



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
18.07.2018 Bulletin 2018/29

(51) Int Cl.:
H04R 3/00 (2006.01) **H04S 7/00** (2006.01)

(21) Application number: **17151645.3**

(22) Date of filing: **16.01.2017**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
MA MD

(71) Applicant: **Vestel Elektronik Sanayi ve Ticaret A.S.**
45030 Manisa (TR)

(72) Inventors:
• **ULUAG, Onur**
45030 Manisa (TR)
• **BAKANOGLU, Kagan**
45030 Manisa (TR)

(74) Representative: **Ascherl, Andreas et al**
KEHL, ASCHERL, LIEBHOFF & ETTMAYR
Patentanwälte - Partnerschaft
Emil-Riedel-Strasse 18
80538 München (DE)

(54) **VIDEO DISPLAY APPARATUS AND METHOD OF OPERATING THE SAME**

(57) The present invention provides a video display apparatus (100) at least comprising a display screen (10), at least one loudspeaker for emitting a sound in association with at least one still or moving image displayed on the display screen (10), at least two spatially separated microphones (1, 2, 3), and an audio signal processing unit (20) configured to separate the sound emitted by the loudspeaker from a sound received by the microphones. The present invention also provides a method of operating a video display apparatus, wherein the method at least comprises displaying on a display screen of the

apparatus at least one still or moving image, emitting a sound from a loudspeaker of the apparatus in association with displaying the at least one still or moving image, receiving a sound by at least two spatially separated microphones of the apparatus, and separating the sound emitted from the loudspeaker from the sound received by the microphones. Such a method allows for three-dimensional localization and separation of sound sources to receive and execute voice commands for control of a video display apparatus, such as a television, without the need for a remote control.

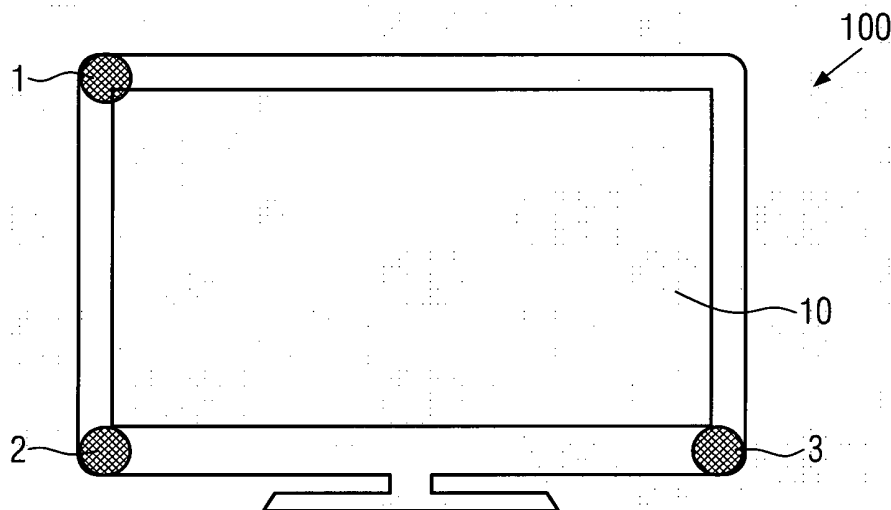


FIG. 3

Description

[0001] The present invention relates to a video display apparatus according to claim 1 and to a method of operating a video display apparatus according to claim 12.

Background of the Invention

[0002] At present, video display apparatuses, such as televisions, video games machines and computer monitors, are typically operated either by touch, using one or more push buttons, a keyboard, keypad, joystick, and/or touch screen of the display apparatus, or by transmitting electromagnetic signals to the apparatus, for example, infrared or radio waves, using a separate device, such as a dedicated remote control and/or smart phone. It is hard to operate a video display apparatus using sound, such as voice commands, because apart from comprising a display screen, such a video display apparatus typically also comprises at least one loudspeaker, which itself emits sound in association with still or moving images displayed on the display screen, for example as a sound track accompanying a film or television programme or as sound effects accompanying a video game. It is difficult for sound signals intended to operate the video display apparatus to be discriminated from these sounds emitted by the display apparatus itself, as well as from echoes, which are hard to model and predict, and from background noise.

Object of the Invention

[0003] It is therefore an object of the invention to provide a video display apparatus and a method of operating a video display apparatus.

Description of the Invention

[0004] The object of the invention is solved by a video display apparatus according to claim 1. The video display apparatus at least comprises a display screen, at least one loudspeaker for emitting a sound in association with at least one still or moving image displayed on the display screen, at least two spatially separated microphones, and an audio signal processing unit configured to separate the sound emitted by the loudspeaker from a sound received by the microphones.

[0005] This solution is beneficial since such a video display apparatus is relatively much larger than a portable device like a remote control unit or a smart phone, so that the at least two microphones can be positioned sufficiently far apart from each other to give good spatial resolution for discriminating sound sources from each other. Moreover, since the audio signal processing unit can receive the sound emitted by the loudspeaker directly as an electronic signal before, during or after its emission, the sound emitted by the loudspeaker can be separated from the sound received by the microphones with a high degree of certainty and echoes can be easily identified and accounted for.

[0006] Advantageous embodiments of the invention may be configured according to any claim and/or part of the following description.

[0007] At least one of the microphones is preferably located adjacent the display screen, facing the same direction as the display screen. This improves the chances that at least one of the microphones will be facing a viewer of the display screen. More preferably, at least two of the microphones are located on either side of the display screen with the display screen between them, facing the same direction as the display screen. This is beneficial in increasing the horizontal resolution of the microphones.

[0008] Preferably, at least one pair of the at least two microphones are spatially separated by at least 400 mm, more preferably by at least 500 mm, more preferably still by at least 600 mm, and most preferably by at least 700 mm from each other. This is advantageous because the spatial resolution of the microphones increases in proportion to their spatial separation.

[0009] In a preferred embodiment, the at least two microphones comprise three microphones arranged in a triangle. This is beneficial because it allows sound sources to be discriminated from each other in two dimensions. For example, if the triangle has one horizontal and one vertical side, this will give corresponding spatial resolution of sound sources in the horizontal and vertical directions.

[0010] Preferably, the video display apparatus further comprises a sound source locating unit. The sound source locating unit may locate the source of sounds, based upon the differences between the sound signals received by different ones of the at least two microphones. For example, the sound source locating unit may locate the source of sounds based on the different times of arrival of the sound from a single, common source at different ones of the at least two microphones.

[0011] Preferably, the video display apparatus further comprises a voice recognition unit. This is beneficial because it can allow the video display apparatus to adopt one of a plurality of different user profiles according to the voice of a

user recognized by the voice recognition unit.

[0012] If so, the video display apparatus preferably further also comprises a voice command execution unit. This is beneficial because it can allow the video display apparatus to be controlled by a user issuing voice commands, such as "switch to channel A", "increase volume", and so on, without the need for a separate control device, such as a dedicated remote control or a smart phone. It also allows the video display apparatus to be used in hands-free multimedia and gaming applications.

[0013] In one possible embodiment, the video display apparatus further comprises a multi-view display unit configured to display at least two different still or moving images on the display screen simultaneously. Multi-view is an existing display technology allowing at least two different still or moving images to be displayed on the display screen simultaneously, for example by displaying the different images with different polarizations from each other. A plurality of viewers with multi-view glasses of correspondingly different polarizations may then watch respective ones of the different still or moving images simultaneously without the need for a split screen. For example, one viewer may watch a film or television programme whilst another viewer browses an album of photos or plays a video game on the same display screen. Typically in such a case, one or more viewers may wear head or earphones supplied by the video display apparatus with a respective sound signal appropriate to the image or images being watched by the viewer in question.

[0014] If so, and if the video display apparatus also comprises a sound source locating unit and a voice command execution unit, the voice command execution unit is preferably configured to execute a command in relation to a respective one of the simultaneously displayed still or moving images and/or a sound signal generated by the video display apparatus according to a location of a sound source issuing the command identified by the sound source locating unit. This is beneficial because a plurality of viewers of the display screen may then control whatever they are watching by issuing one or more voice commands which only affect the images they are viewing and/or the sound signal they are receiving and not the different images or sound of another simultaneous viewer. It also allows for display of the different images and/or the corresponding sound signals to be adapted to the respective locations of the simultaneous viewers. For example, the respective images may track the location of a viewer as they move. According to this example, if two simultaneous viewers swap positions, one of the viewers may call out "I'm over here" as a voice command, and the video display apparatus may then redirect the displayed images and/or the accompanying sound signals accordingly.

[0015] In one possible embodiment, the video display apparatus may further comprise a television receiver. This allows the video display apparatus to display television programmes and for the programmes to be selected and controlled using voice commands, instead of using a separate device, such as a dedicated remote control or smart phone.

[0016] Preferably, the audio signal processing unit is further configured to separate environmental noise from the sound received by the microphones. This is beneficial because it can be used to improve the accuracy of sound source location, voice recognition and execution of voice commands. The separation of environmental noise from the sound received by the microphones may be carried out by sampling the sound received by the microphones at times when the loudspeaker of the video display apparatus is silent and when no rapid variations in the volume of sound received by the microphones is detected, which might otherwise be indicative of a user's voice, and then using these samples as examples of environmental noise.

[0017] The present invention further relates to a method of operating a video display apparatus. The method at least comprises displaying on a display screen of the apparatus at least one still or moving image, emitting a sound from a loudspeaker of the apparatus in association with displaying the at least one still or moving image, receiving a sound by at least two spatially separated microphones of the apparatus, and separating the sound emitted from the loudspeaker from the sound received by the microphones.

[0018] Preferably, the method further comprises locating at least one source of the sound received by the microphones.

[0019] Preferably, the method further comprises recognizing at least one voice in the sound received by the microphones.

[0020] If so, the method preferably further comprises executing a command issued by the at least one voice according to the location of the sound source issuing the command.

[0021] The present invention further relates to a computer program product or a program code or system for executing one or more than one of the herein described methods.

[0022] Further features, goals and advantages of the present invention will now be described in association with the accompanying drawings, in which exemplary components of the invention are illustrated. Components of the apparatuses and methods according to the invention which are at least essentially equivalent to each other with respect to their function can be marked by the same reference numerals, wherein such components do not have to be marked or described in all of the drawings.

[0023] In the following description, the invention is described by way of example only with respect to the accompanying drawings.

Brief Description of the Drawings

[0024]

Fig. 1 is a schematic plan view of different viewer positions relative to a display screen of a video display apparatus;

Fig. 2 is a schematic diagram of separating and processing sound signals received from a plurality of different sources by stereo microphones;

Fig. 3 is a schematic representation of an embodiment of a video display apparatus comprising a plurality of spatially separated microphones;

Fig. 4 schematically represents a three-dimensional method of calculating the distances of a plurality of spatially separated microphones from a single source of sound;

Fig. 5 is a schematic block diagram of signal processing sound signals from two different sources;

Fig. 6 is a graph representing sound signals received from a single source by two spatially separated microphones; and

Fig. 7 is a graph representing sound wave power dissipation over distance.

Detailed Description

[0025] Fig. 1 schematically shows a plan view of different positions P0, P1, P2, P3 of a viewer relative to a display screen 10 of a video display apparatus. Only when the viewer is positioned somewhere in a plane equidistant between the two horizontal extremities of the display screen 10, is the viewer in a "sweet spot", as represented in Fig. 1 by position P0. In this position, a pair of spatially separated microphones, each respectively located adjacent one of the two horizontal extremities of the display screen 10, will receive the same sound emitted by the viewer as each other. In all other positions, such as those represented by P1, P2, P3 in Fig. 1, the viewer is at a greater distance from one of the two horizontal extremities of the display screen 10 than from the other. In any one of these other positions, the sound received by one of the pair of spatially separated microphones located adjacent one of the horizontal extremities of the display screen 10 will be different from the sound received by the other such microphone so that it is possible to effectively separate the sound emitted by the loudspeaker from a surrounding sound received by the microphones.

[0026] Fig. 2 schematically represents separating and processing sound signals received from a plurality of different sound sources Source 1, Source 2, Source 3 by stereo microphones 1, 2 by means of an audio signal processing unit 20. The stereo microphones 1, 2 produce left and right channel audio signals as illustrated in Fig. 2. The audio signal processing unit 20 compares these left and right channel audio signals and extracts from them estimates Estimate 1, Estimate 2, Estimate 3, each of which respectively corresponds to one of the sounds produced by Source 1, Source 2, Source 3.

[0027] Fig. 3 schematically represents an embodiment of a video display apparatus 100. The video display apparatus 100 comprises a display screen 10 and a plurality of spatially separated microphones 1, 2, 3. The microphones 1, 2, 3 are located adjacent the display screen and are arranged in a triangle. The pair of microphones 1, 2 are spatially separated from each other by more than 400 mm, the pair of microphones 2, 3 are spatially separated from each other by more than 500 mm, and the pair of microphones 1, 3 are spatially separated from each other by more than 600 mm.

[0028] The video display apparatus 100 further comprises several loudspeakers (not visible in Fig. 3) for emitting a sound in association with the display of at least one still or moving image on the display screen 10. The video display apparatus 100 also contains a television receiver and an audio signal processing unit, neither of which are visible in Fig. 3. The audio signal processing unit is configured to separate the sound emitted by the loudspeakers from a sound received by the microphones 1, 2, 3.

[0029] Fig. 4 schematically represents a method of calculating the distances in three-dimensions of a plurality of spatially separated microphones, Mic 0, Mic 1, Mic 2, Mic 3 from a single source of sound, S. In this example, the sound source, S, is located at co-ordinates x, y, z in an arbitrarily defined three-dimensional Cartesian co-ordinate system and emits a sound at time, t. As may be seen from Fig. 4, the plurality of spatially separated microphones, Mic 0, Mic 1, Mic 2, Mic 3 comprises four different combinations of three microphones arranged in a triangle. Mic 0 is located at co-ordinates x0, y0, z0 and receives the sound from source S at time t0. Mic 1 is located at co-ordinates x1, y1, z1 and receives the sound from source S at time t1. Similarly, Mic 2 is located at co-ordinates x2, y2, z2 and receives the sound from source S at time t2. Finally, Mic 3 is located at co-ordinates x3, y3, z3 and receives the sound from source S at

time t_3 . The distance = $(x_0 - x, y_0 - y, z_0 - z)$ of Mic 0 from the sound source S is therefore given by the speed of sound, c , multiplied by the difference between the time, t_0 , of reception of the sound by Mic 0 and the time, t , of its emission: $c \cdot (t_0 - t)$. Similarly, the distance of Mic 1 from the sound source S is given by $c \cdot (t_1 - t)$, the distance of Mic 2 from the sound source S is given by $c \cdot (t_2 - t)$, and the distance of Mic 3 from the sound source S is given by $c \cdot (t_3 - t)$. Thus by comparing the different times of reception of the sound at the different microphones Mic 0, Mic 1, Mic 2, Mic 3, the location x, y, z of the sound source in the co-ordinate system may be calculated. Such a method as that described in relation to Fig. 4 may be carried out by a sound source locating unit of a video display apparatus according to an embodiment of the invention.

[0030] Fig. 5 schematically represents how two different sound signals respectively received from two different sources of sound may be modelled. In the example shown in Fig. 5, a first sound signal sine1 having a frequency of 10 Hz is emitted from a first source of sound and a second sound signal sine2 having a frequency of 20 Hz is emitted from a second source of sound. Purely for the sake of this example, sine1 and sine2 are both represented as having a sinusoidal waveform, although in practice, they may have any waveform and any other audio frequency or range of frequencies. sine1 and sine2 are both received by each one of a pair of spatially separated microphones. The first microphone may be modelled by two amplifiers gain1 , gain2 and by an adder labelled add1 in Fig. 5. The second microphone may be modelled by two further amplifiers gain3 , gain4 and by a second adder labelled add2 . Since the first sound source is nearer to the first microphone than it is to the second, the sound signal sine1 may be modelled as passing through the amplifier gain1 with a gain, $k = 0.9$ and through the amplifier gain3 with a gain of only $k = 0.3$. On the other hand, since the second sound source is nearer to the second microphone than it is to the first, the sound signal sine2 may instead be modelled as passing through the amplifier gain2 with a gain of only $k = 0.3$ and through the amplifier gain4 with a gain, $k = 0.9$. Subsequent to these respective amplifications, the sound signals sine1 , sine2 are added to each other by the adders add1 , add2 of each microphone as shown in Fig. 5.

[0031] Fig. 6 schematically represents two sound signals 61, 62, each respectively received by one of two spatially separated microphones from a single, common source. The graph of Fig. 6 plots the amplitude, A , of the two sound signals 61, 62 on the y-axis or ordinate against time, t , on the x-axis or abscissa. As may be seen from Fig. 6, whereas the two sound signals 61, 62 have the same frequency as each other and a similar waveform to each other (which, for the sake of this example, is a sinusoid) the amplitude A of the sound signal 61 differs from that of the sound signal 62, since the common source of the two signals 61, 62 is located further from one of the two microphones than the other.

[0032] Fig. 7 schematically represents sound wave power dissipation over distance. The graph of Fig. 7 plots the amplitude, A , of a sound wave 71 on the y-axis or ordinate against distance, x , on the x-axis or abscissa from the emission of the sound wave 71 by a source, S, to its reception, R. As may be seen from Fig. 7, the amplitude, A , of the sound wave 71 progressively diminishes between the source, S, and its reception, R. The power of the sound wave 71, which is proportional to the square of the amplitude, A , therefore also dissipates accordingly.

[0033] In summary, therefore, the present invention provides a video display apparatus at least comprising a display screen, at least one loudspeaker for emitting a sound in association with at least one still or moving image displayed on the display screen, at least two spatially separated microphones, and an audio signal processing unit configured to separate the sound emitted by the loudspeaker from a sound received by the microphones. The present invention also provides a method of operating a video display apparatus, wherein the method at least comprises displaying on a display screen of the apparatus at least one still or moving image, emitting a sound from a loudspeaker of the apparatus in association with displaying the at least one still or moving image, receiving a sound by at least two spatially separated microphones of the apparatus, and separating the sound emitted from the loudspeaker from the sound received by the microphones. Such a method allows for three-dimensional localization and separation of sound sources to receive and execute voice commands for control of a video display apparatus, such as a television, without the need for a remote control.

Reference Numerals:

1,2,3	Spatially separated microphones	$\text{gain1}, \text{gain2}, \text{gain3}, \text{gain4}$	Amplifiers
10	Display screen		
20	Audio signal processing unit	Mic 0, Mic 1, Mic 2,	Plurality of spatially separated microphones
30	Model of two sound signals	Mic 3	
61	First audio signal		
62	Second audio signal	P0, P1, P2, P3	Different positions of viewer
71	Sound wave		
100	Video display apparatus	R	Reception
A	Amplitude	S	Sound source

(continued)

	add1, add2	Adders	Source 1, Source 2, Source 3	Plurality of sound sources
5	Estimate 1, Estimate 2, Estimate 3	Estimates of sounds produced by sound sources	sine1, sine2 t x	Different sound signals Time Distance

10 Claims

1. A video display apparatus (100) at least comprising:

a display screen (10);

at least one loudspeaker for emitting a sound in association with at least one still or moving image displayed on the display screen (10);

at least two spatially separated microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3); and

an audio signal processing unit (20) configured to separate the sound emitted by the loudspeaker from a sound received by the microphones.

2. A video display apparatus according to claim 1, wherein at least one of the microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3) is located adjacent the display screen (10), facing the same direction as the display screen.

3. A video display apparatus according to claim 1 or claim 2, wherein at least one pair (1, 2; 2, 3; 1, 3) of the at least two microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3) are spatially separated from each other by at least 400 mm.

4. A video display apparatus according to any one of the preceding claims, wherein the at least two microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3) comprise three microphones arranged in a triangle.

5. A video display apparatus according to any one of the preceding claims, further comprising a sound source locating unit.

6. A video display apparatus according to any one of the preceding claims, further comprising a voice recognition unit.

7. A video display apparatus according to claim 6, further comprising a voice command execution unit.

8. A video display apparatus according to any one of the preceding claims, further comprising a multi-view display unit configured to display at least two different still or moving images on the display screen (10) simultaneously.

9. A video display apparatus according to claim 8 as dependent on claim 7 and claim 5, wherein the voice command execution unit is configured to execute a command in relation to at least one of a respective one of the simultaneously displayed still or moving images and a sound signal generated by the video display apparatus according to a location of a sound source (S) issuing the command identified by the sound source locating unit.

10. A video display apparatus according to any one of the preceding claims, further comprising a television receiver.

11. A video display apparatus according to any one of the preceding claims, wherein the audio signal processing unit (20) is further configured to separate environmental noise from the sound received by the microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3).

12. A method of operating a video display apparatus (100), the method at least comprising:

displaying on a display screen (10) of the apparatus at least one still or moving image;

emitting a sound from a loudspeaker of the apparatus in association with displaying the at least one still or moving image;

receiving a sound by at least two spatially separated microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3) of the apparatus; and

separating the sound emitted from the loudspeaker from the sound received by the microphones.

EP 3 349 480 A1

13. A method according to claim 12, further comprising locating at least one source (S) of the sound received by the microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3).
- 5 14. A method according to claim 12 or claim 13, further comprising recognizing at least one voice in the sound received by the microphones (1, 2, 3; Mic 0, Mic 1, Mic 2, Mic 3).
15. A method according to claim 14 as dependent on claim 13, further comprising executing a command issued by the at least one voice according to the location of the sound source (S) issuing the command.

10

15

20

25

30

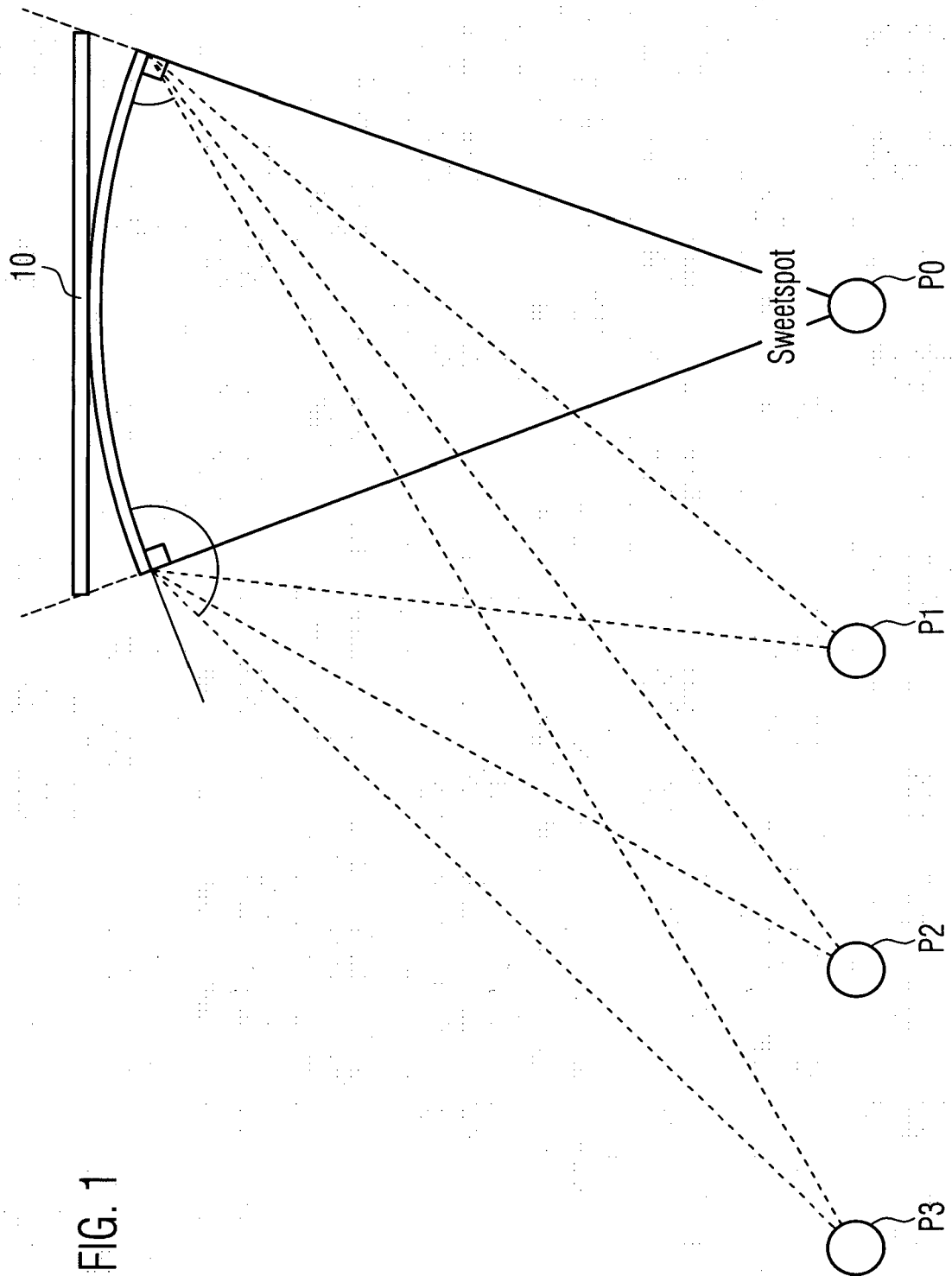
35

40

45

50

55



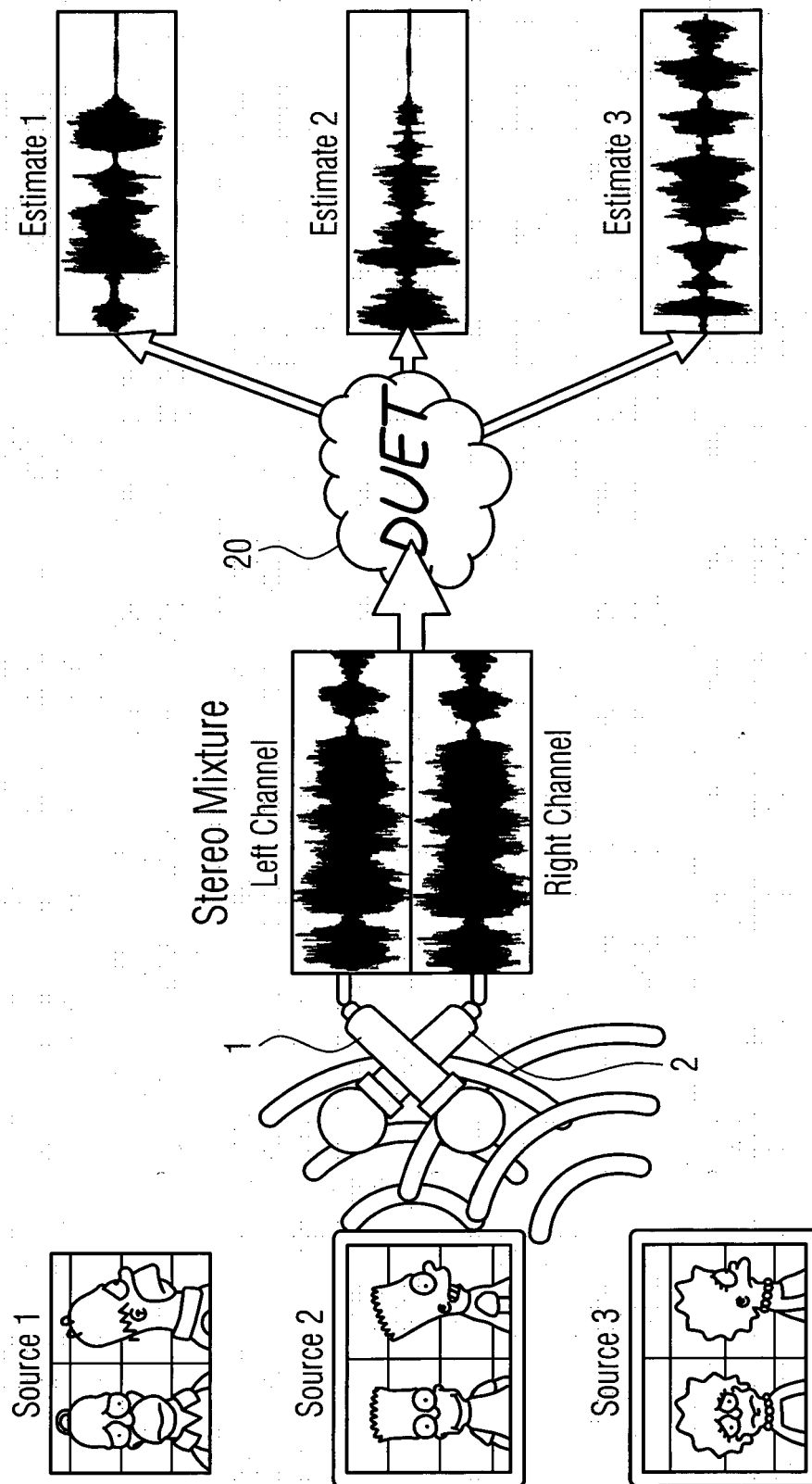


FIG. 2

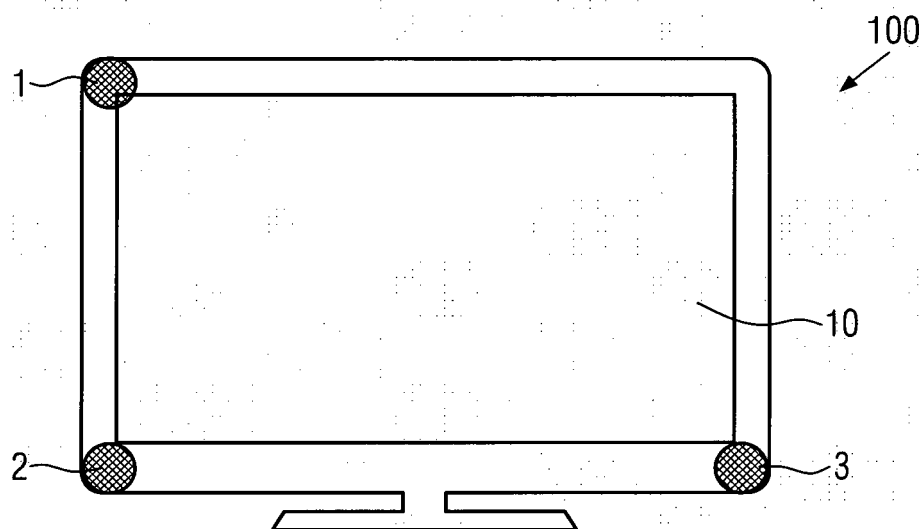


FIG. 3

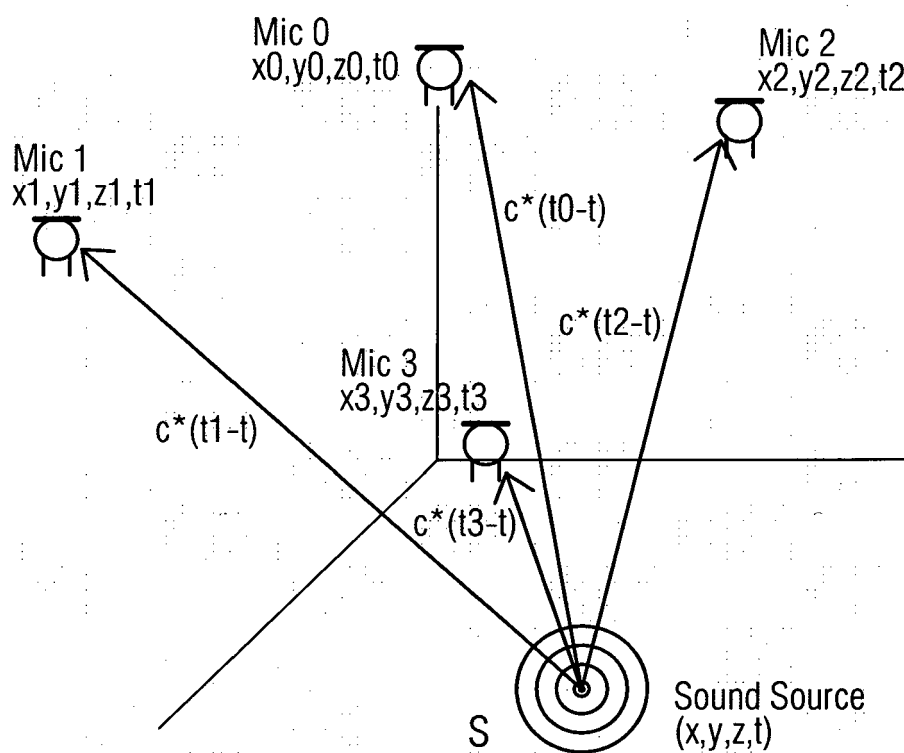


FIG. 4

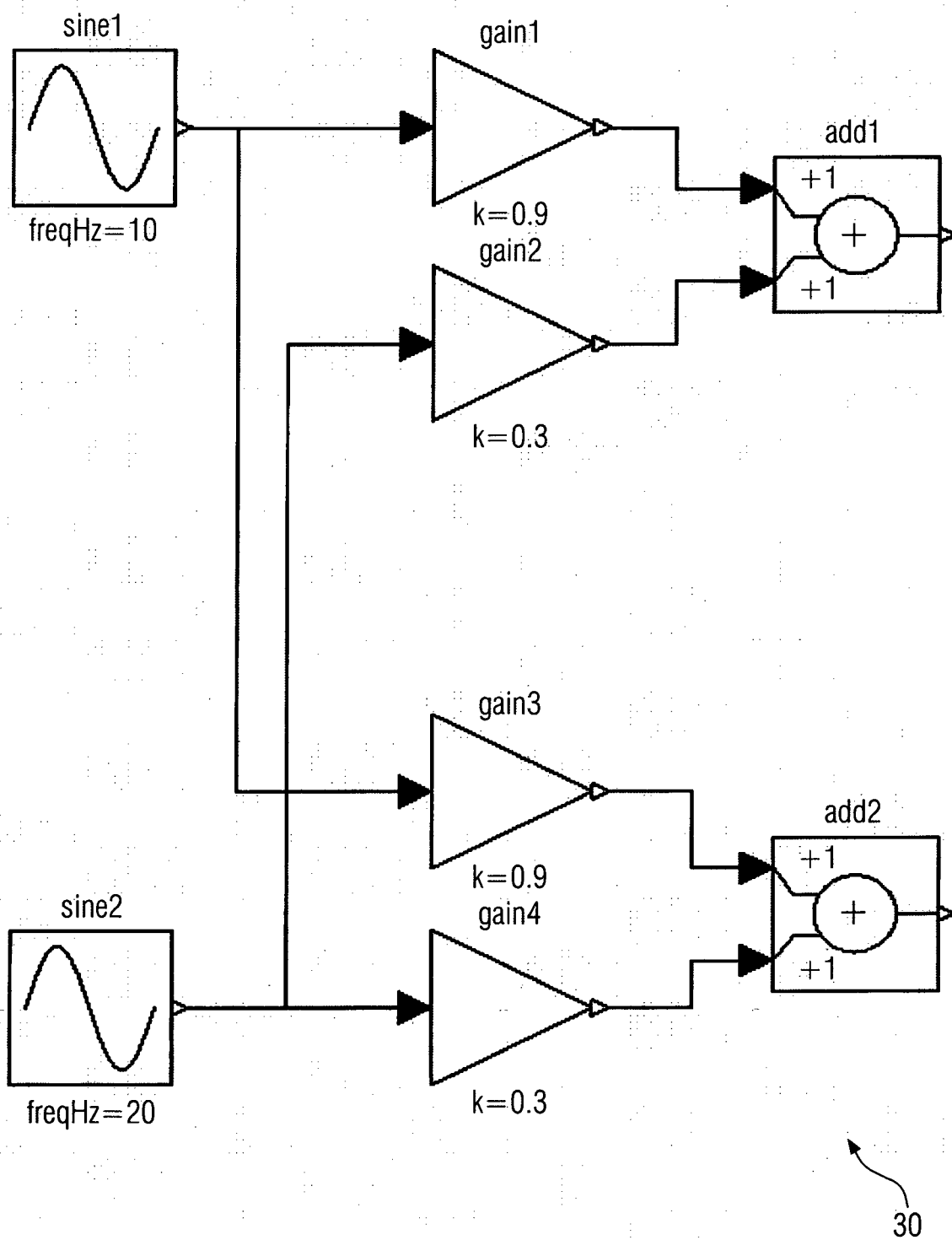


FIG. 5

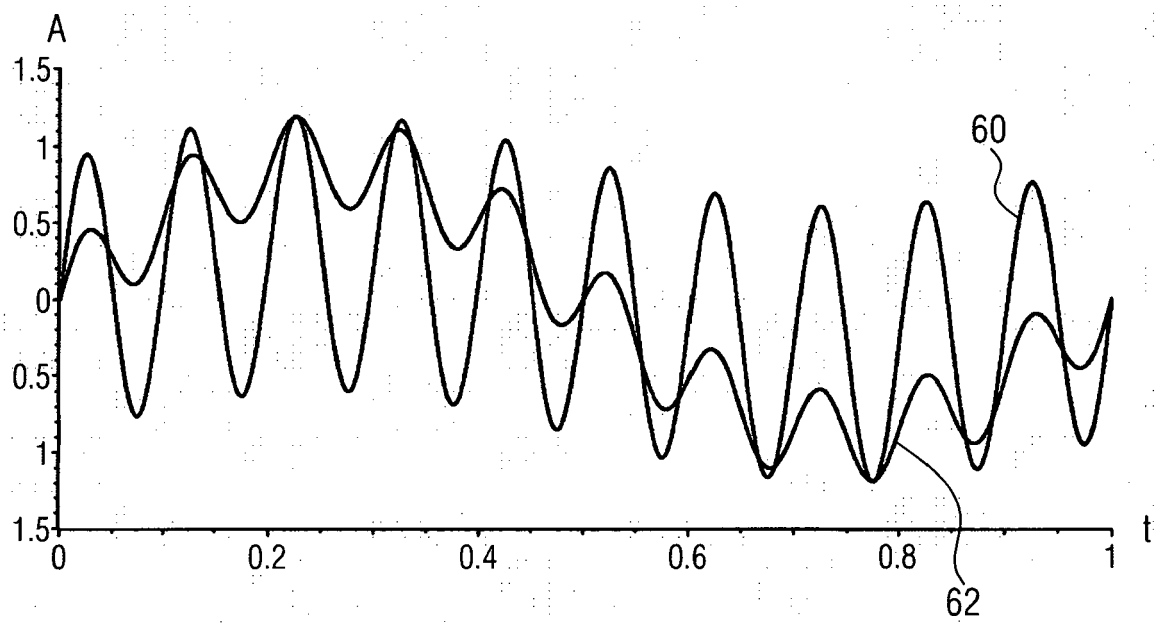


FIG. 6

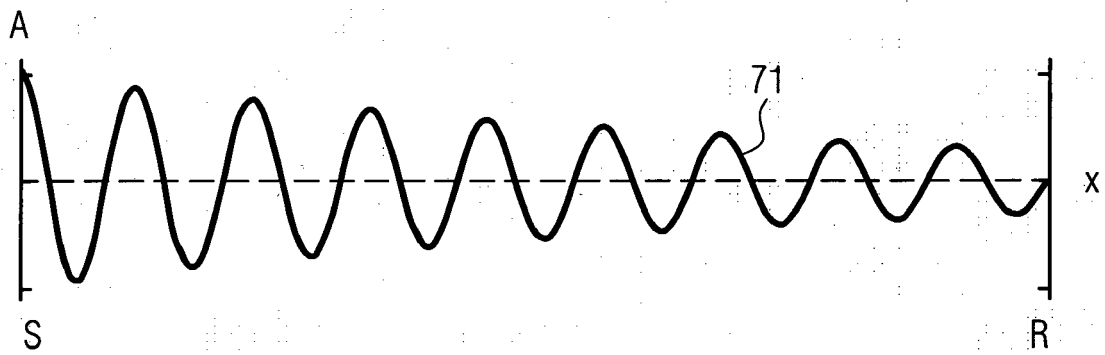


FIG. 7



EUROPEAN SEARCH REPORT

 Application Number
 EP 17 15 1645

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	EP 2 923 634 A1 (STORZ ENDOSKOP PROD GMBH [DE]) 30 September 2015 (2015-09-30) * paragraphs [0014], [0017], [0019]; figure 1 *	1-9, 11-15	INV. H04R3/00 ADD. H04S7/00
X	US 2016/212525 A1 (NAKADAI KAZUHIRO [JP]) 21 July 2016 (2016-07-21) * paragraphs [0141] - [0142], [0148], [0151], [0163] - [0164]; figures 12-13 *	1,3-8, 10-15	
A	US 2016/259305 A1 (CHEN SHUO [CN] ET AL) 8 September 2016 (2016-09-08) * figure 3 *	3,4	
A	EP 1 835 478 A1 (FUJITSU TEN LTD [JP]) 19 September 2007 (2007-09-19) * figures 1-2 *	1-15	
			TECHNICAL FIELDS SEARCHED (IPC)
			H04R H04S G10L H04N
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 23 June 2017	Examiner Rogala, Tomasz
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1
 EPO FORM 1503 03.02 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 17 15 1645

5 This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

23-06-2017

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 2923634 A1	30-09-2015	EP 2923634 A1	30-09-2015
		US 2015279368 A1	01-10-2015
US 2016212525 A1	21-07-2016	JP 2016133304 A	25-07-2016
		US 2016212525 A1	21-07-2016
US 2016259305 A1	08-09-2016	CN 104240606 A	24-12-2014
		US 2016259305 A1	08-09-2016
		WO 2016026204 A1	25-02-2016
EP 1835478 A1	19-09-2007	CN 101080755 A	28-11-2007
		EP 1835478 A1	19-09-2007
		JP 2007041489 A	15-02-2007
		KR 20070092732 A	13-09-2007
		US 2008106595 A1	08-05-2008
		WO 2006064619 A1	22-06-2006