

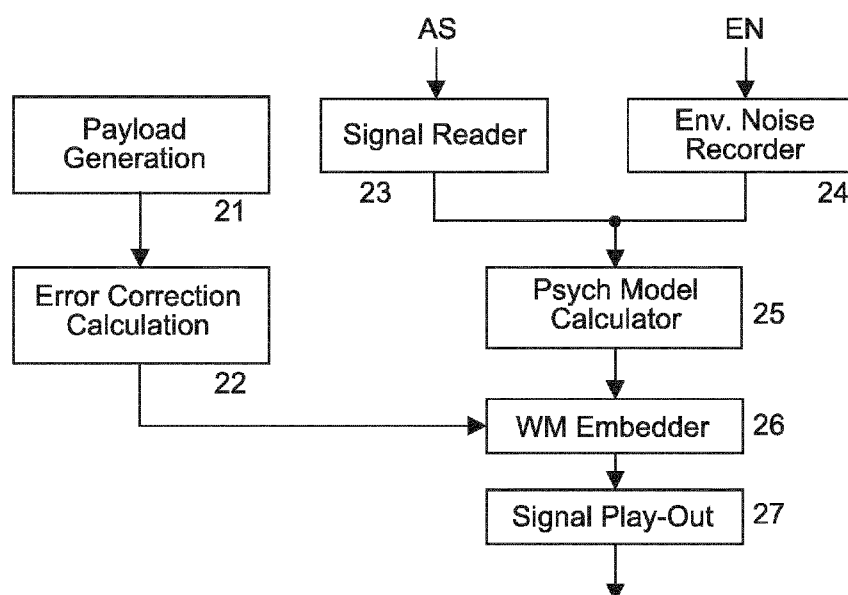


- (51) **International Patent Classification:**  
*G10L 19/018* (2013.01) *G10L 19/16* (2013.01)
- (21) **International Application Number:**  
PCT/EP2014/076108
- (22) **International Filing Date:**  
1 December 2014 (01.12.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
13306687.8 9 December 2013 (09.12.2013) EP
- (71) **Applicant:** THOMSON LICENSING [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).
- (72) **Inventors:** BAUM, Peter Georg; Aschendorfer Str. 25, 30539 Hannover (DE). CHEN, Xiaoming; Vahrenwalderstr. 101, 30165 Hannover (DE). ARNOLD, Michael; Eichendorffstr. 1b, 30916 Isernhagen (DE). GRIES, Ulrich; Friedenauer Str. 20b, 30419 Hannover (DE).
- (74) **Agent:** HARTNACK, Wolfgang; Deutsche Thomson OHG, European Patent Operations, Karl-Wiechert-Allee 74, 30625 Hannover (DE).

- (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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, 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, KM, ML, MR, NE, SN, TD, TG).

**Published:**

— with international search report (Art. 21(3))

(54) **Title:** METHOD AND APPARATUS FOR WATERMARKING AN AUDIO SIGNAL**Fig. 2**

(57) **Abstract:** Improvement of watermark detection in watermarked microphone audio signals picked up in the presence of surrounding noise is achieved by using at encoder side not only the originally received signal for the calculation of the masking threshold and the watermarking strength, but by also taking into account the level of the surrounding noise. This enables an adaptation of the watermarking strength to the current sound pressure level SPL of the surrounding noise. If the SPL of the surrounding noise is increased, the watermarking strength will be increased accordingly. The resulting advantage is a significantly improved audio watermark detection in the presence of surrounding noise.

**Method and Apparatus for watermarking an audio signal**Technical field

5 The invention relates to a method and to an apparatus for watermarking an audio signal taking also into account surrounding noise.

10 Background

Audio watermarking is the process of embedding in an in-audible way information into an audio signal. The embedding is performed by changing the audio signal for example by  
15 adding pseudo-random noise or echoes. To make the embedding in-audible, the strength of the embedding is controlled by a psycho-acoustical analysis of the signal. At receiver side, the watermark can be detected by performing correlation with a pseudo-random noise bit sequence.

20

Summary of invention

The main challenge of current audio watermarking systems is  
25 the robustness against microphone pickup. Especially if there is surrounding noise, it is very difficult to detect the watermark in a watermarked signal that is played back via loudspeaker.

30 A problem to be solved by the invention is to provide improved watermark detection capabilities for microphone audio signals picked-up in the presence of surrounding noise. This problem is solved by the method disclosed in claim 1. An apparatus that utilises this method is disclosed in claim 7.

The inventive improvement of watermark detection in watermarked microphone audio signals picked up in the presence of surrounding noise is achieved by using at encoder side not only the originally received signal for the calculation of the masking threshold and the watermarking strength, but by also taking into account the level of the surrounding noise. This enables an adaptation of the watermarking strength to the current sound pressure level (SPL) of the surrounding noise. If the SPL of the surrounding noise is increased, the watermarking strength will be increased accordingly. The resulting advantage is a significantly improved audio watermark detection in the presence of surrounding noise.

In principle, the inventive method is suited for watermarking an audio signal, including the steps:

- receiving an audio signal and receiving surrounding noise signal or data about a surrounding noise signal;
- calculating a masking threshold for said audio signal, wherein said masking threshold is to be used for embedding watermark payload data and related error correction data, and wherein for calculating said masking threshold the characteristics of said audio signal as well as the characteristics of said surrounding noise are taken into account;
- embedding said watermark payload data and said error correction data into said audio signal and providing the correspondingly watermarked audio signal.

In principle the inventive apparatus is suited for watermarking an audio signal, said apparatus including:

- means being adapted for receiving an audio signal and for receiving surrounding noise signal or data about a surrounding noise signal;
- means being adapted for calculating a masking threshold for said audio signal, wherein said masking threshold is to

be used for embedding watermark payload data and related error correction data, and wherein for calculating said masking threshold the characteristics of said audio signal as well as the characteristics of said surrounding noise are taken into account;

- means being adapted for embedding said watermark payload data and said error correction data into said audio signal and for providing the correspondingly watermarked audio signal.

10

Advantageous additional embodiments of the invention are disclosed in the respective dependent claims.

#### 15 Brief description of drawings

Exemplary embodiments of the invention are described with reference to the accompanying drawings, which show in:

- Fig. 1 Different masking regions;
- 20 Fig. 2 Block diagram of the inventive processing;
- Fig. 3 Example application.

#### Description of embodiments

25

Even if not explicitly described, the following embodiments may be employed in any combination or sub-combination.

For the inventive processing the following application is assumed:

30

- The watermark information embedding into an original audio signal is carried out in real-time in a device connected to a loudspeaker, or a device generating watermarked audio signals intended for a presentation by a

loudspeaker or loudspeakers;

- The corresponding watermarked audio signal is played back by that loudspeaker or loudspeakers;
- A separate device picks up the sound and detects the embedded watermark information, which watermark information is used for example for second-screen applications synchronisation.

Such application happens for example if 2nd screen watermarking embedding is performed in a set-top box or a TV receiver (or any other device emitting sound). The original audio signal to be watermarked is the non-watermarked audio signal received. A listener watching the TV program has a device including a screen (e.g. a tablet computer or a smart phone), which device receives the watermarked acoustic waves from the loudspeaker of the TV receiver.

In a store, a shopper has a mobile device which receives watermarked acoustic waves from one or more loudspeakers arranged nearby his current position within the store, and the watermarked acoustic waves are used for video merchandising or advertising products presented at his current position within that store (like IZ•ON in the USA).

Usually the audio signal is analysed at watermark encoder side and the strength of the embedding is selected based on such analysis, such that the watermark is not audible. This works quite well if there is no surrounding noise. However, if there is surrounding noise (at a listener position), the ratio between watermark amplitude and disturbing noise amplitude (i.e. signal to noise ratio SNR) gets smaller, which means that the correct-detection rate of the watermark detector will decrease.

Usually, the strength of watermark information embedding is controlled by a masking threshold which quantitatively

measures the effect of masking. The maskee depicted in Fig. 1 is the tone which masks out other sound, whereas the test sound is the sound which will be masked (i.e. the watermark signal).

5 However, in general, two different situations can be distinguished regarding the time relation  $\Delta t$  between the masker and the test sound:

- Simultaneous masking as depicted in region II of Fig. 1;
  - Non-simultaneous masking as depicted in regions I and III
- 10 of Fig. 1: pre-masking in region I and post-masking in region III.

The masking threshold of the original signal is derived from the simultaneous masking region, since the original audio signal is available at the time of embedding, whereby the

15 analysis is carried out in blocks having a time resolution of about 10-20ms.

According to the invention, the embedding device evaluates the signal of a microphone which picks up the surrounding

20 noise. For the calculation of the embedding strength not only (the level of) the audio content itself is used, but also (the level of) the surrounding noise. Since the surrounding noise has the effect of an additional psycho-acoustical masker, the watermark strength can be increased without be-

25 coming audible.

Since the surrounding noise has to be recorded or stored before the analysis of the corresponding noise masking threshold can be derived, it naturally fits into the non-simultaneous post-masking region, i.e. into region III in Fig. 1. Although there will be a decay of the post-masking threshold in

30 comparison to the masking threshold within the simultaneous masking region, that decay is limited for  $\Delta t < 50 \text{ ms}$ .

If there is no surrounding noise, the embedding strength is

the same as in the prior art. If there is surrounding noise, the embedding strength will be increased, which means that the watermark robustness will be higher and the detection rate of the audio watermark detector will be better. I.e.,  
5 the more surrounding noise the higher the embedding strength, which mitigates the above-mentioned surrounding noise prior art problems.

In Fig. 2 a step or stage 21 generate payload data for a watermarking to be carried out, followed by a corresponding  
10 error correction data calculation step or stage 22.

A signal reader step or stage 23, which can be a device including a microphone, receives an audio signal AS to be watermarked. Further, an environment or surrounding or ambient  
15 noise recorder 24 receives the related environment noise EN. Recorder 24 can be included in the device with the microphone. A psycho-acoustical model calculating step or stage 25 calculates for each section of the audio signal AS a combined masking threshold for watermark signal insertion,  
20 thereby taking into account the current audio signal magnitude level as well as the corresponding surrounding noise level. Following masking threshold calculation, in a watermark embedding step or stage 26 the payload data including the error correction data are embedded into the audio signal  
25 with a strength according to the combined masking threshold. The correspondingly watermarked audio signal is thereafter played out by a device 27, e.g. an amplifier and a loudspeaker.

Normally the masker is frequency dependent, and the frequency distribution of the original audio microphone signal and  
30 of the ambient noise microphone signal is taken into account.

There are several ways for taking the ambient noise into account. If the microphone is located at the same position as

the listener (for example, a microphone included in a TV remote control or a tablet computer or a smart phone), the psycho-acoustical model can be calculated based on the - possibly weighted - sum of the original signal and the ambient noise signal. The current characteristics of the ambient noise are transferred to the watermark embedder. The mobile device (e.g. the remote control) can send e.g. via infrared signal or via electromagnetic waves like Bluetooth or WLAN or via ultrasound (i.e. any kind of transmission except acoustic waves in the human audible range) data about the current ambient noise characteristics to the TV receiver or to the set top box, i.e. to the device that emits the watermarked sound signal or acoustic waves. The remote control includes an IR command transmitter and a microphone, which microphone receives an audio signal (i.e. the surrounding noise), and the microphone-received audio signal or data about that audio signal can be transmitted via the IR command transmitter.

Another solution is to calculate for both signals one psycho-acoustical model and to calculate the final masking threshold by adding - possibly weighted - both masking thresholds.

If it is important to keep low the complexity of the calculation, it is also possible to calculate the full psycho-acoustical model only for the original audio microphone signal and to calculate a scalar value for the ambient noise microphone signal, for example the - possibly frequency weighted (for example A-weighted) - sound pressure level. The final masking threshold is then the masking threshold of the original audio microphone signal shifted by the scalar value derived from the ambient noise microphone signal.

Fig. 3 shows a person watching a TV 31 and a tablet display or device 32. Nearby the person a remote control 33 is lo-

cated which includes a microphone receiving surrounding noise and which sends a corresponding surrounding noise data signal to a receiving unit 34 of the TV 31. The received signal is evaluated in a block 35 which may comprise the  
5 processing blocks shown in Fig. 2. The TV 31 produces correspondingly watermarked sound that is received in device 32 and can be used for 2nd screen applications.

The described processing (in device 31) can be carried out  
10 by a single processor or electronic circuit, or by several processors or electronic circuits operating in parallel and/or operating on different parts of the complete processing.

The instructions for operating the processor or the processors according to the described processing can be stored in  
15 one or more memories. The at least one processor is configured to carry out these instructions.

Claims

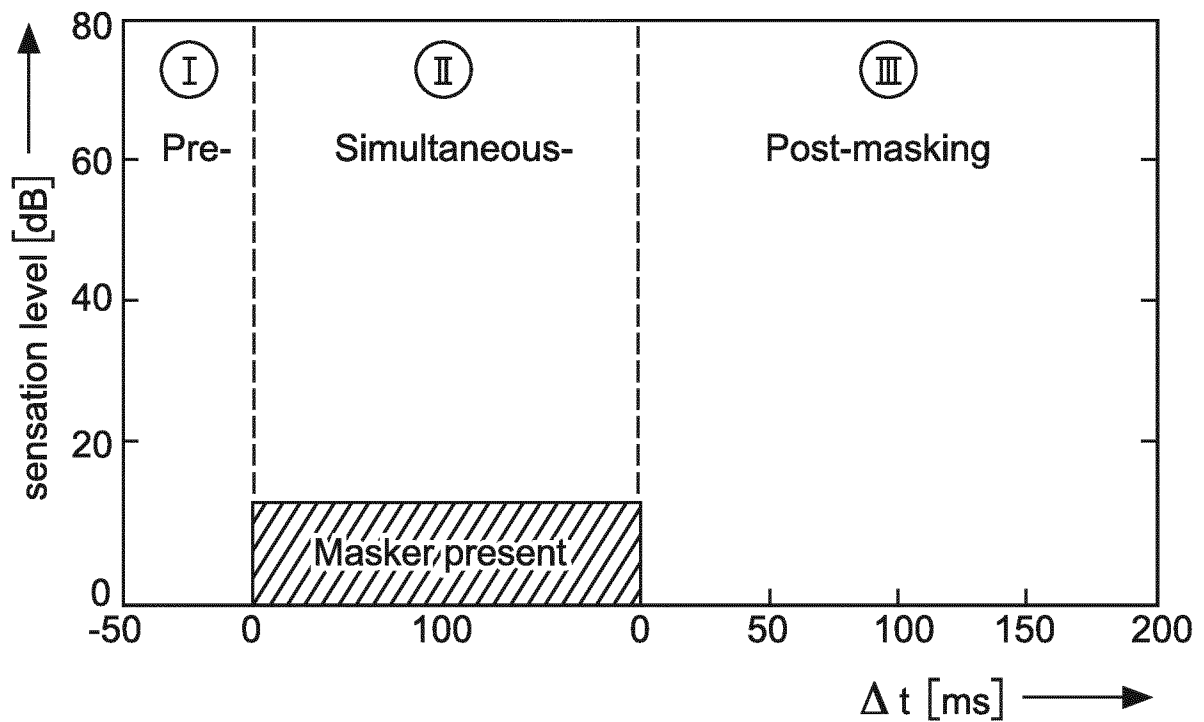
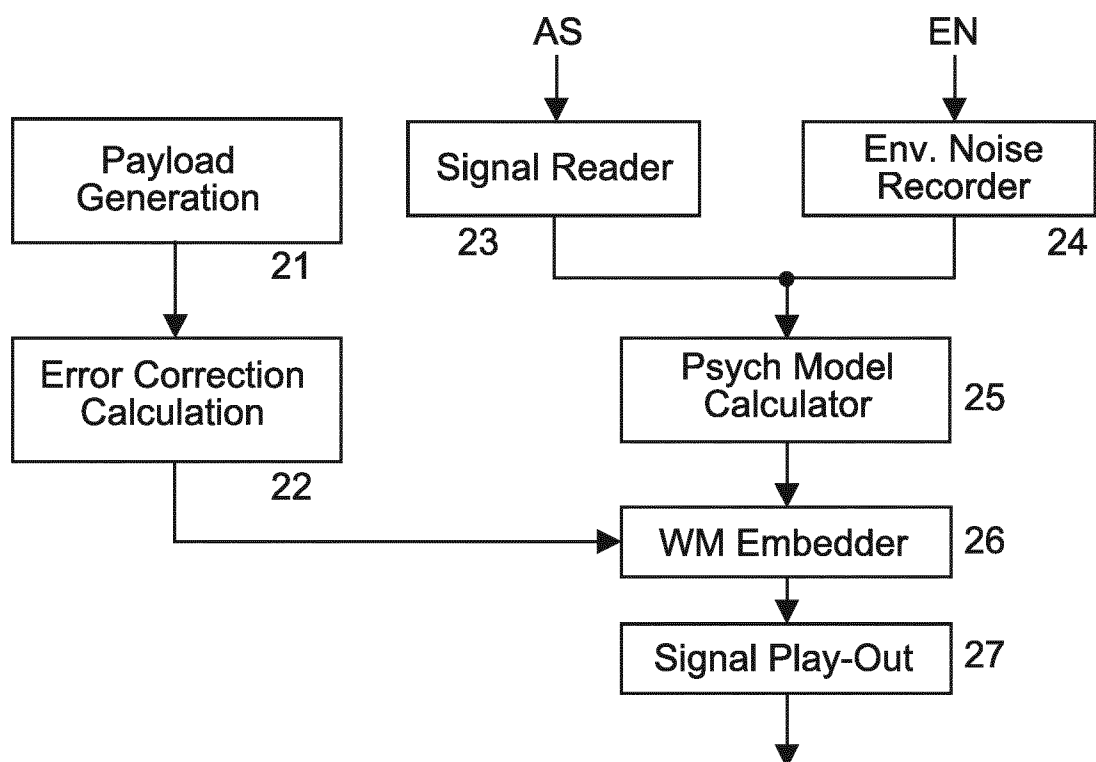
1. Method for watermarking an audio signal, including:
  - receiving (23) an audio signal (AS) and receiving (24) a  
5 surrounding noise signal (EN);
  - calculating (25) a masking threshold for said audio signal, wherein said masking threshold is used for embedding (26) watermark payload data (21), and wherein for calculating (25) said masking threshold the characteristics of  
10 said audio signal as well as the characteristics of said surrounding noise (EN) are taken into account;
  - embedding (26) said watermark payload data into said audio signal and providing (27) the correspondingly watermarked audio signal.
- 15 2. Method according to claim 1, further including:  
using said masking threshold also for embedding error correction data (22) for said watermark payload data into said audio signal.
- 20 3. Method according to claim 1 or 2, further including:  
recording or storing said surrounding noise signal, corresponding to a post-masking region, before said masking threshold calculation.
- 25 4. Method according to one of claims 1 to 3, further including:  
using said watermark payload data for second-screen applications synchronisation.
- 30 5. Method according to one of claims 1 to 4, wherein the masking threshold is frequency dependent, and the frequency distribution of said audio signal and of said surrounding noise signal are taken into account for deter-

mining the watermarking strength to be applied.

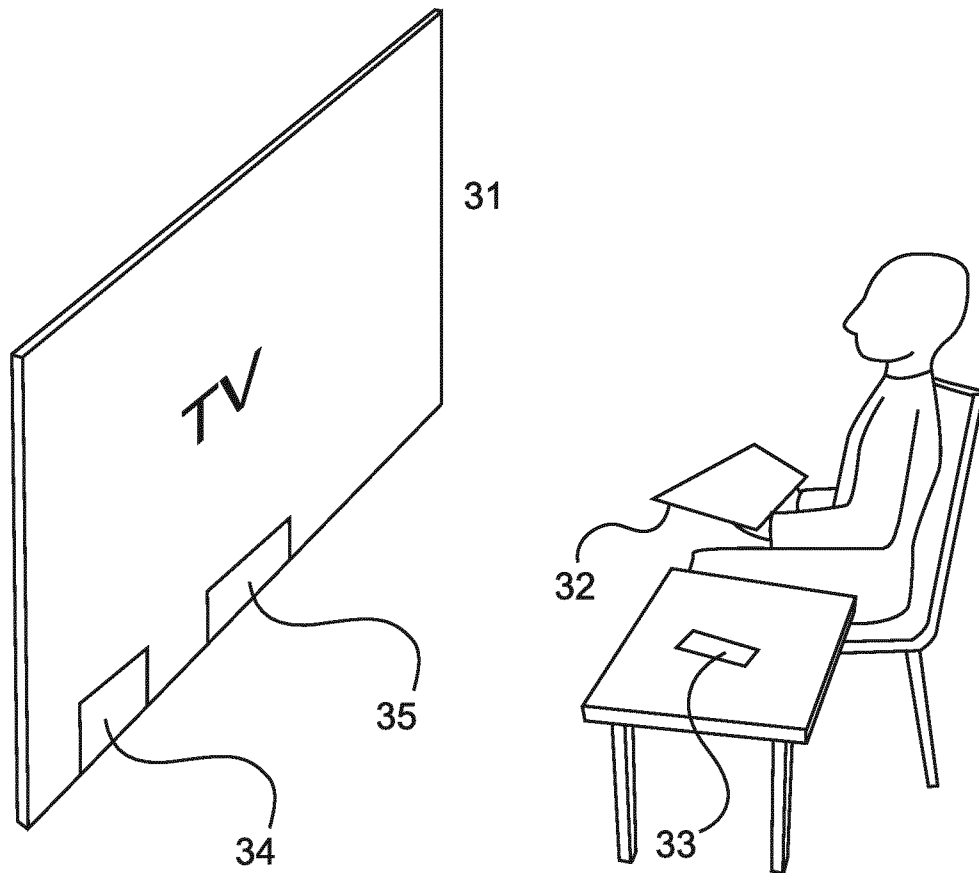
6. Method according to one of claims 1 to 5, wherein a microphone is included in a mobile device - e.g. a remote control - and said mobile device sends - e.g. via infrared signal - surrounding noise signal data, which can represent data about the current ambient noise characteristics, to a device which performs said audio signal watermarking, e.g. to a TV receiver or to a set top box.
7. Apparatus for watermarking an audio signal, said apparatus including:
- means (23, 24) adapted for receiving an audio signal (AS) and for receiving a surrounding noise signal (EN);
  - means (25) adapted for calculating a masking threshold for said audio signal, wherein said masking threshold is used for embedding watermark payload data, and wherein for calculating said masking threshold characteristics of said audio signal as well as characteristics of said surrounding noise are taken into account;
  - means (26, 27) adapted for embedding said watermark payload data into said audio signal and for providing the correspondingly watermarked audio signal.
8. Apparatus according to claim 7, wherein said masking threshold is also used in said embedding means (26, 27) for embedding error correction data (22) for said watermark payload data into said audio signal.
9. Apparatus according to claim 7 or 8, wherein said surrounding noise signal is recorded or stored before said masking threshold calculation and corresponds to a post-masking region.

10. Apparatus according to one of claims 7 to 9, wherein said watermark payload data is used for second-screen applications synchronisation.
- 5 11. Apparatus according to one of claims 7 to 10, wherein the masking threshold is frequency dependent, and the frequency distribution of said audio signal and of said surrounding noise signal are taken into account for determining the watermarking strength to be applied.
- 10 12. Apparatus according to one of claims 7 to 11, which receives said data about surrounding noise from non-acoustic waves emitted from a mobile device with a microphone, which surrounding noise signal data can represent data about the current ambient noise characteristics.
- 15 13. An apparatus for watermarking an audio signal, comprising:
- 20 - a memory that stores data which control operation of a processor;
- said processor, which executes a procedure comprising: receiving (23) an audio signal (AS) and receiving (24) a surrounding noise signal (EN);
- 25 calculating (25) a masking threshold for said audio signal, wherein said masking threshold is used for embedding (26) watermark payload data (21) and related error correction data (22), and wherein for calculating (25) said masking threshold the characteristics of said audio signal as well as the characteristics of said surrounding noise (EN) are taken into account;
- 30 embedding (26) said watermark payload data and said error correction data into said audio signal and providing (27) the correspondingly watermarked audio signal.

1/2

**Fig. 1****Fig. 2**

2/2

**Fig. 3**

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2014/076108

A. CLASSIFICATION OF SUBJECT MATTER  
INV. G10L19/018 G10L19/16  
ADD.

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)  
G10L

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

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

EPO-Internal, COMPENDEX, INSPEC, 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 2012/274459 A1 (JAISIMHA SHREE [US] ET AL) 1 November 2012 (2012-11-01)         | 1-3,5,<br>7-9,11,<br>13 |
| Y         | figure 1<br>paragraphs [0029], [0031] - [0033],<br>[0038], [0046]<br>-----<br>-/-- | 4,6,10,<br>12           |



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 application or patent 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

"&" document member of the same patent family

Date of the actual completion of the international search

19 February 2015

Date of mailing of the international search report

27/02/2015

Name and mailing address of the ISA/

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

Authorized officer

Chétry, Nicolas

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2014/076108

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT |                                                                                                                                                                                                                                                                                                                                                                                                                         |                       |
|------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| Category*                                            | Citation of document, with indication, where appropriate, of the relevant passages                                                                                                                                                                                                                                                                                                                                      | Relevant to claim No. |
| Y                                                    | <p>"Audio Watermarking for Second Screen SyncNow with the Synapse DAW77 module, A Synapse application note",</p> <p>21 May 2012 (2012-05-21), pages 1-13, XP055087681,<br/>Retrieved from the Internet:<br/>URL:http://www.axon.tv<br/>[retrieved on 2013-11-11]<br/>p.1, section "Introduction", second paragraph<br/>p. 2, section "Benefits of the audio watermarking technology", second paragraph</p> <p>-----</p> | 4,6,10, 12            |
| A                                                    | <p>DI CHANG ET AL: "Location Based Robust Audio Watermarking Algorithm for Social TV System",<br/>4 December 2012 (2012-12-04), ADVANCES IN MULTIMEDIA INFORMATION PROCESSING PCM 2012, SPRINGER BERLIN HEIDELBERG, BERLIN, HEIDELBERG, PAGE(S) 726 - 738, XP047014261,<br/>ISBN: 978-3-642-34777-1<br/>figure 1</p> <p>-----</p>                                                                                       | 1-13                  |
| A                                                    | <p>WO 95/06309 A1 (VOICE POWERED TECHNOLOGY INTER [US]) 2 March 1995 (1995-03-02)<br/>figure 1<br/>page 7, line 31 - page 8, line 7</p> <p>-----</p>                                                                                                                                                                                                                                                                    | 1-13                  |
| A                                                    | <p>US 7 454 327 B1 (NEUBAUER CHRISTIAN [DE] ET AL) 18 November 2008 (2008-11-18)<br/>column 8, lines 39-60</p> <p>-----</p>                                                                                                                                                                                                                                                                                             | 1-13                  |

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2014/076108

| Patent document<br>cited in search report | Publication<br>date | Patent family<br>member(s) | Publication<br>date         |
|-------------------------------------------|---------------------|----------------------------|-----------------------------|
| US 2012274459                             | A1                  | 01-11-2012                 | EP 2701946 A2 05-03-2014    |
|                                           |                     |                            | US 2012274459 A1 01-11-2012 |
|                                           |                     |                            | WO 2012148901 A2 01-11-2012 |
| -----                                     |                     |                            |                             |
| WO 9506309                                | A1                  | 02-03-1995                 | AU 7637194 A 21-03-1995     |
|                                           |                     |                            | WO 9506309 A1 02-03-1995    |
| -----                                     |                     |                            |                             |
| US 7454327                                | B1                  | 18-11-2008                 | AT 234533 T 15-03-2003      |
|                                           |                     |                            | DE 19947877 A1 10-05-2001   |
|                                           |                     |                            | DE 50001467 D1 17-04-2003   |
|                                           |                     |                            | EP 1212857 A2 12-06-2002    |
|                                           |                     |                            | US 7454327 B1 18-11-2008    |
|                                           |                     |                            | US 2009076801 A1 19-03-2009 |
|                                           |                     |                            | US 2009138259 A1 28-05-2009 |
|                                           |                     |                            | US 2012065964 A1 15-03-2012 |
|                                           |                     |                            | WO 0126262 A2 12-04-2001    |
| -----                                     |                     |                            |                             |