

(51) International Patent Classification:
G10K 15/04 (2006.01) *H04N 7/14* (2006.01)(21) International Application Number:
PCT/SG2011/000013(22) International Filing Date:
11 January 2011 (11.01.2011)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
12/698,453 2 February 2010 (02.02.2010) US(71) Applicant (for all designated States except US): **CRE-
ATIVE TECHNOLOGY LTD** [SG/SG]; 31 International
Business Park, Creative Resource, Singapore 609921
(SG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **SIM, Wong, Hoo**
[SG/SG]; 5 Chestnut Terrace, Singapore 679123 (SG).
TOO, Yew, Teng [SG/SG]; Block 628a, Woodlands,
Ring Road #07-280, Singapore 731682 (SG). **GOH,
Tieh, Cheng** [SG/SG]; C/o Creative Technology Ltd, 31
International Business Park, Creative Resource, Singa-
pore 609921 (SG).(74) Agent: **ATMD BIRD & BIRD LLP**; 2 Shenton Way
#18-01, SGX Centre 1, Singapore 068804 (SG).(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 (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, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

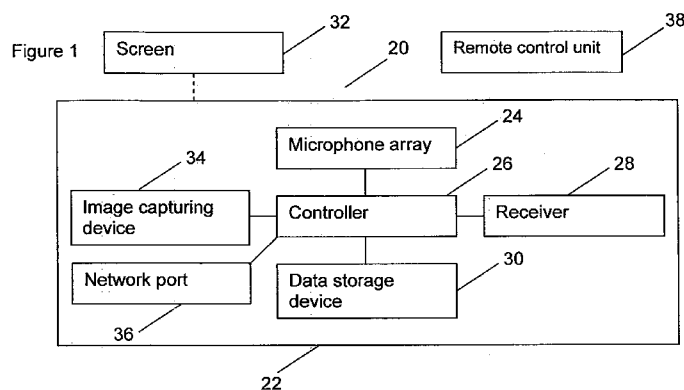
— as to applicant's entitlement to apply for and be granted
a patent (Rule 4.17(ii))

— of inventorship (Rule 4.17(iv))

Published:

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

(54) Title: AN APPARATUS FOR ENABLING KARAOKE

(57) Abstract: There is provided an apparatus for enabling karaoke. The apparatus includes a casing for the apparatus; a micro-
phone array incorporated within the casing, the microphone array being concealed from a user by the casing; a controller coupled
to the microphone array to at least process audio signals from the user input at the microphone array; a data storage device cou-
pled to the controller, the data storage device being for storing pre-recorded songs used for karaoke, and for storing songs as sung
by the user; and an image capturing device coupled to the controller, the image capturing device being for capturing images of the
user while the user is singing, with the captured images of the user being stored on the data storage device.

AN APPARATUS FOR ENABLING KARAOKE

FIELD OF INVENTION

- 5 This invention relates to an apparatus that facilitates the activity of karaoke for a user, particularly in relation to ease of access to the activity of karaoke for the user.

BACKGROUND

- 10 Nowadays, it is typical for many people to associate the activity of karaoke to be an activity which involves the setting up of voice manipulation/processing type of equipment together with other audio visual equipment. As such, there exists a general perception that the activity of karaoke involves a rather
15 complicated set-up process involving the functional connection of several pieces of hardware, and this consequently has dissuaded people from taking up the activity altogether.

- It should be noted that there have been devices which attempt to address the
20 aforementioned complexity concerns pertaining to the activity of karaoke, but such devices have not been popular due to issues relating to, for example, undesirable user experience, limited device functionality in relation to the activity of karaoke, lack of portability, and the like. In view of the
25 aforementioned, it is evident that there is a dearth of equipment available to users which facilitates the activity of karaoke both in a convenient manner for users and in a manner which provides a positive experience for the users.

The present invention aims to address the absence of such a device and correspondingly meet the unresolved needs of the users.

30

SUMMARY

There is provided an apparatus for enabling karaoke. The apparatus includes a casing for the apparatus; a microphone array incorporated within the casing,

the microphone array being concealed from a user by the casing; a controller coupled to the microphone array to at least process audio signals from the user input at the microphone array; a data storage device coupled to the controller, the data storage device being for storing pre-recorded songs used for karaoke, and for storing songs as sung by the user; and an image capturing device coupled to the controller, the image capturing device being for capturing images of the user while the user is singing, with the captured images of the user being stored on the data storage device.

It is advantageous that the microphone array detects the audio signals from the user when the user is within a predefined field of detection of the microphone array. Preferably, the microphone array includes either omnidirectional microphones or uni-directional microphones. The microphone array may be configured such that the predefined field of detection is spread over a wide area.

The apparatus may further include a receiver coupled to the controller to receive audio signals from the user using a wireless microphone.

It is preferable that the controller also performs a method for processing the pre-recorded songs. The method may include assessing audible lyrics of the pre-recorded songs; converting the audible lyrics to text for display on a screen connected to the apparatus; and muting audible lyrics from the pre-recorded songs. The method may also further include controlling audio output of the pre-recorded songs; and providing the audio output of the pre-recorded songs. The controller may control the audio output by adjusting parameters of audio output such as, for example, voice echo, music volume, voice volume, music pitch and so forth. The parameters of audio output may be adjusted using either a remote control device or voice control. The controller may also perform echo cancellation for secondary audio input at the microphone array, the secondary audio input being playback of the user's voice during the activity of karaoke.

The image capturing device may capture images of the user when the user is within a predefined field of capture of the image capturing device. The image capturing device may be either integrated within the casing or separate from the apparatus. The captured images of the user may be used for creating a video.

Preferably, the casing protects components of the apparatus.

The apparatus may further including a network port coupled to the controller, the network port enabling the connection of the apparatus to a network.

It is preferable that the captured images of the user singing and a recording of the singing by the user are merged into a music video. The music video may be transmitted over the network and distributed to either at least one recipient or a central video depository for storage. Advantageously, access to the stored videos on the central video depository by third parties may enable playback, and grading of the stored videos by the third parties.

DESCRIPTION OF FIGURES

In order that the present invention may be fully understood and readily put into practical effect, there shall now be described by way of non-limitative example only preferred embodiments of the present invention, the description being with reference to the accompanying illustrative drawings.

Figure 1 shows a schematic view of the apparatus of the present invention. Figure 2 shows a process flow of a controller of the apparatus of Figure 1 when the controller is processing audio.

DESCRIPTION OF PREFERRED EMBODIMENTS

With reference to Figure 1, there is provided an apparatus 20 for enabling karaoke. It should be appreciated that the apparatus 20 is directed at enabling convenience for a user in relation to the activity of karaoke but not at the

expense of the user experience. The apparatus 20 may be a standalone unit which may be portable without a need for much assembly/disassembly of parts of the apparatus 20 to the user as this also aids in the convenience to the user.

5

The apparatus 20 includes a casing 22 for the apparatus 20. The casing 22 may protect components of the apparatus 20 such that the apparatus 20 may be considered to be durable and robust. The casing 22 may be made from a either a metal or a plastic material as both materials provide strength without incurring an excessive penalty in relation to weight. The casing 22 may also incorporate a handle to aid in the portability of the apparatus 20.

There is a microphone array 24 incorporated within the casing 22 of the apparatus 20. The microphone array 24 may be concealed from the user by the casing 22, where the concealment of the microphone array 24 may be to protect the microphone array 24 from damage and may also be for maintaining aesthetic appearances of the apparatus 20. The microphone array 24 may detect the sound signals from the user when the user is within a predefined field of detection of the microphone array 24. It should be appreciated that the predefined field of detection of the microphone array 24 may be dependent on the technical specifications of microphones employed in the microphone array 24.

The microphone array 24 includes either omni-directional microphones or uni-directional microphones. The microphone array 24 may be able to capture audio input from either a spherical region or a hemispherical region around the apparatus 20 regardless of the type of microphones used in the microphone array 24. It is evident that using omni-directional microphones would provide a spherical region for capturing audio input, but it should be noted that varying a facing of uni-directional microphones may also provide a spherical region for capturing audio input. The microphone array 24 may be configured such that the predefined field of detection is spread over a wide area. When the predefined field of detection is spread over a wide area, a "sweet spot" of audio detection for the microphone array 24 is also

correspondingly enlarged. This allows the user to be moving around while singing during the activity of karaoke and also enables audio signals to be able to be picked up from more positions around the apparatus 20. In an enclosed environment like a cabin of a vehicle, passengers in the cabin of the vehicle may participate in the activity of karaoke regardless of their seating position in the cabin and regardless of the location of the apparatus 20 in the cabin of the vehicle. It should be appreciated that the apparatus 20 may be employed in any enclosed environment, with an effective range of the apparatus 20 being dependent on the predefined field of detection for the microphone array 24.

The apparatus 20 should also include a controller 26 coupled to the microphone array 24 to at least process audio signals from the user input at the microphone array 24. The controller 26 may process the audio signals from the user input at the microphone array 24. The processing carried out by the controller 26 may include echo cancellation for secondary audio input at the microphone array, the secondary audio input being playback of the user's voice during the activity of karaoke. This ensures that a resultant post-processed output of the user is not processed again by the controller 26 as this would distort the intended audio output by the apparatus 20.

Furthermore, the controller 26 also performs a method 50 for processing pre-recorded songs used for karaoke as shown in Figure 2. The method 50 advantageously allows any audio file to be used for the activity of karaoke. This is because the method 50 is able to generate "sing-along" text for the user even if there is no such text at an initial juncture. In this regard, the user is able to sing along to any audio file, which provides the user with a greater variety compared to being restricted to typical karaoke video sources.

The method 50 includes assessing audible lyrics of the pre-recorded songs (52) and converting the audible lyrics to text (54) for display on a screen 32 connected to the apparatus 20. The screen 32 may be either a device which is separate from the apparatus 20 or incorporated in the apparatus 20. The

screen 32 may be, for example, an LCD panel, a TFT panel, a plasma panel, and so forth.

It should be appreciated that the assessment and conversion of the audible lyrics may be carried out using known speech-to-text conversion techniques and an actual speech-to-text conversion technique may be decided upon based on a processing capability of the controller 26. The method 50 also includes muting the audible lyrics (56) from the pre-recorded songs such that only the musical track of the pre-recorded songs is played back for the user for the user to sing-along to.

In addition, the method 50 also further includes controlling audio output (58) of the pre-recorded songs; and providing the audio output of the pre-recorded songs (60) to the user. The audio output may be controlled by adjusting parameters of audio output such as, for example, voice echo, music volume, voice volume, music pitch and so forth. The parameters of audio output may be adjusted using either a remote control device 38 or voice control. Voice control may be in the form of firstly, identification of control keywords such as, for example, echo, music volume, voice volume, pitch, and the like followed by identification of action keywords such as, for example, up, down, more, less, higher, lower and so forth. In order to avoid inadvertent adjustment of audio output parameters during use of voice control, it should be appreciated that voice control may be carried out only when the user is not singing as both the action keywords and control keywords may be found in the lyrics of a song being sung by the user.

The apparatus 20 may also include a data storage device 30 coupled to the controller 26, where the data storage device 30 may be for storing the pre-recorded songs used for karaoke, and for storing songs as sung by the user.

The data storage device 30 may be non-volatile memory, and may be either in a form of a hard disk drive or in a form of a flash memory module. The storage of songs sung by the user may be to aid the user in assessing/reviewing the user's singing at a later juncture. It is also convenient for the user that the user's singing may be stored without using additional equipment, and that

recording of the user's singing may be stored concurrently while the user is singing.

There may also be an image capturing device 34 coupled to the controller 26, whereby the image capturing device 34 may be for capturing images of the user when the user is singing, with the captured images of the user being stored on the data storage device 30. The image capturing device 34 may be either able to capture wide angle images which allow zooming into the user's image after requisite wide angle image correction processes are applied, or able to track the user's movements such that the user's expressions are fully captured during a duration when the user is singing. It should be noted that the image capturing device 34 may capture images of the user when the user is within a predefined field of capture (ie. field of view) of the image capturing device 34. It should also be noted that the image capturing device 34 may be either incorporated within the casing 22 or it may be a separate device that is able to be coupled to the controller 26 of the apparatus 20. The captured images of the user may subsequently be used for creating a video during an instance(s) while the user is singing. The video may merge images of the user while singing together with a recording of the singing by the user in a form of a music video.

The apparatus 20 may further include a receiver 28 coupled to the controller 26 to receive audio signals from the user using a wireless microphone. It should be appreciated that when the receiver 28 receives signals, the controller 26 would subsequently disable the microphone array 24, such that audio input is only input through the wireless microphone. It should be noted that a mobile phone may function as the wireless microphone.

Finally, the apparatus 20 may include a network port 36 coupled to the controller, where the network port 36 allows the apparatus 20 to be connectable to a network. The network port 36 may enable the apparatus 20 to be connectable to a network via either a cable connection or a wireless connection. When the apparatus 20 is connected to a network, the aforementioned video created using the captured images of the user may be

transmitted over the network and distributed to either at least one recipient or a central video depository for storage. Access to the stored videos on the central video depository by third parties may enable at least playback, and grading of the stored videos by the third parties.

5

In view of the aforementioned, it should be appreciated that the apparatus 20 is able to meet the needs of users who wish to have an apparatus which allows convenient access to the activity of karaoke. However, it is evident that there are many aspects to the present invention which would further enhance the user experience during the activity of karaoke which would also appeal to the users of the apparatus 20.

10

Whilst there has been described in the foregoing description preferred embodiments of the present invention, it will be understood by those skilled in the technology concerned that many variations or modifications in details of design or construction may be made without departing from the present invention.

15

CLAIMS

1. An apparatus for enabling karaoke, the apparatus including:
 - a casing for the apparatus;
 - 5 a microphone array incorporated within the casing, the microphone array being concealed from a user by the casing;
 - a controller coupled to the microphone array to at least process audio signals from the user input at the microphone array;
 - a data storage device coupled to the controller, the data storage device
 - 10 being for storing pre-recorded songs used for karaoke, and for storing songs as sung by the user; and
 - an image capturing device coupled to the controller, the image capturing device being for capturing images of the user while the user is singing, with the captured images of the user being stored on the data storage
 - 15 device;
 - wherein the microphone array detects the audio signals from the user when the user is within a predefined field of detection of the microphone array.
2. The apparatus of claim 1, wherein the microphone array includes either
- 20 omni-directional microphones or uni-directional microphones.
3. The apparatus of claim 1, further including a receiver coupled to the controller to receive audio signals from the user using a wireless microphone.
- 25 4. The apparatus of claim 1, wherein the controller also performs a method for processing the pre-recorded songs, the method including:
 - assessing audible lyrics of the pre-recorded songs;
 - converting the audible lyrics to text for display on a screen connected
 - to the apparatus; and
 - 30 muting audible lyrics from the pre-recorded songs.
5. The apparatus of claim 4, wherein the controller performs the method which further includes:

controlling audio output of the pre-recorded songs; and
providing the audio output of the pre-recorded songs.

6. The apparatus of claim 5, wherein the controller controls the audio
output by adjusting parameters of audio output selected from a group
including: voice echo, music volume, voice volume, and music pitch.

7. The apparatus of claim 6, wherein the parameters of audio output are
adjusted using either a remote control device or voice control.

8. The apparatus of claim 1, wherein the image capturing device captures
images of the user when the user is within a predefined field of capture of the
image capturing device.

9. The apparatus of claim 1, wherein the captured images of the user is
used for creating a video.

10. The apparatus of claim 1, wherein the casing protects components of
the apparatus.

11. The apparatus of claim 1, wherein the controller performs echo
cancellation for secondary audio input at the microphone array, the secondary
audio input being playback of the user's voice during the activity of karaoke.

12. The apparatus of claim 1, wherein the microphone array is configured
such that the predefined field of detection is spread over a wide area.

13. The apparatus of claim 1, wherein the image capturing device is either
integrated within the casing or separate from the apparatus.

14. The apparatus of claim 1, further including a network port coupled to
the controller, the network port enabling the connection of the apparatus to a
network.

15. The apparatus of claim 14, wherein the captured images of the user singing and a recording of the singing by the user are merged into a music video.

5 16. The apparatus of claim 15, wherein the music video is transmitted over the network and distributed to either at least one recipient or a central video depository for storage.

10 17. The apparatus of claim 16, wherein access to the stored videos on the central video depository by third parties may enable playback, and grading of the stored videos by the third parties.

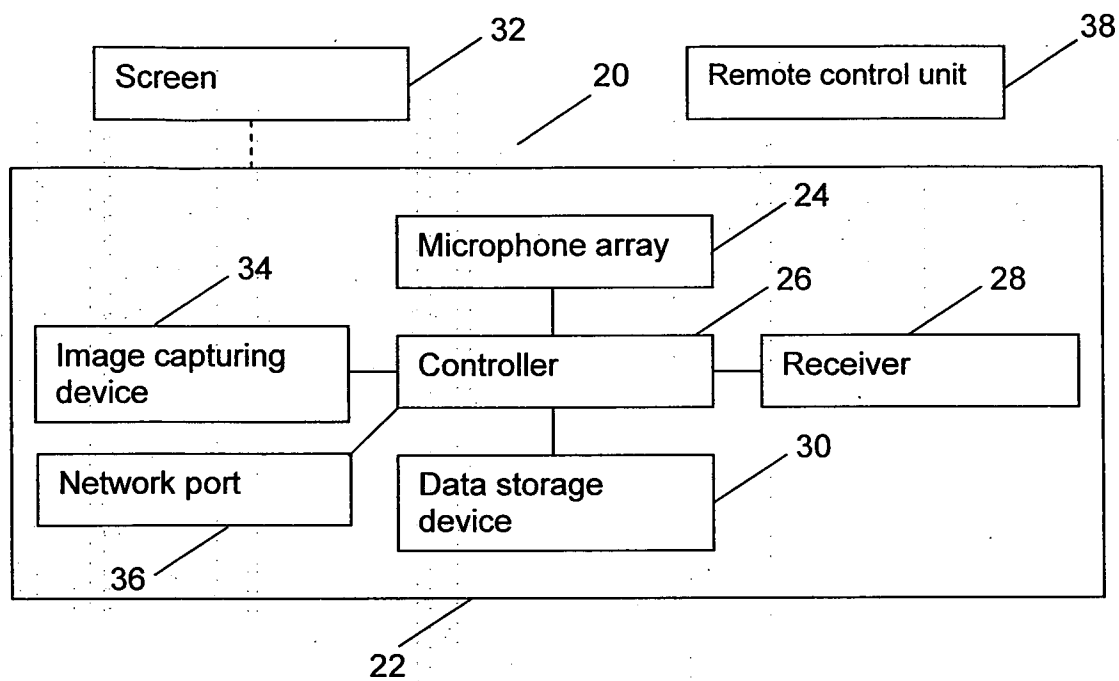


Figure 1

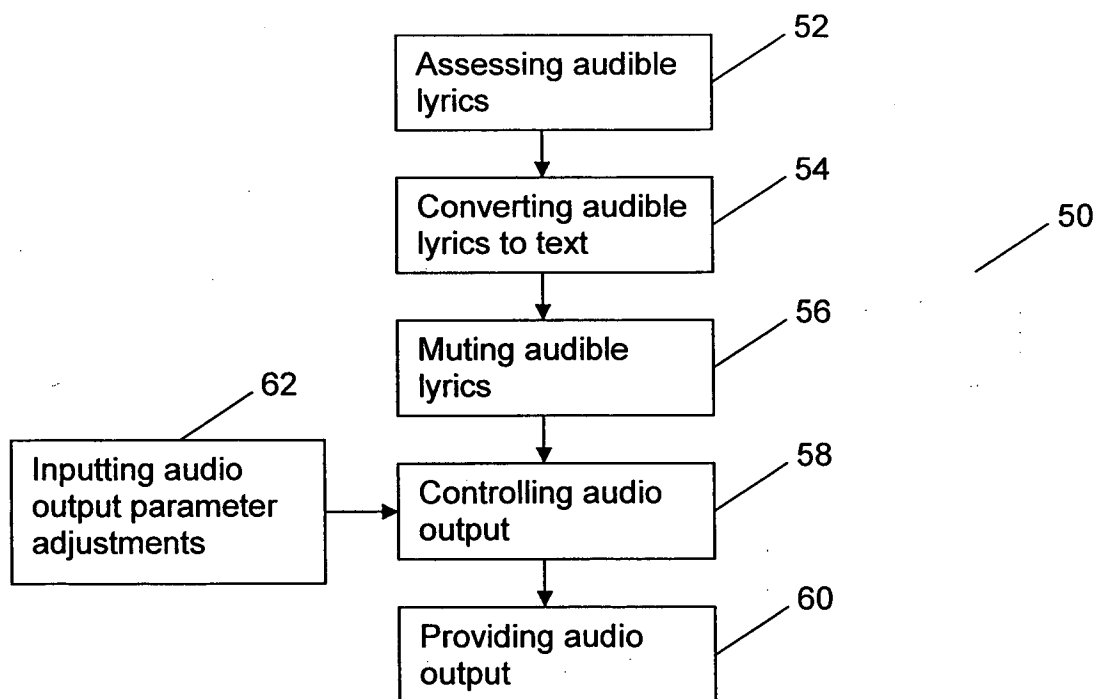


Figure 2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2011/000013

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

G10K 15/04 (2006.01)

H04N 7/14 (2006.01)

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)

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)

EPODOC, WPI, Google, Google Patents: karaoke, sing, speech, invisible, hidden, conceal, microphone array, camera, image, casing, housing, portable, convenience and like keywords

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2004/0041902 A1 (WASHINGTON) 4 March 2004 Whole document, in particular: paragraphs [0008, 0010, 0023, 0025, 0027, 0028, 0029, 0031, 0034, 0039, 0044, 0050, 0060], figures [2, 3, 4A, 4B, 8], claims [1, 15]	1-3, 8-14
Y	Whole document	4-7, 15-17
Y	US 6086380 A (CHU <i>et al.</i>) 11 July 2000 Whole document, in particular: abstract, col 2 ln 39-41, col 8 ln 2-6, col 8 ln 63 to col 9 ln 7, col 11 ln 1-7, col 13 ln 7-16, figure [3], claim [1]	4-7, 15-17
A	US 2004/0001137 A1 (CUTLER <i>et al.</i>) 1 January 2004 Whole document	
A	EP 0694833 A1 (AT&T CORP.) 31 January 1996 Whole document	



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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"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
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"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
22 February 2011

Date of mailing of the international search report

24 FEB 2011

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustalia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer

LILI ENG
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 2 6283 3138

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/SG2011/000013

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member			
US	2004041902	NONE			
US	6086380	NONE			
US	2004001137	CN	1479525	EP	1377041
		KR	20040002639	US	7852369
EP	0694833	CA	2151073	JP	8106374
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.					
END OF ANNEX					