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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
4 November 2010 (04.11.2010)

PCT

(10) International Publication Number
WO 2010/124926 A1

(51) International Patent Classification:
H04N 13/00 (2006.01)

(21) International Application Number:
PCT/EP2010/054773

(22) International Filing Date:
12 April 2010 (12.04.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09305372.6 29 April 2009 (29.04.2009) EP

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BOISSON, Guillaume** [FR/FR]; 12 rue Jean Malo-Renault, F-35000 Rennes (FR). **KERBIRIOU, Paul** [FR/FR]; 9 rue Yves Kerguelen, F-35235 Throrigne Fouillard (FR). **LE MEUR, Olivier** [FR/FR]; La Souche, F-35160 Talensac (FR).

(74) Agent: **RITTNER, Karsten**; 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, 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, PE, PG, PH, PL, PT, RO, RS, RU, 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, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, 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))*

(54) **Title**: METHOD AND A DEVICE FOR GENERATING AND PROCESSING DEPTH DATA

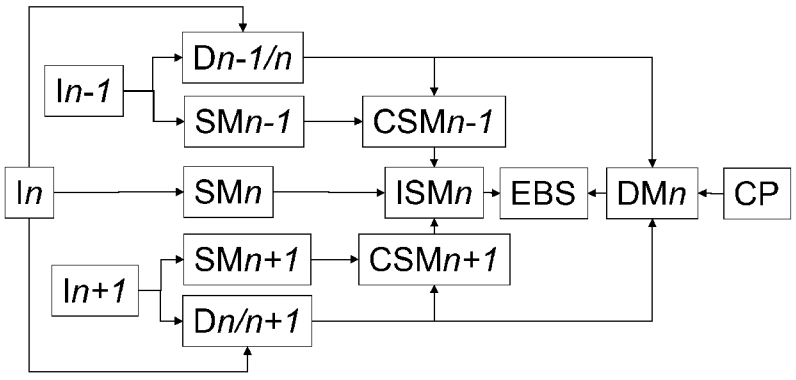


Fig. 1

(57) **Abstract**: The invention proposes a method for generating and processing depth data associated with a mapping of first image data comprised in a first image onto second image data comprised in at least a second image. Said method comprising the steps of generating a control signal which depends on a first saliency value determined using the first image data, wherein said control signal further depends on at least a second saliency value determined using the second image data, generating depth data using said mapping and processing the depth data wherein generating and/or processing of depth data is dependent on said control signal. This allows for more accurate generation and/or processing of the depth data in regions which are visually attracting and thus more likely in the focus of a viewer's attention.

WO 2010/124926 A1

Method and a Device for generating and processing depth data

Technical Field

The invention is related to a method and a device for generating and processing depth data. Said depth data is associated with a mapping of first image data comprised in a first image onto second image data comprised in at least a second image.

Background

An image is a two dimensional (2D) projection on an image plane located at a view point wherein the view point has a location and an orientation.

Due to the projection, some or all depth information comprised in a scene depicted by an image is lost.

It is known that a second image may be used for preserving the depth information, implicitly, and for allowing a viewer to vary the orientation within a small range. That is, a scene is stereoscopically depicted by help of two images taken from two slightly different view points, like the eyes of a human, the difference depending on the focal length used. One of the images is provided to the right eye and the other is provided to the left eye. This allows a viewer for experiencing the depth of the scene and to vary the orientation of the view point within said small range. But the location of the viewpoint remains fixed.

There is a trend in the art towards flexible viewpoint applications like free viewpoint TV (FTV) which allow a viewer to freely vary not only the orientation of the viewpoint but also the location.

For making the location of the viewpoint freely variable as well, depth data has to be available. Depth data is an explicit representation of depth information, for instance in form of a depth map, in form of a three-dimensional model of structures comprised in a depicted scene or in form of a set of layered images or depth planes.

Using one or more images and depth data, views of a scene from arbitrary located and/or orientated viewpoints may be generated automatically in response to a viewer's input determining a desired viewpoint.

Depth data may be collected using dedicated depth scanning devices based on laser ranging, ultrasonic ranging or the like. Another way of generating depth data is based on a mapping between images without using dedicated depth sensors. Such mapping -also known as disparity mapping- is based on the fact that relative dislocation between a place of depiction of a feature in one of the images and a place of depiction of the same feature in the other of the images is reciprocally proportional to the distance of said feature to the images.

Depth data may be corrupted in different ways. For instance, due to the reciprocal relationship minor errors in determining disparities of closed by features may result in severely erroneous depth data. Other sources of corruption are quantization during encoding and packet loss during transmission.

Invention

The inventors recognized that such corruptions more or less impede the viewer's viewing experience in dependency on the degree of attention the viewer spends on image aspects generated using said corrupted depth data.

Therefore, the invention proposes a method for generating and processing depth data associated with a mapping of

first image data comprised in a first image onto second image data comprised in at least a second image, said method comprising the features of independent claim 1. A corresponding device is proposed which comprises the features of independent claim 6.

Said method comprises the steps of generating a control signal which depends on a first saliency value determined using the first image data, wherein said control signal further depends on at least a second saliency value determined using the second image data, generating depth data using said mapping and processing the depth data wherein generating and/or processing of depth data is dependent on said control signal.

Making generation and/or processing of depth data dependent on saliency values which are determined for that image data whose mapping is associated with said depth data allows for more accurate generation and/or processing of the depth data in regions which are visually attracting and thus more likely in the focus of a viewer's attention.

Further advantageous embodiments of the method are characterized by the features of one of the claims depending on claim 1 and further advantageous embodiments of the device are characterized by the features of one of the claims depending on claim 6.

Drawings

Exemplary embodiments of the invention are illustrated in the drawings and are explained in more detail in the following description.

In the figures:

Fig. 1 depicts an exemplarily flow chart of generating and processing of a depth map which may be used in single frame plus depth (2D+z) or a multiview plus depth (MVD) encoding, for instance, and

Fig. 2 depicts an exemplarily flow chart of generating and processing of a mesh-based three dimensional model, for instance.

Exemplary embodiments

An exemplarily flow chart of generating and processing of an integrated depth map DM*n* according to an exemplary embodiment of the inventive principle is depicted in Fig. 1.

The integrated depth map DM*n* is based on camera parameters CP and generated from disparity maps D*n*-1/*n* and D*n*+1/*n* which are generated from image I*n* and image I*n*-1, respectively, from image I*n* and image I*n*+1.

Disparity maps D*n*-1/*n* and D*n*+1/*n* are not only used for generating the integrated depth map DM*n* but also for warping or compensating saliency maps S*n*+1 and S*n*-1 which are saliency maps generated from image I*n*+1 and image I*n*-1. Compensation or warping results in compensated saliency maps CS*n*+1 and CS*n*-1 which are aligned with saliency map S*n* generated from image I*n* and therefore can be fused with or integrated into said saliency map S*n* generated from image I*n* which results in an integrated saliency map ISM*n*.

Thus, there are the integrated depth map DM*n* for image *n* as well as an integrated saliency map which allows for a region-of-interest based encoding of the integrated depth map DM*n* into an encoded bitstream EBS.

Another exemplarily flow chart of the inventive principle is depicted in Fig. 2.

From a set SoV of two or more views another set SoSM of a corresponding number of saliency maps is generated. The generated set SoSM of saliency maps as well as the set SoV of views are used for extracting one or more three dimensional models TdVM of one or more objects or structures at least partly depicted in the views of the set SoV of views.

The modelling is based on depth information determined using two or more views of said set SoV of views and controlled by saliency information determined using corresponding saliency maps of said set SoSM of saliency maps. That is, the more salient a surface region of a modelled object or structure is, the more detailed said surface region is modelled. For instance, in a mesh model the number of facets, nodes and/or edges is increased such that the size of facets in salient areas is reduced and approximation of the surface of the modelled object or structure is more precise in salient areas.

Commonly, modelling is based on disparities of pairs of image regions of which one is comprised in a first view of the modelled object or structure and one is comprised in a second view of the modelled object or structure wherein said second view depicts the object or scene from a different viewpoint than said first view and the images regions of each pair are determined as depicting a same aspect of the object or structure to-be-modelled.

The extracted three dimensional models TdM may then be encoded into an encoded bitstream EBS wherein a region-of-interest based encoding is used which is based on one or more three dimensional saliency models TdSM of the modelled objects or structures.

The claimed method may be implemented on a processing device comprised in a notebook, a desktop computer, a server or any other computing device. The computing device

may further comprise a storage medium like a hard disk, a BluRay-disk or a flash memory carrying image data of images.

The processing device may be operated as a saliency value determining device which uses image data for determining corresponding saliency values.

The processing device may further be operated as a control signal generator which generates control signals in dependency on determined saliency values.

And, the processing device may further be operated as a mapping device which maps first image data comprised in a first image onto second image data comprised in at least a second image.

The processing device may also be operated as a depth data generator which generates depth data in dependency on determined mappings and the generated control signal and/or the processing device may also be operated as a depth data processor which processes depth data in dependency on the generated control signal.

The exemplary embodiments of the invention are described for illustrative purposes only and shall not be construed as limiting the extent of projection as requested which is solely defined by the independent claims.

Claims

1. Method for generating and processing depth data associated with a mapping of first image data comprised in a first image onto second image data comprised in at least a second image, wherein processing of depth data comprises encoding of the depth data and said method comprises the steps of

- generating a control signal which depends on a first saliency value determined using the first image data, wherein said control signal further depends on at least a second saliency value determined using the second image data,
- generating depth data using said mapping and
- processing the depth data wherein
- encoding of depth data is dependent on said control signal.

2. Method according to claim 1, wherein generating said depth data is also dependent on said control signal and comprises generating one or more nodes of a mesh model and assigning one or more further saliency values to the generated nodes with the number of generated nodes and/or values of the further saliency values being dependent on said control signal.

3. Method according to one of the preceding claims, wherein processing of the depth data further comprises lossy compression of the depth data and said control signal is further used for controlling a degree of compression loss.

4. Method according to one of the preceding claims, wherein encoding of the depth data comprises ordered encoding of the depth data together with further depth data in a code sequence and said control signal is used for determining a rank in the encoding order.

5. Method according to one of the preceding claims, wherein the control signal is generated dependent on a sum of said first saliency value and said at least a second saliency value.

6. Device for generating and processing depth data associated with a mapping of first image data comprised in a first image onto second image data comprised in at least a second image, wherein processing comprises encoding of the depth data and said device comprises

- means for generating a control signal which depends on a first saliency value determined using the first image data, wherein said control signal further depends on at least a second saliency value determined using the second image data,
- means for generating depth data using said mapping,
- means for processing depth data and
- a controller adapted for using said control signal for controlling said encoding of the depth data.

7. Device according to claim 6, wherein said means for generating depth data is adapted for using said control signal for controlling said generating one or more nodes of a mesh model and assigning one or more further saliency

values to the generated nodes with the controller being adapted for controlling the number of generated nodes and/or values of the further saliency values in dependence on said control signal.

8. Device according to one of the claims 6 or 7, wherein processing of the depth data further comprises lossy compression of the depth data and the controller is further adapted for using said control signal for controlling a degree of compression loss.

9. Device according to one of the claims 6-8, wherein encoding of the depth data comprises ordered encoding of the depth data together with further depth data in a code sequence and the controller is adapted for using said control signal for determining a rank in an encoding order.

10. Device according to one of the claims 6-9, wherein means for generating a control signal is adapted for generating the control signal dependent on a sum of said first saliency value and said at least a second saliency value.

11. Device according to claim 10 or method according to claim 5, wherein said sum is a weighted sum.

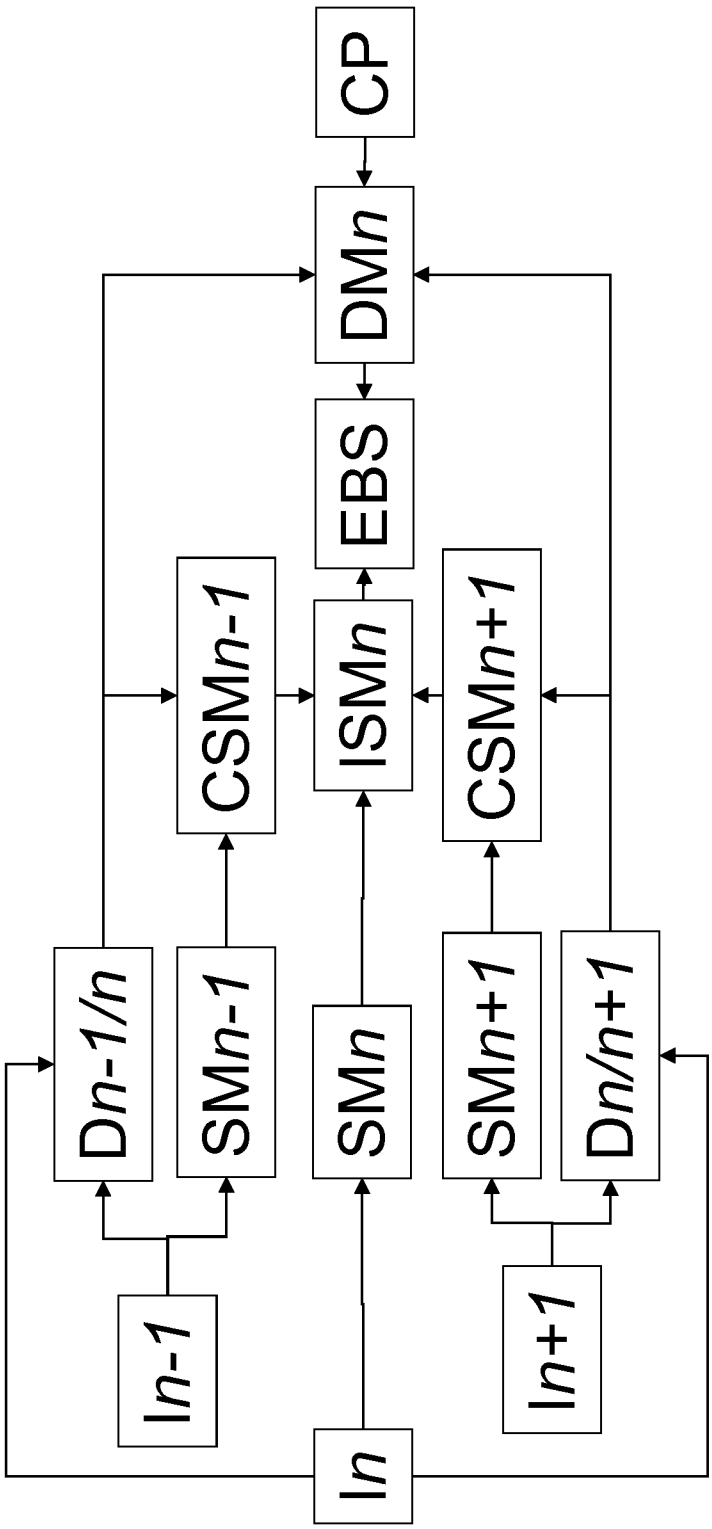


Fig. 1

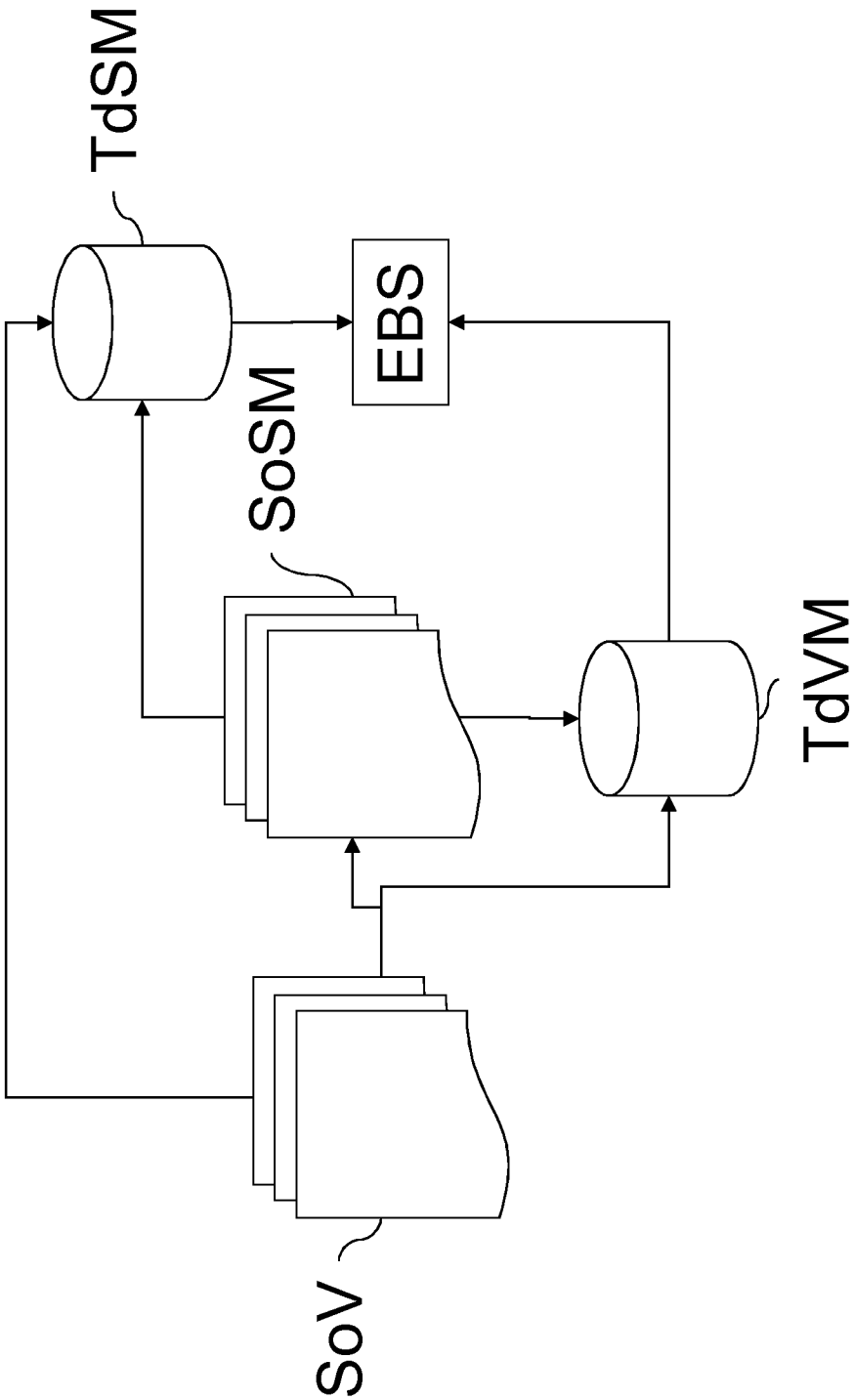


Fig. 2

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2010/054773

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N13/00
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)
G06T H04N

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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FELDMANN I ET AL: "REAL-TIME SEGMENTATION FOR ADVANCES DISPARITY ESTIMATION IN IMMERSIVE VIDEOCONFERENCE APPLICATIONS" JOURNAL OF WSCG, VACLAV SKALA-UNION AGENCY, PILZEN, CZ, vol. 10, no. 1, 1 January 2002 (2002-01-01), pages 171-178, XP008017665 ISSN: 1213-6972 page 172 - page 177	1-11
A	EP 1 544 792 A (THOMSON LICENSING SA [FR]) 22 June 2005 (2005-06-22) * abstract	1,6

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" 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 June 2010	07/07/2010

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Cooke, Edward

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2010/054773

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YONGQING ZENG: "Perceptual segmentation algorithm and its application to image coding" IMAGE PROCESSING, 1999. ICIP 99. PROCEEDINGS. 1999 INTERNATIONAL CONFE RENCE ON KOBE, JAPAN 24-28 OCT. 1999, PISCATAWAY, NJ, USA,IEEE, US, vol. 2, 24 October 1999 (1999-10-24), pages 820-824, XP010369077 ISBN: 978-0-7803-5467-8 * abstract --- LE MEUR O ET AL: "Efficient Saliency-Based Repurposing Method" IMAGE PROCESSING, 2006 IEEE INTERNATIONAL CONFERENCE ON, IEEE, PI, 1 October 2006 (2006-10-01), pages 421-424, XP031048663 ISBN: 978-1-4244-0480-3 * abstract ---	3,8 1,6

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2010/054773

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 1544792 A	22-06-2005	BR PI0417594 A	20-03-2007
		CN 1894721 A	10-01-2007
		DE 602004006306 T2	10-01-2008
		WO 2005059832 A1	30-06-2005
		JP 2007515009 T	07-06-2007
		US 2007116361 A1	24-05-2007