the energy intervals vary with frequency between 1200 and 2200 Hz. Overlaid in black are discrete, narrowband code tones 602 (e.g., CBET) opportunistically inserted into the audio using the principles of psychoacoustic masking. For encoded tones, one of which is illustrated as 602 in FIG. 6, the energy of the inserted code tone varies with the level of audio, so more quiet portions of the frequency spectrum (e.g., 604) receive little encoding energy and compared to louder portions (e.g., 605), which get proportionally more.

[00063] In contrast, the simplified encoding ("twinkle") 603 is encoded and inserted at constant levels across the frequency spectrum, where the levels are independent of the audio levels. This allows the simplified encoding to be pre-recorded, easily generated and capable of being reused across various and/or different content. The simplified encoding could have the same message structure as the CBET encoding described above, utilizing a 10-tone symbol set. Alternately, other message structures are possible as well. As mentioned above, the twinkle may be transmitted automatically at regular intervals. Alternately the twinkle may be invoked by an ActionScript. If two players are utilized (i.e., one for the media, and one for the twinkle), the ActionScript could relay a beacon for the media from user device 130 to portable device 140, while simultaneously requesting a second

- 19 -

(preferably invisible) Flash Player in the user device 130 to transmit the twinkle to portable device 140. Under a preferred embodiment, the ActionScript should invoke both players at a common volume setting.

[00064] In certain embodiments, it is advantageous to configure the audio beacon or "twinkle" on a device so that it is not intrusive and/or distracting to the user. Additionally, the beacon may be arranged to have audio characteristics that make it easier and more robust to encode. For example, simulated environmental sounds, such as power supply fans, air vent exhaust, crowd/audience noise, ocean waves and such may be used as the audio beacon sound in which encoded messages may be inserted. In another example, it is known that many computing devices, such as personal computers, phones, tablets and laptops, contain a pre-stored library of audio that is used to alert or notify users. By creating and storing sounds advantageously suited to encode data in the on-device sound library, a more robust beacon may be utilized. Furthermore, since the sound used for the beacon is known *a priori*, the encoding on the device may be simplified, since the time, frequency, masking and other encoding processes will be known.

[00065] Turning to FIG. 7, an exemplary embodiment is illustrated where user device 130 is configured with a control panel 700 that allows the device to control various aspects, including sound 701 that includes a resident sound library 720. In a preferred embodiment, a script, operating alone or as part of meter software (see FIG. 1A, ref. 132), may control aspects of sound library 720, such that specific sounds are used for audio beacon encoding. Of course, it is possible for a user to manually change these features either directly through user device 130, or through a remote connection.

[00066] Sound library 720 comprises audio sounds (704-707) that are associated with one or more software applications 700-701 and/or events 702-703. "Events," for the purposes of FIG. 7, include notifications (e.g., receipt of email, social networking software status update, etc.) and alerts (e.g., start-up, shut-down, application error, etc.) that are made pursuant to the device's operating system and associated software running on the device. In this example, a first application

- 20 -

(APP1) 700 and a second application (APP2) 701 are associated with respective sounds 704-705 that may be triggered when the application is activated, or at predetermined times determined by the specific application. Applications 700-701 may include metering software 132 discussed above in connection with FIG. 1A.

[00067] In one embodiment, each sound 704, 705 is associated with a respective application 700, 701. The sounds (704-705) are used for encoding messages to form respective beacons (708-709). As data for the audio beacon is being collected on user device 130 (e.g., user device ID, web hash, etc.), the sound (e.g., 704) is copied, data is encoded into it and the encoded copy is stored in a buffer or other suitable memory. As further data for an audio beacon is collected, another copy of the original sound (704) is made, and the further data is encoded and stored into the buffer/memory. This process repeats for as long as necessary to form a string of encoded sounds. The sounds may be arranged sequentially or in other suitable formats. When an application 700 triggers a sound, a first beacon 708 is audibly transmitted. When the next trigger occurs, the next beacon is audibly transmitted, and so on, until the buffer/memory is empty, or a predetermined amount of time has expired. In another embodiment, one application can control a plurality of sounds (704-705) and produce a plurality of encoded beacons (708-709).

[00068] Continuing with FIG. 7, events 702-703 may be associated with respective sounds 706-707, similar to applications 700-701 discussed above, to produce beacons 710-711. This configuration may be particularly advantageous to beacon information when a user performs an act on the device, such as open/close a browser window, open/close a tab on a browser, open/close an application, etc. As data for the audio beacon is being collected (e.g., user device ID, web hash, etc.), the sound (e.g., 706) is copied, data is encoded into it and the encoded copy is stored in a buffer or other suitable memory. As further data for an audio beacon is collected, another copy of the original sound (706) is made, and the further data is encoded and stored into the buffer/memory. This process repeats for as long as necessary to form a string of encoded sounds. Again, the sounds may be arranged sequentially or in other suitable formats. When an event (702) is detected, a first beacon 710 is audibly

- 21 -

transmitted. When the next event occurs, the next beacon is audibly transmitted, and so on, until the buffer/memory is empty, or a predetermined amount of time has expired.

[00069] As mentioned previously, the sounds 704-707 are preferably predetermined and may simulate an environmental sound so as not to be intrusive or distracting to the user. Additionally, the sound may be selected to contain audio characteristics (e.g., having high masking levels in critical frequency bands) that makes it conducive to robust audio encoding. By using a predetermined sound for the audio beacon encoding, designers can have more flexibility in audibly beaconing data. In addition to audio characteristics, the predetermined sounds may have different lengths as well. In an example where an application (700) controls multiple sounds, the sounds may be the same instance of one sound, but having different lengths (e.g., 5 sec., 10 sec., etc.). In cases where a device's (130) volume is lower, the application may default to a longer sound to increase the probability of the beacon code being detected. If the volume increases, the device 130 can switch to a shorter sound. This configuration has the added benefit of ensuring that users maintain a sufficient volume on their device to avoid longer (and possibly more intrusive) beacons.

[00070] Various embodiments disclosed herein provide devices, systems and methods for performing various functions using an audience measurement system that includes audio beaconing. Although specific embodiments are described herein, those skilled in the art recognize that other embodiments may be substituted for the specific embodiments shown to achieve the same purpose. As an example, although terms like "portable" are used to describe different components, it is understood that other, fixed, devices may perform the same or equivalent functions. Also, while specific communication protocols are mentioned in this document, one skilled in the art would appreciate that other protocols may be used or substituted. This application covers any adaptations or variations of the present invention. Therefore, the present invention is limited only by the claims and all available equivalents.

- 22 -

**CLAIMS**

What is claimed is:

Claim 1. A method for measuring and communicating media exposure, comprising the steps of:

receiving data in a user device from a computer network;

obtaining first characteristic data from the data in the user device;

generating a predetermined sound in the user device in response to receiving the data, wherein the sound comprises environmental audio;

encoding the sound with second characteristic data; and

transmitting the encoded sound to a remote location for media exposure processing.

Claim 2. The method according to claim 1, wherein the first characteristic data is obtained from one of a site ID, URL page, URL file and timestamp.

Claim 3. The method according to claim 2, wherein a unique identifier is appended to the first characteristic data.

Claim 4. The method according to claim 3, wherein the second characteristic data is one of a unique user device ID, a household ID (HHID), a portable device ID (PPMID), and another timestamp.



- 23 -

Claim 5. The method according to claim 4, wherein the data comprises audio data, and wherein the second characteristic data is encoded to the audio data at a constant level.

Claim 6. The method according to claim 1, wherein the encoding is performed by embedding the second characteristic data within the sound where the second characteristic is audibly imperceptible within the audio data.

Claim 7. The method according to claim 4, wherein the data comprises non-audio data, and wherein the second characteristic data is encoded into the sound at a constant level.

Claim 8. A method for measuring and communicating media exposure, comprising the steps of:

receiving media data in a user device, said media data comprising audio data;

obtaining first characteristic data from the media data in the user device;

sampling at least a portion of the audio data in the user device, wherein the sampled portion is processed in the user device to be subsequently formed as an audio signature;

generating predetermined sound in the user device in response to receiving the media data, wherein the sound comprises environmental audio;

encoding the sound with second characteristic data; and

transmitting the encoded sound to a remote location for media exposure processing

- 24 -

Claim 9. The method according to claim 8, wherein the first characteristic data is obtained from one of a site ID, URL page, URL file and timestamp.

Claim 10. The method according to claim 9, wherein a unique identifier is appended to the first characteristic data.

Claim 11. The method according to claim 10, wherein the second characteristic data is one of a unique user device ID, a household ID (HHID), a portable device ID (PPMID), and another timestamp.

Claim 12. The method according to claim 11, wherein the second characteristic data is encoded to the sound at a constant level.

Claim 13. The method according to claim 8, wherein the encoding is performed by embedding the second characteristic data within the sound where the second characteristic is audibly imperceptible within the audio tone.

Claim 14. A method for measuring media exposure in a processing system, the method comprising the steps of:

receiving first characteristic data related to data that was accessed at a user device;

receiving second characteristic data related to the media data, the second characteristic data being different from the first characteristic data, wherein said second characteristic data is formed from previous acoustic encoding performed on an

- 25 -

environmental sound generated at the user device in response to the user device accessing the data; and

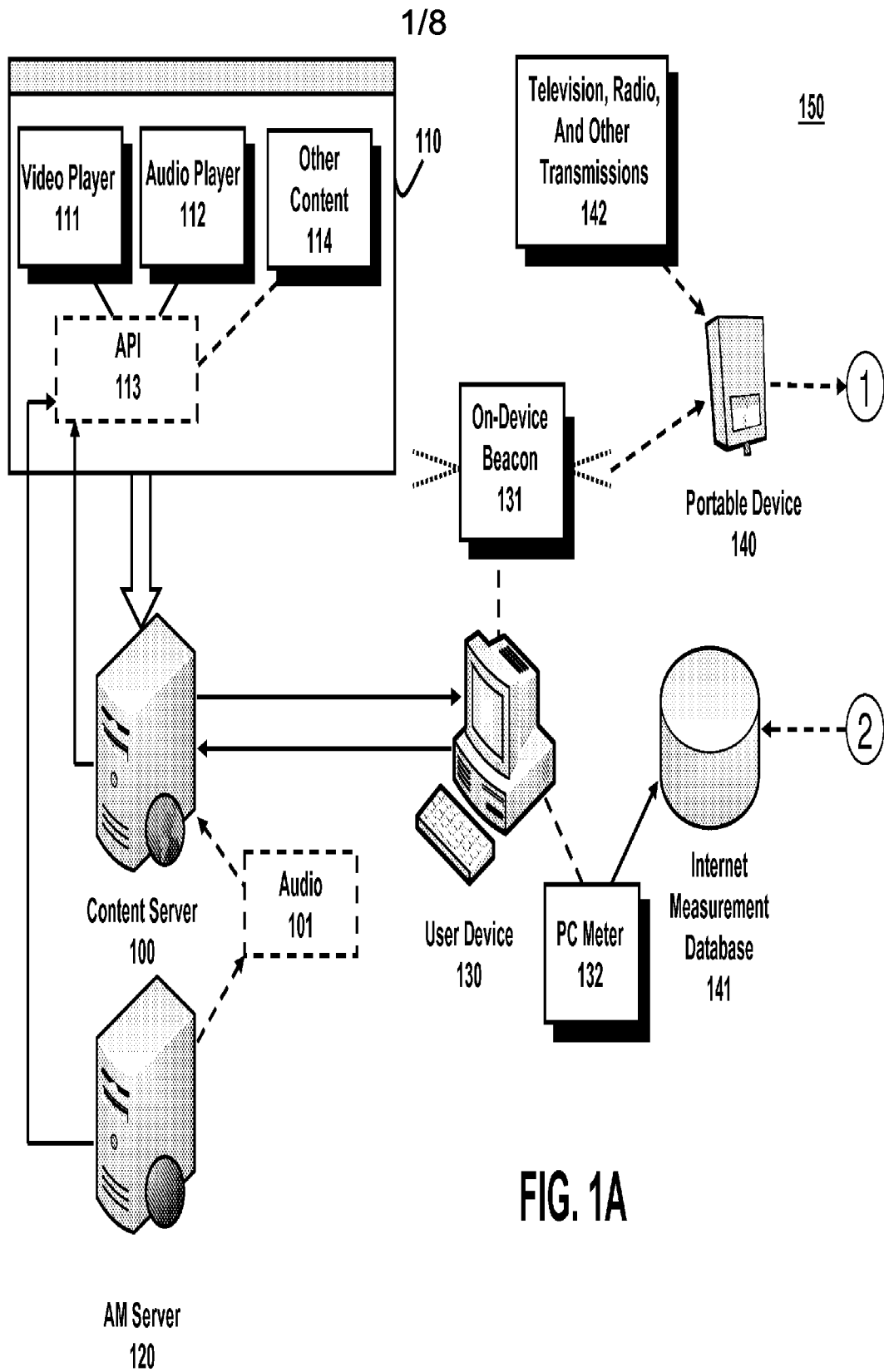
correlating the first and second characteristic data to determine a level of media exposure.

Claim 15. The method according to claim 14, wherein the first characteristic data is obtained from one of a site ID, URL page, URL file and timestamp related to the media data.

Claim 16. The method according to claim 15, wherein the first characteristic data further comprises a unique identifier related to the user device.

Claim 17. The method according to claim 16, wherein the second characteristic data is one of a unique user device ID, a household ID (HHID), a portable device ID (PPMID), and another timestamp related to the user device.

Claim 18. The method according to claim 17, wherein the acoustic encoding of the second characteristic comprises embedding the second characteristic data within the sound at a constant level where the second characteristic is audibly imperceptible within the audio data.



2/8

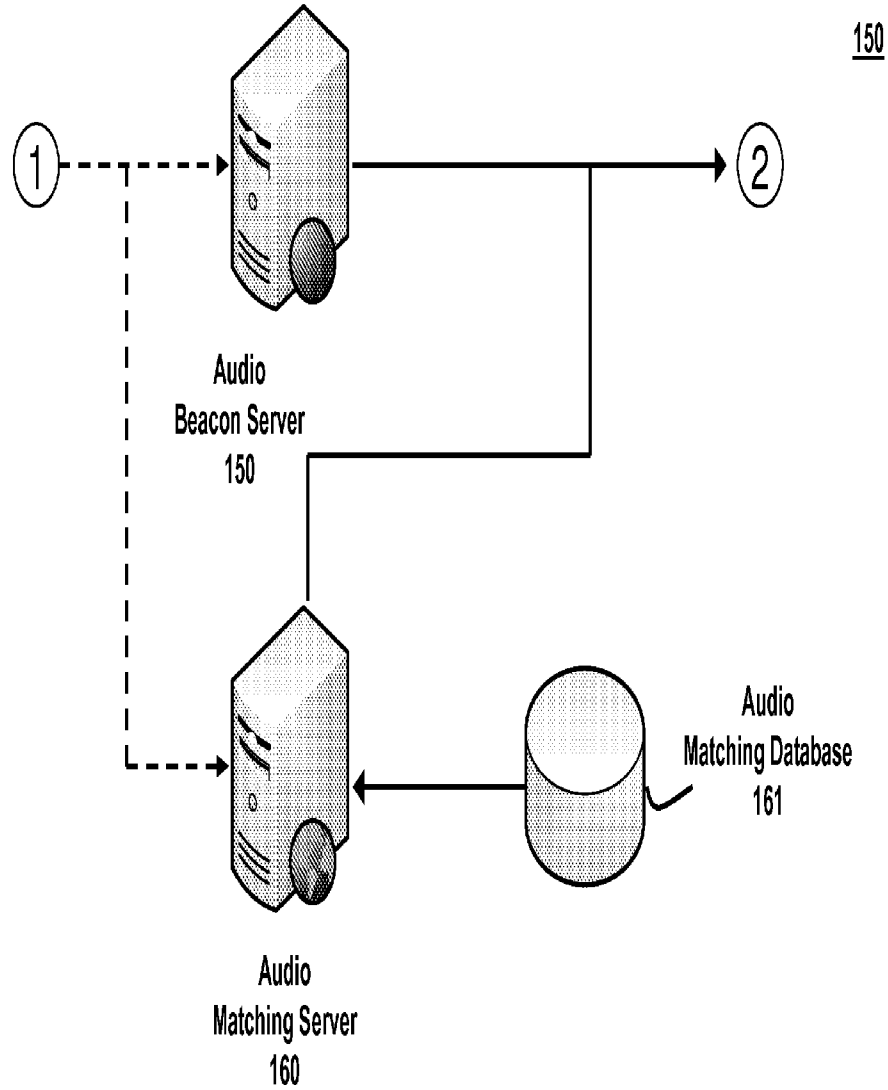


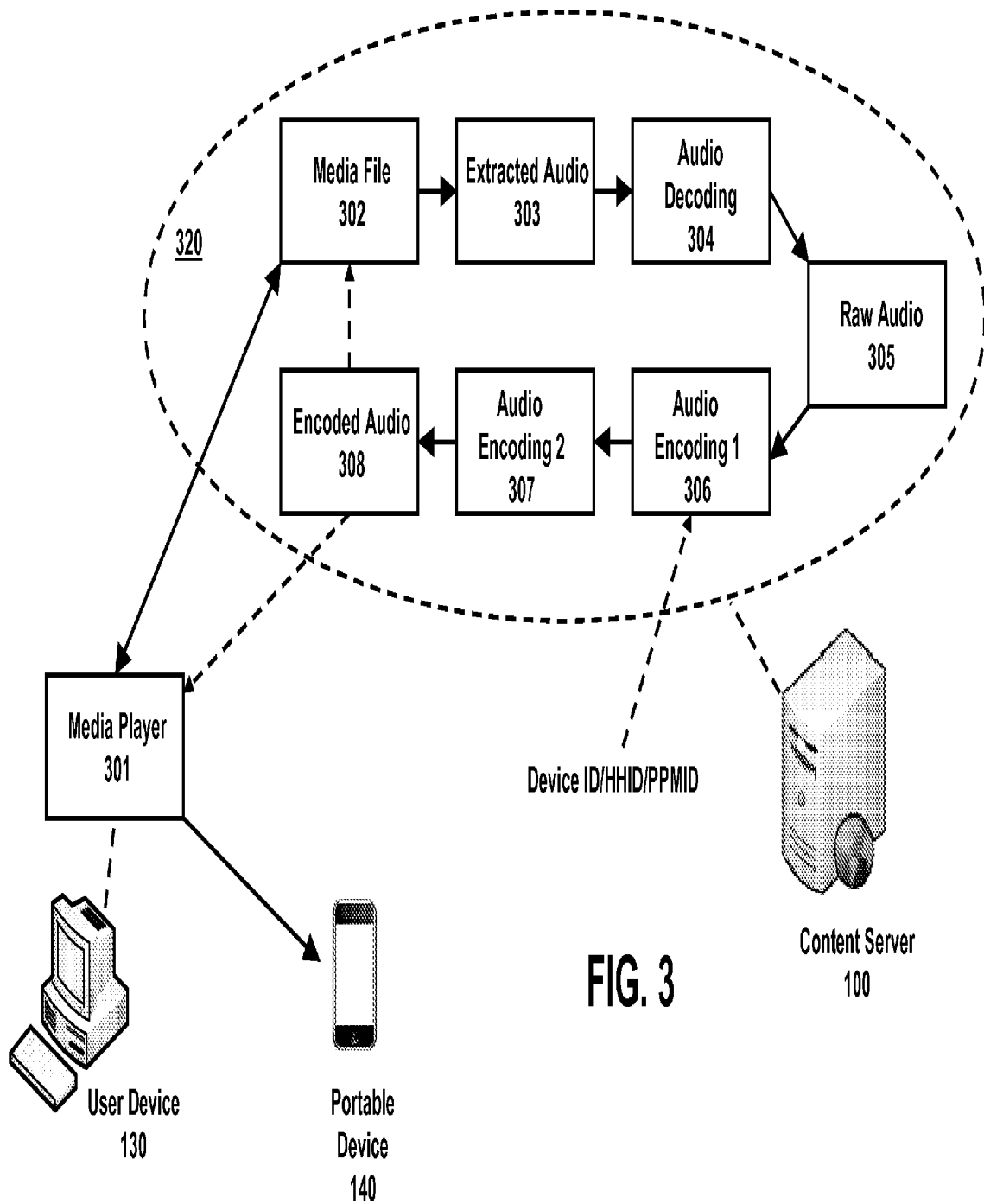
FIG. 1B

3/8

|     |                                                     |                                                            |                               |                          |                                                   |                                     |                                                             |                                             |                                                    |
|-----|-----------------------------------------------------|------------------------------------------------------------|-------------------------------|--------------------------|---------------------------------------------------|-------------------------------------|-------------------------------------------------------------|---------------------------------------------|----------------------------------------------------|
| 200 | Time →                                              | 0-3 sec                                                    | 4 sec ----- 180 ----- 300 sec |                          |                                                   |                                     | 301-315 sec                                                 | 316—599 sec                                 | 600 sec                                            |
| 201 | Activities→                                         | Video Loaded                                               | Video Played                  |                          |                                                   |                                     | Ad served                                                   | Video continues                             | Video End                                          |
| 202 | Audio Beacon detected by portable device →          | Continuous Audio representing the website (e.g., Hulu.com) |                               |                          |                                                   |                                     |                                                             |                                             |                                                    |
| 203 | Additional beacon events →                          | URLs (Hulu.com & Video ID)                                 |                               |                          |                                                   |                                     | AD Event Beacon (co-<br>relation URL)                       |                                             | Video End Beacon                                   |
| 204 | Audio matching events detected by portable device → |                                                            | Audio Signature Sample 1      | Audio Signature Sample 2 | Audio Signature Sample 3 (Content Match Achieved) | Same Signature Sample 4 (No Change) | Audio Signature Sample 5 (Different Content couldn't match) | Audio Signature Sample 5 (Same Video again) | No Signature (But Can't assume video is completed) |
| 205 | Overlay portable device data                        | X                                                          |                               |                          | X                                                 |                                     | X                                                           | X                                           | X                                                  |
| 206 | End- User Experience                                | URL of site where the video was played                     |                               |                          | Content Matched                                   |                                     | Ad Served                                                   |                                             | Video Played till the End                          |

FIG. 2

4/8



5/8

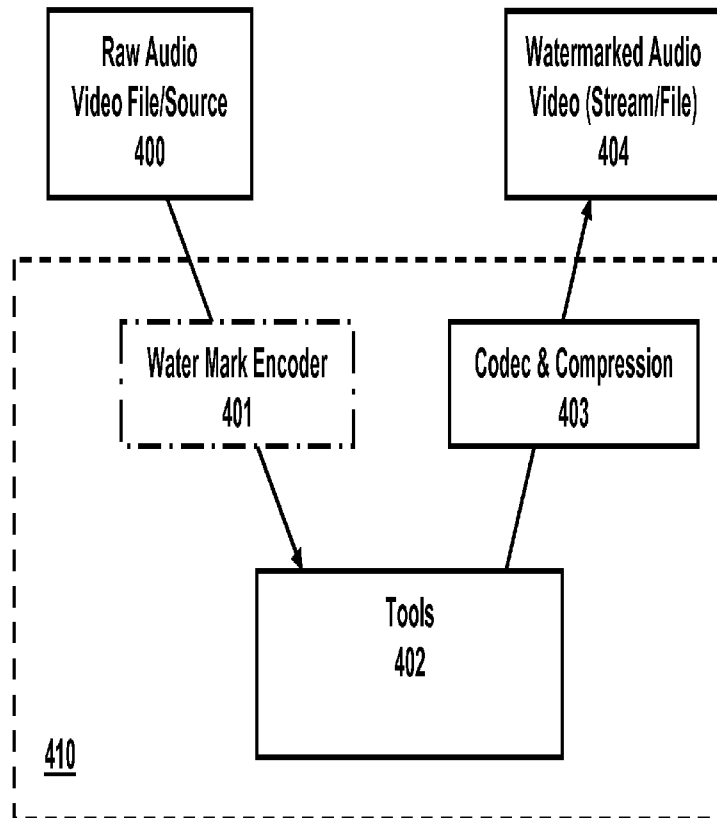
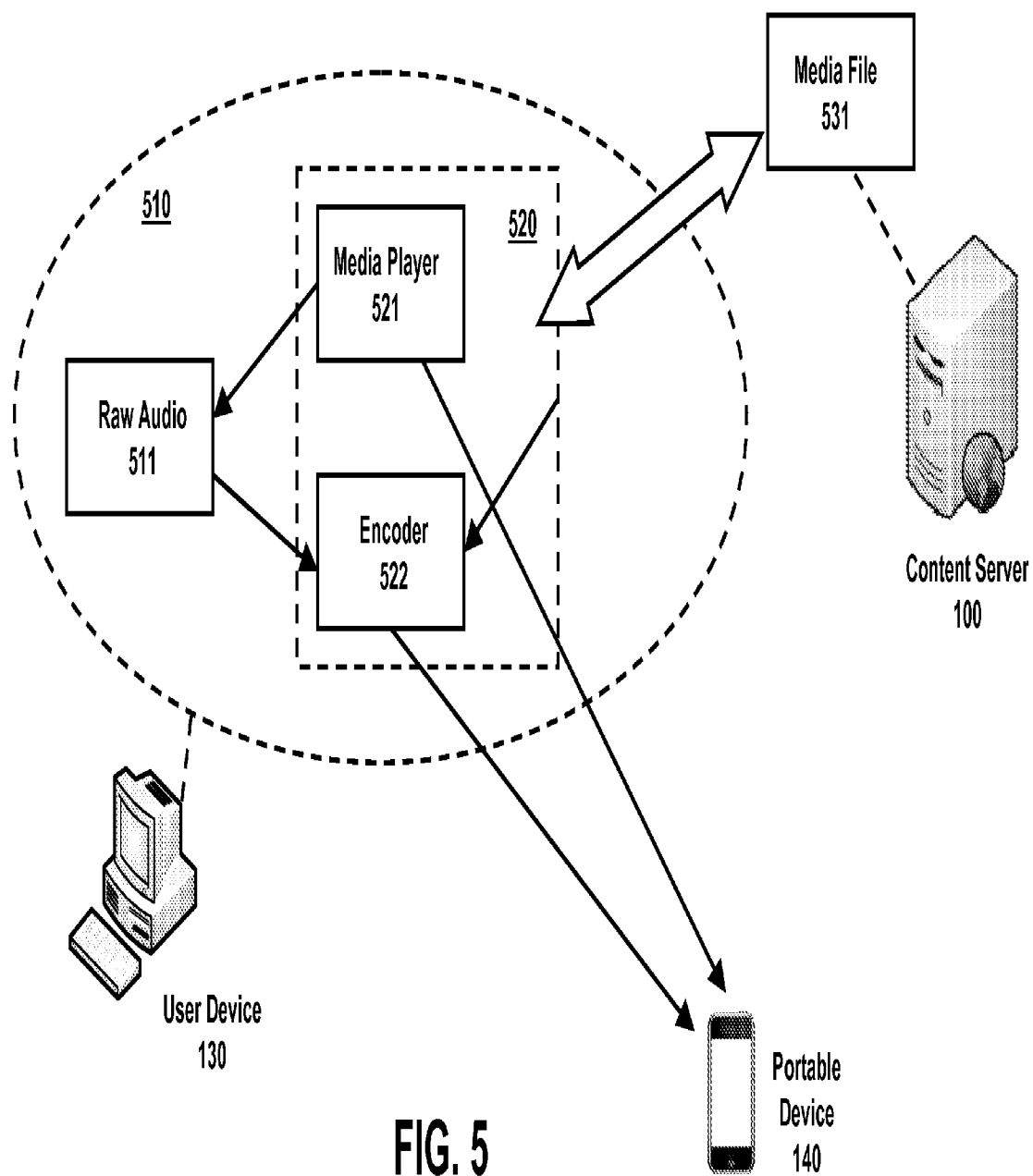


FIG. 4



6/8



7/8

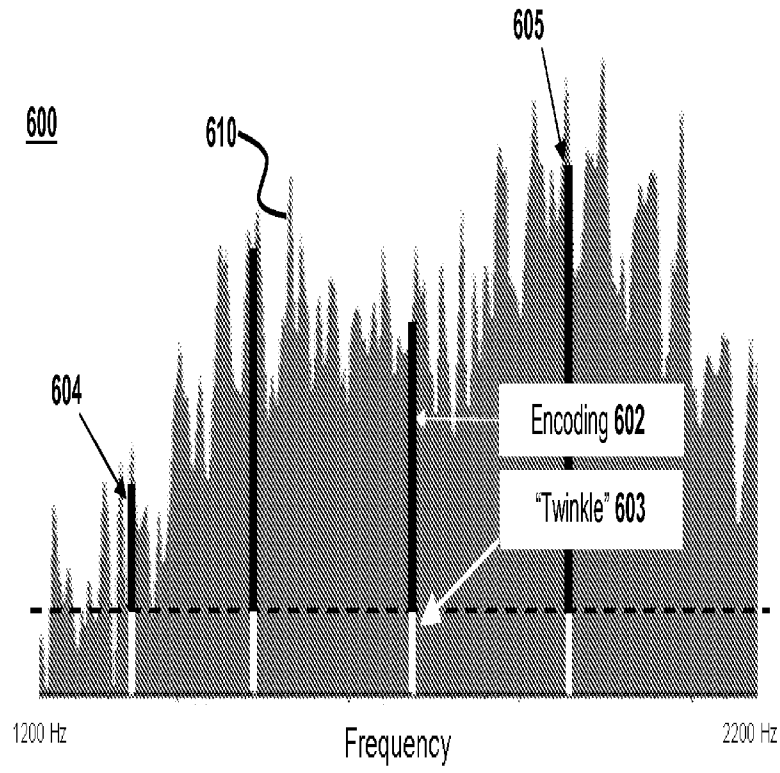
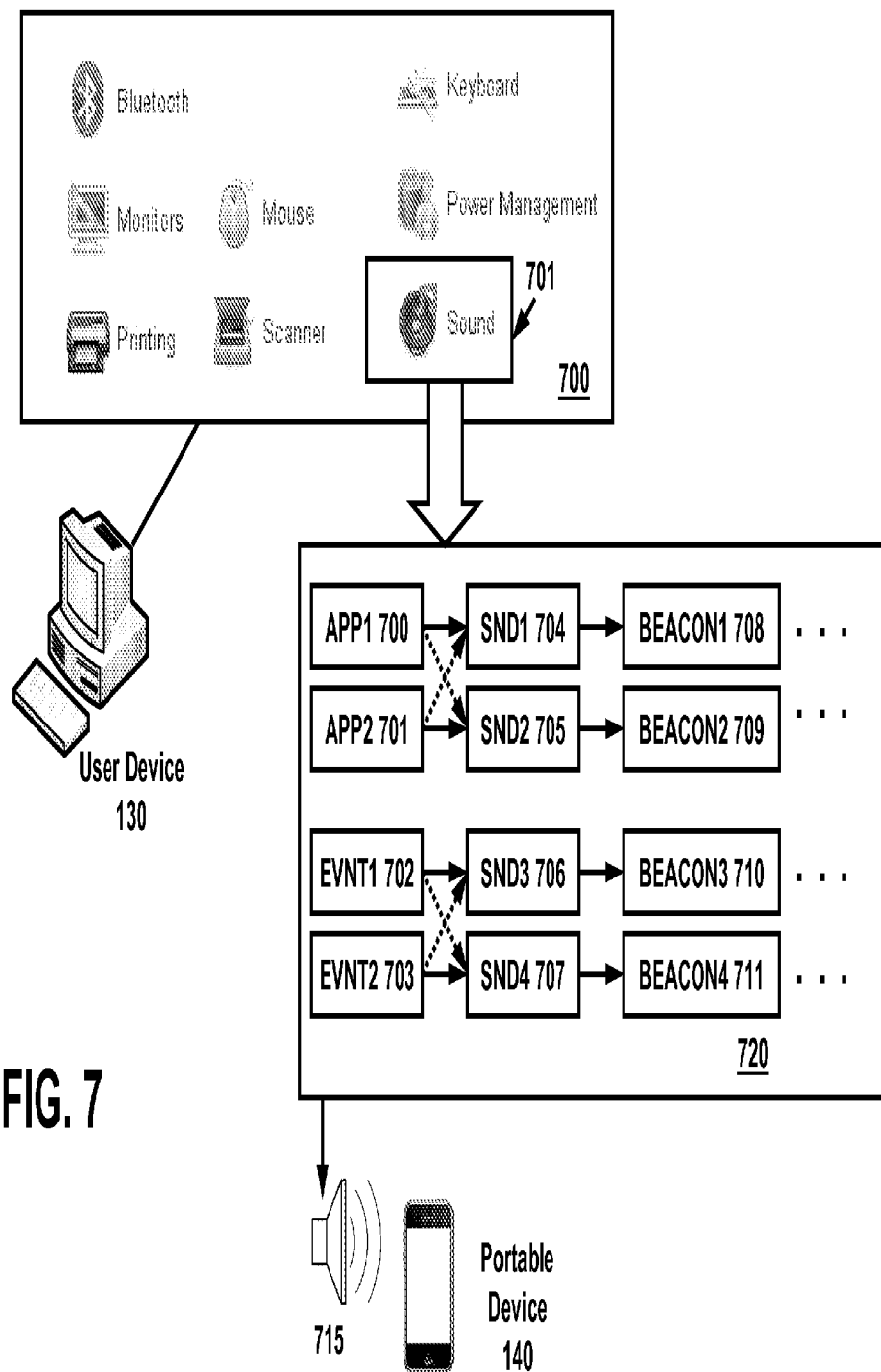


FIG. 6

8/8



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/71966

## A. CLASSIFICATION OF SUBJECT MATTER

**IPC(8) - H03F 1/26 (2013.01)****USPC - 702/189**

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)  
USPC: 702/189Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
USPC: 702/189 or 386/46 or 709/231 or 700/94 or 380/231 or 380/200 or 434/156 or 455/2.01 or 705/1 or 725/34 or 709/203 or 709/206 or 381/119Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
PatBase, Google (Web), Google Patents; inaudible, imperceptible, environmental, background, ambient, noise, sound, audio, auditory, psychoacoustic, masking, signature, fingerprint, characteristic, identifier, id, watermark, URL

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages                         | Relevant to claim No. |
|-----------|------------------------------------------------------------------------------------------------------------|-----------------------|
| Y         | US 2010/0268540 A1 (Arshi et al.) 21 October 2010 (21.10.2010), Fig. 1A, para [0005]-[0006], [0024]-[0034] | 1-18                  |
| Y         | US 2007/0100483 A1 (Kentish et al.) 03 May 2007 (03.05.2007), para [0044]-[0045]                           | 1-18                  |
| A         | US 2003/0081781 A1 (Jensen et al.) 01 May 2003 (01.05.2003) entire document                                | 1-18                  |
| A         | US 2004/0137929 A1 (Jones et al.) 15 July 2004 (15.07.2004) entire document                                | 1-18                  |
| A         | US 7,239,981 B2 (Kolessar et al.) 03 July 2007 (03.07.2007) entire document                                | 1-18                  |

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

|                                                                                                                                                                         |                                                                                                                                                                                                                                                  |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| * 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 application or patent but 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<br>01 February 2013 (01.02.2013)                                                                                      | Date of mailing of the international search report<br><b>01 MAR 2013</b>                              |
| Name and mailing address of the ISA/US<br>Mail Stop PCT, Attn: ISA/US, Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-3201 | Authorized office <sup>^</sup><br>Lee W. Young<br>PCT Helpdesk: 571-272-4300<br>PCT OSP: 571-272-7774 |