

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization

International Bureau



(10) International Publication Number

WO 2014/001137 A1

(43) International Publication Date
3 January 2014 (03.01.2014)

WIPO | PCT

(51) International Patent Classification:

H04N 21/234 (2011.01) H04N 21/8549 (2011.01)
H04N 21/462 (2011.01) G06F 17/30 (2006.01)

VALLIER, Louis; 6,rue des Mimosas, F-35520 La Mézière (FR). PÈREZ, Patrick; 4 rue Eugène Quessot, F-35200 Rennes (FR). HELLIER, Pierre; 16, rue Angélique du Coudray, F-35235 Thorigé Fouillard (FR).

(21) International Application Number:

PCT/EP2013/062568

(74) Agent: SCHMIDT-UHLIG, Thomas; Deutsche Thomson OHG, European Patent Operations, Karl-Wiechert-Allee 74, 30625 Hannover (DE).

(22) International Filing Date:

18 June 2013 (18.06.2013)

(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, 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.

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

12305733.3 25 June 2012 (25.06.2012) EP

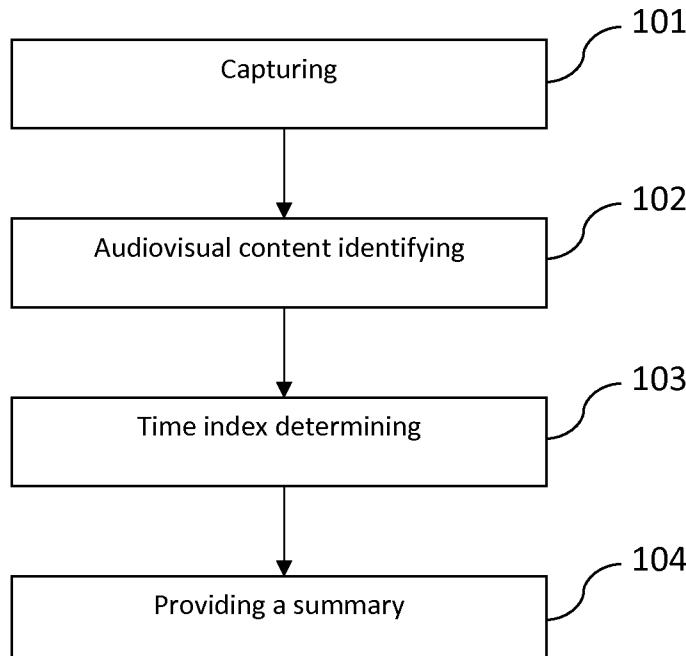
(71) Applicant: THOMSON LICENSING [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

(72) Inventors: OISEL, Lionel; Le petit Pellan, F-35137 La Nouaye (FR). ZEPEDA, Joaquin; 10 Allée Henri Bergson, F-35136 St.-Jacques-de-la-Lande (FR). CHE-

[Continued on next page]

(54) Title: SYNCHRONIZED MOVIE SUMMARY

Fig 1



(57) Abstract: The present invention relates to a method for providing (104) a summary of an audiovisual object. The method comprises the steps of: capturing (101) information from the audiovisual object; identifying (102) the audiovisual object; determining (103) the time index of the captured information relative to the audiovisual object; and providing (104) a summary of a portion of the identified audiovisual object, the portion being comprised between the beginning and the determined time index of the identified audiovisual object.

WO 2014/001137 A1



(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, 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))*

SYNCHRONIZED MOVIE SUMMARY

TECHNICAL FIELD

The present invention relates to a method for providing a summary of an audiovisual object.

5

BACKGROUND

It may occur that a viewer misses the beginning of an audiovisual object being played back. Facing with that problem, the viewer would like to know what is missed. The US 10 patent application 11/568,122 addresses this problem by providing an automatic summarization of a portion of a content stream for a program using a summarization function mapping the program to a new segment space and depending upon whether the content portion is a beginning, intermediate, or 15 ending portion of the content stream.

It is one object of the present invention to provide an end user a summary which is better tailored to the content the end user actually missed.

20 SUMMARY OF THE INVENTION

To this end, the present invention proposes a method for providing a summary of an audiovisual object, comprising the steps of:

- (i) capturing information from the audiovisual object that allows to identify the audiovisual object and allows to determine a time index relative to the audiovisual object;
- (ii) identifying the audiovisual object;

- (iii) determining the time index of the captured information relative to the audiovisual object; and
- (iv) providing a summary of a portion of the identified audiovisual object, the portion being comprised between the beginning and the determined time index of the identified audiovisual object.

The determination of the time index enables to precisely evaluate the portion of the audiovisual object which has been missed by a user, and to generate and to provide a summary tailored to the missed portion. As a result, the user is provided with a summary containing information relevant to what the user missed and bounded by the determined time index. For example, spoilers of an audiovisual object are not disclosed in the provided summary.

15 The invention also relates to a method, wherein:

a database comprising data of time-indexed images of the identified audiovisual object is provided;

the captured information is data of an image of the audiovisual object at the capturing time; and

20 the time index is determined upon a similarity matching between the data of the image of the audiovisual object at the capturing time and the data of the time-indexed images of the identified audiovisual object in the database.

25 Preferably, the nature of the data of the image of the audiovisual object and the nature of the data of the time-indexed images of the identified audiovisual object are of signature nature.

30 The advantage of using signatures, in particular, includes that the data become lighter than the raw data, and allow

therefore a quicker identifying as well as a quicker matching.

Alternatively, the invention relates to method, wherein:

a database comprising data of time-indexed audio signals of the identified audiovisual object is provided;

5 the captured information is data of an audio signal of the audiovisual object at the capturing time; and

10 the time index is determined upon a similarity matching between the data of the audio signal of the audiovisual object at the capturing time and the data of the time-indexed audio signals of the identified audiovisual object in the database.

Preferably, the nature of the data of the audio signal of the audiovisual object and the nature of the data of the time-indexed audio signals of the identified audiovisual object are of signature nature.

Advantageously, the step of capturing is performed by a mobile device.

Advantageously, the step of identifying, the step of determining and the step of providing are performed on a dedicated server.

This way, less processing power is required on the capturing side, and the process of providing a summary is accelerated.

For a better understanding, the invention shall now be explained in more detail in the following description with reference to the figures. It is understood that the invention is not limited to the described embodiments and that specified features can also expediently be combined and/or

modified without departing from the scope of the present invention as defined in the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

5 Figure 1 shows an exemplary flowchart of a method according to the present invention.

Figure 2 shows an example of an apparatus allowing the implementation of the method according to the present invention.

10

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Referring to Fig. 2, an exemplary apparatus configured to implement the method of the present invention is illustrated.

The apparatus comprises a rendering device 201, a capturing

15 device 202 and a database 204, and optionally, a dedicated server 205. A first preferred embodiment of the method of the present invention will be explained in more detail with reference to the flow chart in Fig. 1 and the apparatus in Fig. 2.

20 The rendering device 201 is used for rendering an audiovisual object. For example, the audiovisual object is a movie and the rendering device 201 is a display. Then, information of the rendered audiovisual object, e.g., data of an image of a movie being displayed, is captured 101 by a capturing device

25 202 equipped with capturing means. Such device 202 is for example a mobile phone equipped with a digital camera. The captured information is used for identifying 102 the audiovisual object and determining 103 a time index relative to the audiovisual object. Subsequently, a summary of a

30 portion of the identified audiovisual object is provided 104,

wherein the portion of the object is comprised between the beginning and the determined time index of the identified audiovisual object.

Specifically, the captured information, i.e. the data of an

5 image of the movie, is sent to a database 204, via for example a network 203. The database 204 comprises data of time-indexed images of the identified audiovisual objects, such as a set of movies in this preferred embodiment.

Preferably, the data of the image of the audiovisual object

10 and the data of the time-indexed images of the identified audiovisual object in the database are signatures of the images. For example, such a signature may be extracted using a key point descriptor, e.g. SIFT descriptor. Then, the steps of identifying 102 the audiovisual object and determining 103 the time index of the captured information is performed upon 15 a similarity matching between the data of the image of the audiovisual object at capturing time and the data of the time-indexed images in the database 204, i.e. between the signatures of the images. The most similar time-indexed image 20 in the database 204 for the image of the audiovisual object at capturing time is identified, allowing to identify the audiovisual object and to determine the time index of the captured information relative to the audiovisual object. Then a summary of a portion of the identified audiovisual object, 25 which is comprised between the beginning and the determined time index of the identified audiovisual object, is obtained and provided 104 to the user.

The data of the image of the audiovisual object, e.g., the image signature, can be captured either directly by the

30 capturing device 202 equipped with the capturing means or alternatively on a dedicated server 205. Similarly, the steps of identifying 102 the audiovisual object, determining 103

the time index of the captured information, and providing 104 a summary can be alternatively performed on a dedicated server 205.

An advantage of performing the image signature capture 5 directly on the device 202 is that the weight of the data sent to the dedicated server 205 is lighter in terms of memory.

An advantage of performing the signature capture on the dedicated server 205 is that the nature of the signature may 10 be controlled on the server side. Thus the nature of the signature of the image of the audiovisual object and the nature of the signatures of the time-indexed images in the database 204 are the same, and can be directly compared.

The database 204 can be located in the dedicated server 205.

15 It can of course also be located outside the dedicated server 205.

In the above preferred embodiment, the captured information is the data of an image. In a more general manner, the information can be any data that is able to be captured by a 20 capturing device 202 possessing the adapted capturing means, provided the captured data enables identifying 102 of the audiovisual object and determining 103 the time index of the captured information relative to the audiovisual object.

In a second preferred embodiment for the method of this 25 invention, the captured information is data of an audio signal of an audiovisual object at the capturing time. The information can be captured by a mobile device equipped with a microphone or a loudspeaker. The data of the audio signal of the audiovisual object can be a signature of the audio signal, which is then matched to the most similar audio 30 signature among the collection of audio signatures contained

in the database 204. The similarity matching is thus used for identifying 102 the audiovisual object and determining 103 the time index of the captured information relative to the audiovisual object. A summary of a portion of the identified 5 audiovisual object is subsequently provided 104, wherein the portion of the object is comprised between the beginning and the determined time index of the identified audiovisual object.

An example for the database 204 and a summary of a portion of 10 the identified audiovisual object will now be described. An offline process is performed in order to generate the database 204, with the help of existing and/or public database. An exemplary database for a collection of a set of movies will be explained now, but the invention is not 15 limited to the description below.

For the summary database of the database 204, a temporally synchronized summary of the full movie is generated. This relies, for example, on an existing synopsis, such as those available on the Internet Movie Database (IMDB). Such 20 synopsis may be retrieved directly from the name of the movie. Synchronization can be performed by synchronizing a textual description of a given movie with an audiovisual object of the given movie, by using for example a transcription of an audio track of the given movie. Then, a 25 matching of the words and concepts extracted from both the transcription and the textual description is performed, resulting in a synchronized synopsis for the movie. The synchronized synopsis may of course be obtained manually.

Optionally, additional information is also extracted. A face 30 detection and a clustering process are applied on the full movie, thus providing clusters of faces which are visible in the movie. Each of the clusters is composed of faces

corresponding to the same character. This clustering process may be performed using the techniques detailed in *M.*

Everingham, J. Sivic, and A. Zisserman "Hello! My name is... Buffy" - Automatic naming of characters in TV video"

5 *Proceedings of the 17th British Machine Vision Conference (BMVC 2006)*. A list of characters associated with a list of movie time codes associated to the presence of a particular character is then obtained. The obtained clusters may be matched against with an IMDB character list of the given

10 movie for a better clustering result. This matching process may comprise manual steps.

The obtained synchronized synopsis summary and the cluster lists are stored in the database 204. The movies in the database 204 are divided into a plurality of frames, and each 15 of the frames is extracted. The frames of the movie are then indexed for facilitating post-synchronization processes, such as determining 103 a time index of the captured information relative to the movie. Alternatively, instead of extracting each frame of the movie, only a part of the frames are

20 extracted by an adequate sub-sampling, in order to reduce the amount of data to be processed. For each extracted frame, an image signature, e.g., a fingerprint based on key point description, is generated. Those key points and their associated descriptions are indexed in an efficient way,

25 which may be done using the techniques described in *H. Jégou, M. Douze, and C. Schmid - Hamming embedding and weak geometric consistency for large scale image search - ECCV, October 2008*. The frames of the movies associated with the image signatures are then stored in the database 204.

30 To obtain the summary of a portion of an identified audiovisual object (i.e. a movie), information of the audiovisual object, e.g., data of an image thereof, is

captured by a capturing device 202. The information is then sent to the database 204, and compared to the database 204 for identifying the audiovisual object. For example, a frame of the movie is identified in the database 204 corresponding 5 to the captured information. The identified frame facilitates the matching between the captured information and the synchronized synopsis summary in the database 204, thus determining the time index of the captured information relative to the movie. A synchronized summary of a portion of 10 the movie is then provided to a user, wherein the portion of the movie is comprised between the beginning and the determined time index of the identified movie. For example, the summary can be provided by being displayed on the mobile device 202 and being read by the user. Optionally, the 15 summary can include cluster lists of characters appearing in the portion of the movie.

CLAIMS

1. A method for providing (104) a summary of an audiovisual object, comprising the steps of:

- (i) capturing (101) information from the audiovisual object that allows to identify the audiovisual object and allows to determine a time index relative to the audiovisual object;
- (ii) identifying (102) the audiovisual object;
- (iii) determining (103) the time index of the captured information relative to the audiovisual object; and
- (iv) providing (104) a summary of a portion of the identified audiovisual object, the portion being comprised between the beginning and the determined time index of the identified audiovisual object.

15

2. The method of claim 1, wherein:

a database (204) comprising data of time-indexed images of the identified audiovisual object is provided;

the captured information is data of an image of the

20 audiovisual object at the capturing time; and

the time index is determined upon a similarity matching between the data of the image of the audiovisual object at the capturing time and the data of the time-indexed images of the identified audiovisual object in the database (204).

25

3. The method of claim 2, wherein:

the nature of the data of the image of the audiovisual object and the nature of the data of the time-indexed images of the identified audiovisual object are of signature nature.

5 4. The method of claim 1, wherein:

a database (204) comprising data of time-indexed audio signals of the identified audiovisual object is provided;

the captured information is data of an audio signal of the audiovisual object at the capturing time; and

10 the time index is determined upon a similarity matching between the data of the audio signal of the audiovisual object at the capturing time and the data of the time-indexed audio signals of the identified audiovisual object in the database (204).

15

5. The method of claim 2, wherein:

the nature of the data of the audio signal of the audiovisual object and the nature of the data of the time-indexed audio signals of the identified audiovisual object are of signature

20 nature.

6. The method of any one of the aforementioned claims, wherein the step of capturing (101) is performed by a mobile device (202).

25

7. The method of any one of the aforementioned claims, wherein the step of identifying (102), the step of

determining (103) and the step of providing (104) are performed on a dedicated server (205).

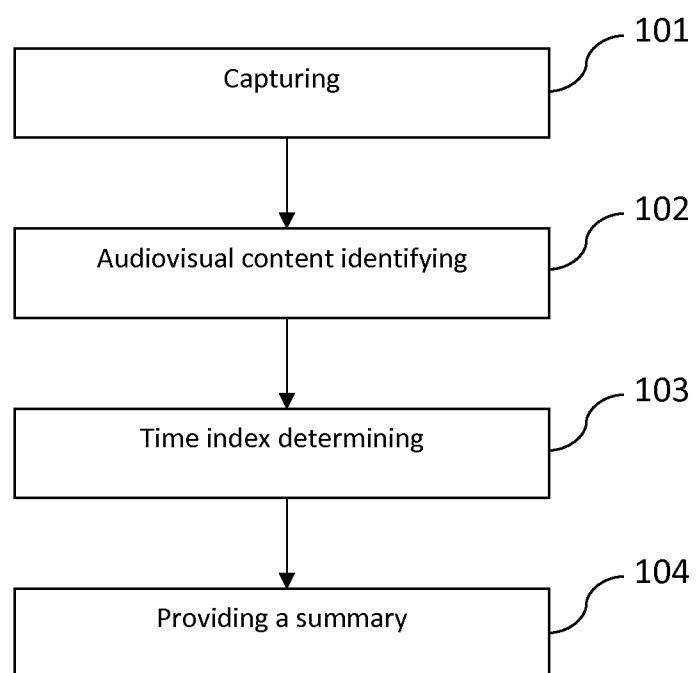


Fig 1

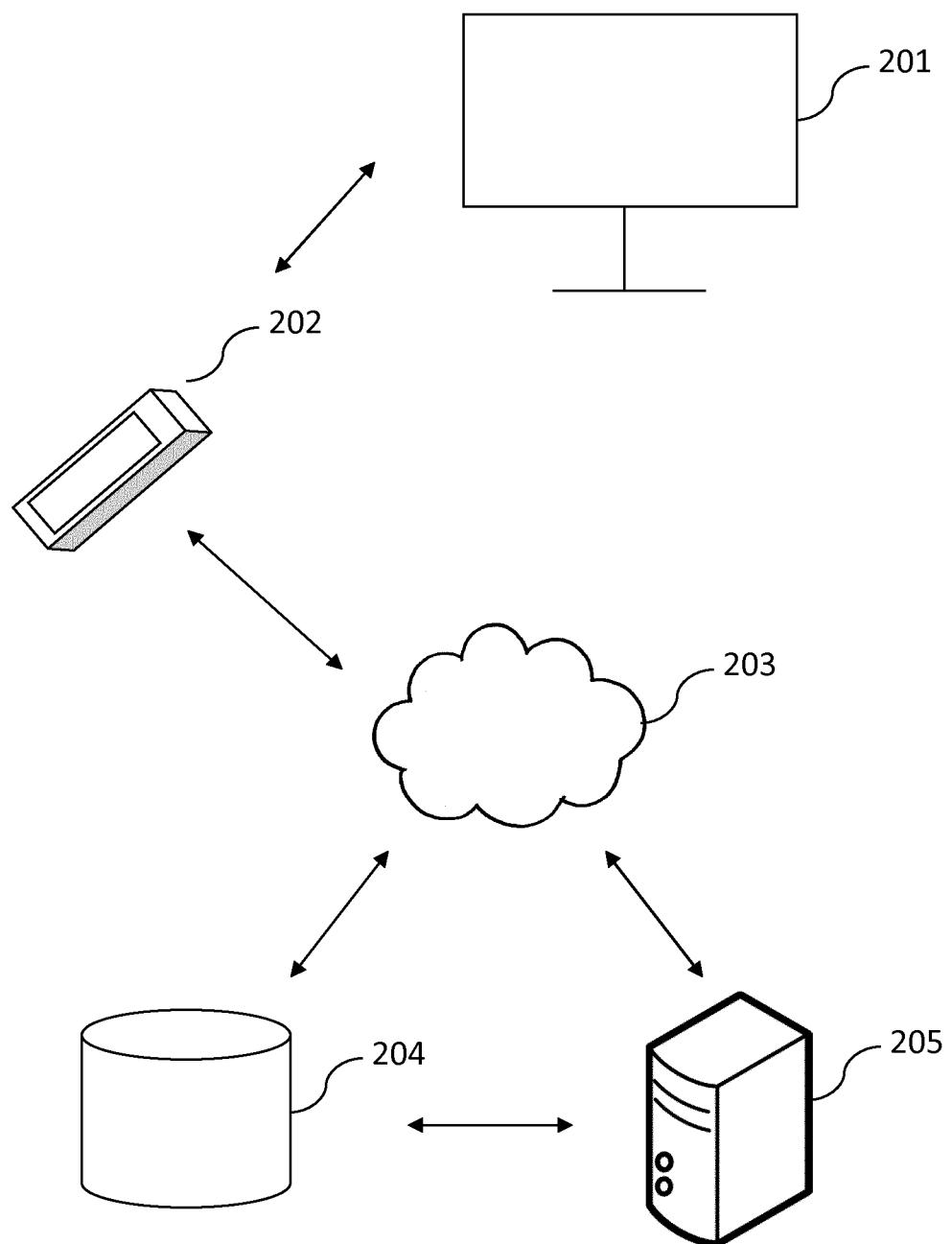


Fig 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/062568

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N21/234 H04N21/462 H04N21/8549 G06F17/30
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)
H04N G06F

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011/276157 A1 (WANG AVERY LI-CHUN [US] ET AL) 10 November 2011 (2011-11-10) paragraph [0034] - paragraph [0038] paragraph [0055] - paragraph [0059] paragraph [0067] - paragraph [0076] paragraph [0087] - paragraph [0088] figures 1-11 ----- WO 2005/103954 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; DIMITROVA NEVENKA [US]; BARBIERI) 3 November 2005 (2005-11-03) cited in the application page 6, line 5 - line 10 page 10, line 4 - line 12 figure 1 claim 1 ----- - / -	1-7
Y		1-7

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	Date of mailing of the international search report
29 August 2013	05/09/2013
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) 340-3046	Authorized officer Arpacı, Mutlu

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/062568

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2005/062610 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; BARBIERI MAURO [NL]; MEKENKAMP GE) 7 July 2005 (2005-07-07) abstract ----- US 2012/033876 A1 (MOMEYER BRIAN [US] ET AL) 9 February 2012 (2012-02-09) abstract -----	1-7
A		1-7

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No PCT/EP2013/062568

Patent document cited in search report	Publication date	Patent family member(s)			Publication date
US 2011276157	A1 10-11-2011	CA 2798093 A1			10-11-2011
		CN 102959543 A			06-03-2013
		EP 2567332 A1			13-03-2013
		KR 20130029082 A			21-03-2013
		US 2011276157 A1			10-11-2011
		WO 2011140269 A1			10-11-2011
<hr style="border-top: 1px dashed black;"/>					
WO 2005103954	A1 03-11-2005	CN 1977262 A			06-06-2007
		EP 1743258 A1			17-01-2007
		JP 2007534261 A			22-11-2007
		KR 20070001240 A			03-01-2007
		US 2009119704 A1			07-05-2009
		WO 2005103954 A1			03-11-2005
<hr style="border-top: 1px dashed black;"/>					
WO 2005062610	A1 07-07-2005	CN 1894964 A			10-01-2007
		EP 1698174 A1			06-09-2006
		JP 2007519321 A			12-07-2007
		KR 20060126508 A			07-12-2006
		US 2007109443 A1			17-05-2007
		WO 2005062610 A1			07-07-2005
<hr style="border-top: 1px dashed black;"/>					
US 2012033876	A1 09-02-2012	CN 103052961 A			17-04-2013
		EP 2601616 A2			12-06-2013
		KR 20130060280 A			07-06-2013
		US 2012033876 A1			09-02-2012
		WO 2012019163 A2			09-02-2012