

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(10) **DK/EP 3566468 T3**

(12) Oversættelse af
europæisk patentskrift

-
- (51) Int.Cl.: **H 04 R 25/00 (2006.01)** **G 10 L 21/0216 (2013.01)** **G 10 L 25/60 (2013.01)**
H 04 R 3/00 (2006.01) **H 04 R 27/00 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2021-05-10**
- (80) Dato for Den Europæiske Patentmyndigheds
bekendtgørelse om meddelelse af patentet: **2021-03-10**
- (86) Europæisk ansøgning nr.: **17700268.0**
- (86) Europæisk indleveringsdag: **2017-01-09**
- (87) Den europæiske ansøgnings publiceringsdag: **2019-11-13**
- (86) International ansøgning nr.: **EP2017050341**
- (87) Internationalt publikationsnr.: **WO2018127298**
- (84) Designerede stater: **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**
- (73) Patenthaver: **Sonova AG, Laubisrütistrasse 28, 8712 Stäfa, Schweiz**
- (72) Opfinder: **JOST, Timothée, Route de Peseux 3, 2012 Auvernier, Schweiz**
GIGANDET, Xavier, Impasse Côte-des-Esserts 15, 1774 Cousset, Schweiz
- (54) Benævnelse: **MIKROFONAGGREGAT TIL AT BÆRE PÅ BRUGERENS BRYST**
- (56) Fremdragne publikationer:
US-A1- 2012 239 385
US-A1- 2014 093 091
US-A1- 2016 255 444

DESCRIPTION

[0001] The invention relates to microphone assembly to be worn at a user's chest for capturing the user's voice.

[0002] Typically, such microphone assemblies are worn at the user's chest either by using a clip for attachment to the user's clothing or by using a lanyard, so as to generate an output audio signal corresponding to the user's voice, with the microphone assembly usually including a beamformer unit for processing the captured audio signals in a manner so as to create an acoustic beam directed towards the user's mouth. Such microphone assembly typically forms part of a wireless acoustic system; for example, the output audio signal of the microphone assembly may be transmitted to a hearing aid. Typically, such wireless microphone assemblies are used by teachers of hearing impaired pupils / students wearing hearing aids for receiving the speech signal captured by the microphone assembly from the teacher's voice.

[0003] By using such chest-worn microphone assembly, the user's voice can be picked up close to the user's mouth (typically at a distance of about 20 cm), thus minimizing degradation of the speech signal in the acoustic environment.

[0004] However, while the use of a beamformer may enhance the signal-to-noise ratio (SNR) of the captured voice audio signal, this requires that the microphone assembly is placed in such a way that the acoustic microphone axis is oriented towards the user's mouth, while any other orientation of the microphone assembly may result in a degradation of the speech signal to be transmitted to the hearing aid. Consequently, the user of the microphone assembly has to be instructed so as to place the microphone assembly at the proper location and with the proper orientation. However, in case that the user does not follow the instructions, only a less than optimal sound quality will be achieved. Examples of proper and improper use of a microphone assembly are illustrated in Fig. 1a.

[0005] US 2016/0255444 A1 relates to a remote wireless microphone for a hearing aid, comprising a plurality of omnidirectional microphones, a beamformer for generating an acoustic beam directed towards the mouth of the user and an accelerometer for determining the orientation of the microphone assembly relative to the direction of gravity, wherein the beamformer is controlled in such a manner that the beam always points into an upward direction, i.e. in a direction opposite to the direction of gravity.

[0006] US 2014/0270248 A1 relates to a mobile electronic device, such as a headset or a smartphone, comprising a directional microphone array and a sensor for determining the orientation of the electronic device relative to the orientation of the user's head so as to control the direction of an acoustic beam of the microphone array according to the detected orientation relative to the user's head.

[0007] US 9,066,169 B2 relates to a wireless microphone assembly comprising three

microphones and a position sensor, wherein one or two of the microphones are selected according to the position and orientation of the microphone assembly for providing the input audio signal, wherein a likely position of the user's mouth may be taken into account.

[0008] US 9,066,170 B2 relates to a portable electronic device, such as a smartphone, comprising a plurality of microphones, a beamformer and orientation sensors, wherein a direction of a sound source is determined and the beamformer is controlled, based on the signal provided by the orientation sensors, in such a manner that the beam may follow movements of the sound source.

[0009] It is an object of the invention to provide for a microphone assembly to be worn at a user's chest which is capable of providing for an acceptable SNR in a reliable manner. It is a further object to provide for a corresponding method for generating an output audio signal from a user's voice.

[0010] According to the invention, these objects are achieved by a microphone assembly as defined in claims 1 and 15, respectively.

[0011] The invention is beneficial in that, by selecting one acoustic beam from a plurality of fixed acoustic beams (i.e. beams which are stationary with regard to the microphone assembly) by taking into account both the orientation of the selected beam with regard to the direction of gravity (or, more precisely, the direction of the projection of the direction of gravity onto the microphone plane) and an estimated speech quality of the selected beam, an output signal of the microphone assembly having a relatively high SNR can be obtained, irrespective of the actual orientation and position on the user's chest relative to the user's mouth.

[0012] Having fixed beams allows to have a stable and reliable beamforming stage, while at the same time allowing for fast switching from one beam to another, thereby enabling fast adaptations to changes in the acoustic conditions. In particular, compared to systems using an adjustable beam, i.e. rotating beam with adjustable angular target, the present selection from fixed beams is less complex and is less prone to be perturbed by interferers

[0013] (environmental noise, neighbouring talker,...); also, adaptive part of such adjustable beam is also critical: If too slow, the system will take time to converge to the optimal solution and part of the talker's speech may be lost; if too fast, then the beam may target interferers during speech breaks.

[0014] More in detail, by taking into account both the orientation of the selected beam with regard to gravity and the estimated speech quality of the selected beam, not only a tilt of the microphone assembly with regard to the vertical axis but also a lateral offset with regard to the center of the user's chest may be compensated for. For example, when the microphone assembly is laterally offset, the most vertical beam may not always be the optimal choice, since the user's mouth in such case could be located 30° or more off the vertical axis, so that in the most vertical beam the desired voice signal would be already attenuated, while, when taking

into account also the estimated speech quality, a beam close to the most vertical beam may be selected which in such case would provide for a higher SNR than the most vertical beam. Thus, the invention allows for orientation-independent and also partially location-independent positioning of the microphone assembly on the user's chest.

[0015] Preferred embodiments are defined in the dependent claims.

[0016] Hereinafter, examples of the invention will be illustrated by reference to the attached drawings, wherein:

Fig. 1a

is a schematic illustration of the orientation of an acoustic beam of a microphone assembly of the prior art with a fixed beam former relative to the user's mouth;

Fig. 1b

is a schematic illustration of the orientation of the acoustic beam of a microphone assembly according to the invention relative to the user's mouth;

Fig. 2

is a schematic illustration of an example of a microphone assembly according to the invention, comprising three microphones arranged as a triangle;

Fig. 3

is an example of a block diagram of a microphone assembly according to the invention;

Fig. 4

is an illustration of the acoustic beams produced by the beamformer of the microphone assembly of Figs. 2 and 3;

Fig. 5

is an example of a directivity pattern which can be obtained by the beamformer of the microphone assembly of Figs. 2 and 3;

Fig. 6

is a representation of the directivity index (upper part) and of the white noise gain (lower part) of the directivity pattern of Fig. 5 as a function of frequency;

Fig. 7

is a schematic illustration of the selection of one of the beams of Fig. 4 in a practical use case;

Fig. 8

is an example of a use of a wireless hearing system using a microphone assembly according to the invention; and

Fig. 9

is a block diagram of a speech enhancement system using a microphone assembly according to the invention.

[0017] Fig. 2 is a schematic perspective view of an example of a microphone assembly 10 comprising a housing 12 having essentially the shape of a rectangular prism with a first essentially rectangular flat surface 14 and a second essentially rectangular flat surface (not

shown in Fig. 2) which is parallel to the first surface 14. Rather than having a rectangular shape, the housing may have any suitable form factor, such as round shape. The microphone assembly 10 further comprises three microphones 20, 21, 22, which preferably are arranged such that the microphones (or the respective microphone openings in the surface 14) form an equilateral triangle or at least an approximation of a triangle (for example, the triangle may be approximated by a configuration wherein the microphones 20, 21, 22 are distributed approximately uniformly on a circle, wherein each angle between adjacent microphones is from 110 to 130°, with the sum of the three angles being 360°).

[0018] According to one example, the microphone assembly 10 may further comprise a clip on mechanism (not shown in Fig. 2) for attaching the microphone assembly 10 to the clothing of a user at a position at the user's chest close to the user's mouth; alternatively, the microphone assembly 10 may be configured to be carried by a lanyard (not shown in Fig. 2). The microphone assembly 10 is designed to be worn in such a manner that the flat rectangular surface 14 is essentially parallel to the vertical direction.

[0019] In general, there may be more than three microphones. In an arrangement of four microphones, the microphones still may be distributed on a circle, preferably uniformly. For more than four microphones the arrangement may be more complex, e.g. five microphones may be ideally arranged as the figure five on a dice. More than five microphones preferably would be placed on a matrix configuration, e.g. a 2x3 matrix, 3x3 matrix, etc.

[0020] In the example of Fig. 2 the longitudinal axis of the housing 12 is labelled "x", the transverse direction is labelled "y" and the elevation direction is labelled "z" (the z-axis is normal to the plane defined by the x-axis and the y-axis). Ideally, the microphone assembly 10 would be worn in such a manner that the x-axis corresponds to the vertical direction (direction of gravity) and the flat surface 14 (which essentially corresponds to the x-y-plane) is parallel to the user's chest.

[0021] As illustrated by the block diagram shown in Fig. 3, the microphone assembly further comprises an acceleration sensor 30, a beamformer unit 32, a beam selection unit 34, an audio signal processing unit 36, a speech quality estimation unit 38 and an output selection unit 40.

[0022] The audio signals captured by the microphones 20, 21, 22 are supplied to the beamformer unit 32 which processes the captured audio signals in a manner so as to create 12 acoustic beams 1a-6a, 1b-6b having directions uniformly spread across the plane of the microphones 20, 21, 22 (i.e. the x-y-plane), with the microphones 20, 21, 22 defining a triangle 24 in Fig. 4 (in Figs. 4 and 7 the beams are represented / illustrated by their directions 1a-6a, 1b-6b).

[0023] Preferably, the microphones 20, 21, 22 are omnidirectional microphones.

[0024] The six beams 1b-6b are produced by delay-and-sum beam forming of the audio

signals of pairs of the microphones, with these beams being oriented parallel to one of the sides of the triangle 24, wherein these beams are pairwise oriented antiparallel to each other. For example, the beams 1b and 4b are antiparallel to each other and are formed by delay-and-sum beam forming of the two microphones 20 and 22, by applying an appropriate phase difference. Such beamforming process may be written in the frequency domain as:

$$Y(k) = \frac{1}{2} \left(M_x(k) - M_y(k) \cdot e^{-j \frac{2\pi k F_s p}{Nc}} \right) \quad (1)$$

wherein $M_x(k)$ and $M_y(k)$ are the spectra of the first and second microphone in bin k , respectively, F_s is the sampling frequency, N is the size of the FFT, p is the distance between the microphones and c is the speed of sound.

[0025] Further, the six beams 1a to 6a are generated by beam forming by a weighted combination of the signals of all three microphones 20, 21, 22, with these beams being parallel to one of the medians of the triangle 24, wherein these beams are pairwise oriented antiparallel to each other. This type of beam forming may be written in the frequency domain as:

$$Y(k) = \frac{1}{2} \left(M_x(k) - \frac{1}{2} (M_y(k) + M_z(k)) \cdot e^{-j \frac{2\pi k F_s p_2}{Nc}} \right) \quad (2)$$

wherein p_2 is the length of the median of the triangle,

$$p_2 = \frac{\sqrt{3}}{2} p.$$

[0026] It can be seen from Figs. 5 and 6 that the directivity pattern (Fig. 5), the directivity index versus frequency (upper part of Fig. 6) and the white noise gain as a function of frequency (lower part of Fig. 6) are very similar for these two types of beamforming (which are indicated by " $\theta_{ar} = 0$ " and " $\theta_{ar} = 30$ " in Figs. 5 and 6), with the beams 1a-6a produced by a weighted combination of the signals of all three microphones providing for a slightly more pronounced directivity at higher frequencies. In practice, however, such difference is inaudible, so that the two types of beam forming can be considered as equivalent.

[0027] Rather than using 12 beams generated from three microphones, alternative configurations may be implemented. For example, a different number of beams may be generated from the three microphones, for example only the six beams 1a-6a of the weighted combination beamforming or only the six beams 1b-6b of the delay-and-sum beam forming. Further, more than three microphones may be used. Preferably, in any configuration, the beams are uniformly spread across the microphone plane, i.e. the angle between adjacent beams is the same for all beams.

[0028] The acceleration sensor 30 preferably is a three-axes accelerometer, which allows to determine the acceleration of the microphone assembly 10 along three orthogonal axes x , y and z . Under stable conditions, i.e. when the microphone assembly 10 is stationary, gravity will be the only contribution to the acceleration, so that the orientation of the microphone assembly 10 in space, i.e. relative to the physical direction of gravity G , can be determined by combining the amount of acceleration measured along each axis, as illustrated in Fig. 2. The orientation

of the microphone assembly 10 can be described by the orientation angle θ which is given by $\text{atan}(G_y/G_x)$, wherein G_y and G_x are the measured projections of the physical gravity vector G along the x-axis and the y-axis. While in general an additional angle ϕ between the gravity vector and the z-axis would have to be combined with the angle θ so as to fully define the orientation of the microphone assembly 10 with regard to the physical gravity vector G , this angle ϕ is not relevant in the present use case, since the microphone array formed by the microphones 20, 21 and 22 is planar. Thus, the determined direction of gravity used by the microphone assembly is actually the projection of the physical gravity vector onto the microphone plane defined by the microphones 20, 21, 22.

[0029] The output signal of the accelerometer sensor 30 is supplied as input to the beam selection unit 34 which is provided for selecting a subgroup of M acoustic beams from the N acoustic beams generated by the beamformer 32 according to the information provided by the accelerometer sensor 30 in such a manner that the selected M acoustic beams are those whose direction is closest to the direction antiparallel, i.e. opposite, to the direction of gravity as determined by the accelerometer sensor 30. Preferably, the beam selection unit 34 (which actually acts as a beam subgroup selection unit) is configured to select those two acoustic beams whose direction is adjacent to the direction antiparallel to the determined direction of gravity. An example of such a selection is illustrated in Fig. 7, wherein the vertical axis 26, i.e. the projection G_{xy} of the gravity vector G onto the x-y-plane, falls in-between the beams 1a and 6b.

[0030] Preferably, the beam selection unit 34 is configured to average the signal of the accelerometer sensor 30 in time so as to enhance the reliability of the measurement and thus, the beam selection. Preferably, the time constant of such signal averaging may be from 100 ms to 500 ms.

[0031] In the example illustrated in Fig. 7, the microphone assembly 10 is inclined by 10° clockwise with regard to the vertical positions, so that the beams 1a and 6b would be selected as the two most upward beams. The selection, for example, may be made based on a look-up table with the orientation angle θ as the input, returning the indices of the selected beams as the output. Alternatively, the beam selection unit 34 may compute the scalar product between the vector $-G_{xy}$ (i.e. the projection of the gravity vector G into the x-y-plane) and a set of unitary vectors aligned with the direction of each of the twelve beams 1a-6a and 1b-6b, with the two highest scalar products indicating the two most vertical beams:

$$idx_a = \max_i(-G_x B_{a,y,i} - G_y B_{a,x,i}) \quad (3)$$

$$idx_b = \max_i(-G_x B_{b,y,i} - G_y B_{b,x,i}) \quad (4)$$

wherein idx_a and idx_b are the indices of the respective selected beam, G_x and G_y are the estimated projections of the gravity vector and $B_{a,x,i}$, $B_{a,y,i}$, $B_{b,x,i}$ and $B_{b,y,i}$ are the x and y projections of the vector corresponding to the i-th beam of type a or b, respectively.

[0032] It is to be noted that such beam selection process according to the signal provided by

the accelerometer sensor 30 only works under the assumption that the microphone assembly 10 is stationary, since any acceleration induced by movement of the microphone assembly 10 would bias the estimate of the gravity vector and thus lead to a potentially erroneous selection of beams. In order to prevent such errors, a safeguard mechanism may be implemented by using a motion detection algorithm based on the accelerometer data, with the beam selection being locked or suspended as long as the output of the motion detection algorithm exceeds a predefined threshold.

[0033] As illustrated in Fig. 3, the audio signals corresponding to the beams selected by the beam selection unit 34 are supplied as input to the audio signal processing unit 36 which has M independent channels 36A, 36B, ..., one for each of the M beams selected by the beam selection unit 34 (in the example of Fig. 3, there are two independent channels 36A, 36B in the audio signal processing unit 36), with the output audio signal produced by the respective channel for each of the M selected beams being supplied to the output unit 40 which acts as a signal mixer for selecting and outputting the processed audio signal of that one of the channels of the audio signal processing unit 36 which has the highest estimated speech quality as the output signal 42 of the microphone assembly 10. To this end, the output unit 40 is provided with the respective estimated speech quality by the speech quality estimation unit 38 which serves to estimate the speech quality of the audio signal in each of the channels 36A, 36B of the audio signal processing unit 36.

[0034] The audio signal processing unit 36 may be configured to apply adaptive beam forming in each channel, for example by combining opposite cardioids along the direction of the respective acoustic beam, or to apply a Griffith-Jim beamformer algorithm in each channel to further optimize the directivity pattern and better reject the interfering sound sources. Further, the audio signal processing unit 36 may be configured to apply noise cancellation and/or a gain model to each channel.

[0035] According to a preferred embodiment, the speech quality estimation unit 38 uses a SNR estimation for estimating the speech quality in each channel. To this end, the unit 38 may compute the instantaneous broadband energy in each channel in the logarithmic domain. A first time average of the instantaneous broadband energy is computed using time constants which ensure that the first time average is representative of speech content in the channel, with the release time being longer than the attack time at least by a factor of 2 (for example, a short attack time of 12 ms and a longer release time of 50 ms, respectively, may be used).

[0036] A second time average of the instantaneous broadband energy is computed using time constants ensuring that the second time average is representative of noise content in the channel, with the attack time being significantly longer than the release time, such as at least by a factor of 10 (for example, the attack time may be relatively long, such as 1 s, so that it is not too sensitive to speech onsets, whereas the release time is set quite short, such as 50 ms). The difference between the first time average and the second time average of the instantaneous broadband energy provides for a robust estimate of the SNR.

[0037] Alternatively, other speech quality measures than the SNR may be used, such as a speech intelligibility score.

[0038] The output unit 40 preferably averages the estimated speech quality information when selecting the channel having the highest estimated speech quality. For example, such averaging may employ signal averaging time constants of from 1 s to 10 s.

[0039] Preferably, the output unit 40 assesses a weight of 100% to that channel which has the highest estimated speech quality, apart from switching periods during which the output signal changes from a previously selected channel to a newly selected channel. In other words, during times with substantially stable conditions the output signal 42 provided by the output unit 40 consists only of one channel (corresponding to one of the beams 1a-6a, 1b-6b), which has the highest estimated speech quality. During non-stationary conditions, when beam switching may occur, such beam/channel switching by the output unit 40 preferably does not occur instantaneously; rather, the weights of the channels are made to vary in time such that the previously selected channel is faded out and the newly selected channel is faded in, wherein the newly selected channel preferably is faded in more rapidly than the previously selected channel is faded out, so as to provide for a smooth and pleasant hearing impression. It is to be noted that usually such beam switching will occur only when placing the microphone assembly 10 on the user's chest (or when changing the placement).

[0040] Preferably, safeguard mechanisms may be provided for preventing undesired beam switching. For example, as already mentioned above, the beam selection unit 34 may be configured to analyze the signal of the accelerometer sensor 30 in a manner so as to detect a shock to the microphone assembly 10 and to suspend activity of the beam selection unit 34 so as to avoid changing of the subset of beams during times when a shock is detected, when the microphone assembly 10 is moving too much. According to another example, the output unit 40 may be configured to suspend channel selection, by discarding estimated SNR values during acoustical shocks, during times when the variation of the energy of the audio signals provided by the microphones is found to be very high, i.e. is found to be above a threshold, which is an indication of an acoustical shock, e.g. due to hands clap or an object falling on the floor. Further, the output unit 40 may be configured to suspend channel selection during times when the input level of the audio signals provided by the microphones is below a predetermined threshold or speech threshold. In particular, the SNR values may be discarded in case that the input level is very low, since there is no benefit of switching beams when the user is not speaking.

[0041] In Fig.1b examples of the beam orientation obtained by a microphone assembly according to the invention are schematically illustrated for the three use situations of Fig. 1a, wherein it can be seen that also for tilted and/or misplaced positions of the microphone assembly the beam points essentially towards the user's mouth.

[0042] According to one embodiment, the microphone assembly 10 may be designed as (i.e. integrated within) an audio signal transmission unit for transmitting the audio signal output 42

via a wireless link to at least one audio signal receiver unit or, according to a variant, the microphone assembly 10 may be connected by wire to such an audio signal transmission unit, i.e. the microphone assembly 10 in these cases acts as a wireless microphone. Such wireless microphone assembly may form part of a wireless hearing assistance system, wherein the audio signal receiver units are body-worn or ear level devices which supply the received audio signal to a hearing aid or other ear level hearing stimulation device. Such wireless microphone assembly also may form part of a speech enhancement system in a room.

[0043] In such wireless audio systems, the device used on the transmission side may be, for example, a wireless microphone assembly used by a speaker in a room for an audience or an audio transmitter having an integrated or a cable-connected microphone assembly which is used by teachers in a classroom for hearing-impaired pupils/students. The devices on the receiver side include headphones, all kinds of hearing aids, ear pieces, such as for prompting devices in studio applications or for covert communication systems, and loudspeaker systems. The receiver devices may be for hearing-impaired persons or for normal-hearing persons; the receiver unit may be connected to a hearing aid via an audio shoe or may be integrated within a hearing aid. On the receiver side a gateway could be used which relays audio signal received via a digital link to another device comprising the stimulation means.

[0044] Such audio system may include a plurality of devices on the transmission side and a plurality of devices on the receiver side, for implementing a network architecture, usually in a master-slave topology.

[0045] In addition to the audio signals, control data is transmitted bi-directionally between the transmission unit and the receiver unit. Such control data may include, for example, volume control or a query regarding the status of the receiver unit or the device connected to the receiver unit (for example, battery state and parameter settings).

[0046] In Fig. 8 an example of a use case of a wireless hearing assistance system is shown schematically, wherein the microphone assembly 10 acts as a transmission unit which is worn by a teacher 11 in a classroom for transmitting audio signals corresponding to the teacher's voice via a digital link 60 to a plurality of receiver units 62, which are integrated within or connected to hearing aids 64 worn by hearing-impaired pupils/students 13. The digital link 60 is also used to exchange control data between the microphone assembly 10 and the receiver units 62. Typically, the microphone arrangement 10 is used in a broadcast mode, i.e. the same signals are sent to all receiver units 62.

[0047] In Fig. 9 an example of a system for enhancement of speech in a room 90 is schematically shown. The system comprises a microphone assembly 10 for capturing audio signals from the voice of a speaker 11 and generating a corresponding processed output audio signal. The microphone assembly 10 may include, in case of a wireless microphone assembly, a transmitter or transceiver for establishing a wireless - typically digital - audio link 60. The output audio signals are supplied, either by a wired connection 91 or, in case of a wireless microphone assembly, via an audio signal receiver 62, to an audio signal processing unit 94 for

processing the audio signals, in particular in order to apply a spectral filtering and gain control to the audio signals (alternatively, such audio signal processing, or at least part thereof, could take place in the microphone assembly 10). The processed audio signals are supplied to a power amplifier 96 operating at constant gain or at an adaptive gain (preferably dependent on the ambient noise level) in order to supply amplified audio signals to a loudspeaker arrangement 98 in order to generate amplified sound according to the processed audio signals, which sound is perceived by listeners 99.

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US20160255444A1 [0005]
- US20140270248A1 [0006]
- US9066169B2 [0007]
- US9066170B2 [0008]

MIKROFONAGGREGAT TIL AT BÆRE PÅ BRUGERENS BRYST

Patentkrav

1. Et mikrofonaggregat til at bære på brugerens bryst, som omfatter:
 - mindst tre mikrofoner (20, 21, 22) til opsamling af lydsignaler fra brugerens stemme, idet mikrofonerne definerer et mikrofonplan;
 - en accelerationssensor (30) til detektering af tyngdeacceleration i mindst to ortogonale dimensioner, så gravitationsretningen (G_{xy}) kan fastslås;
 - en stråleformerenhed (32) til behandling af de opsamlede lydsignaler på en måde, så der dannes en pluralitet af N akustiske stråler (1a-6a, 1b-6b) med retninger fordelt over mikrofonplanet,
 - en enhed (34) til udvælgelse af en undergruppe af M akustiske stråler fra de N akustiske stråler, kendetegnet ved, at de M akustiske stråler er de af de N akustiske stråler, hvis retning ligger tættest på den retning (26), der er antiparallel med tyngderetningen fastslået via tyngdeaccelerationen opfanget af accelereringssensoren;
 - kendetegnet ved, at**
 - en lydsignalbehandlingsenhed (36) har M uafhængige kanaler (36A, 36B), en for hver af de M akustiske stråler fra undergruppen, til at producere et udgangslydssignal for hver af de M akustiske stråler;
 - en enhed (38) til estimering af lydsignalets talekvalitet på hver af kanalerne; og
 - en udgangsenhed (40) til udvælgelse af signalet på den kanal med den højeste estimerede talekvalitet som udgangssignal (42) for mikrofonaggregatet (10).
2. Mikrofonaggregatet i patentkrav 1, kendetegnet ved, at enheden til udvælgelse af stråleundergruppen (34) er konfigureret til at vælge, som undergruppe, de to akustiske stråler (1a-6a, 1b-6b), hvis retning støder op til den retning (26), der er antiparallel med den fastslåede tyngderetning (G_{xy}), og/eller kendetegnet ved, at enheden til udvælgelse af stråleundergruppen (34) er konfigureret til at tage en gennemsnitsmåling med accelerometersensoren (30) i tid for at forbedre målingens pålidelighed, idet enheden til udvælgelse af stråleundergruppen (34) konfigureres til at bruge et signal med en gennemsnitlig tidskonstant på mellem 100 msek. og 500 msek.
3. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at enheden til udvælgelse af stråleundergruppen (34) er konfigureret til at analysere det signal, der fremskaffes af accelerometersensoren (30) via en bevægelsesdetekteringsalgoritme til detektering af bevægelse i mikrofonaggregatet (10) og til at suspendere udvælgelsen af undergruppen på tidspunkter, hvor bevægelse detekteres, og/eller kendetegnet ved, at enheden til udvælgelse af stråleundergruppen (34) er konfigureret til ved brug af projektionen (G_{xy}) fra den fysiske tyngderetning på mikrofonplanet som ovennævnte

afgjorte tyngderetning til udvælgelse af undergruppen af akustiske stråler (1a-6a, 1b-6b), mens den ser bort fra projektionen af den fysiske tyngderetning på akse (z), der er normal for mikrofonplanet, idet enheden til udvælgelse af stråleundergruppen (34) konfigureres til at beregne et skalarprodukt mellem projektionen af den fysiske tyngderetning på mikrofonplanet og et sæt enhedsvektorer tilpasset i retning af hver af de N akustiske stråler (1a-6a, 1b-6b) og til at vælge disse M akustiske stråler til undergruppen, hvilket resulterer i M -højeste skalarprodukter.

4. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at stråleformerenheden (32) er konfigureret til at behandle de opsamlede lydsignaler på en måde, så retningerne af de N akustiske stråler (1a-6a, 1b-6b) fordeles jævnt over mikrofonplanet.

5. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at mikrofonaggregatet (10) omfatter tre mikrofoner (20, 21, 22), kendetegnet ved, at mikrofonerne er fordelt nogenlunde jævnt i en cirkel, og kendetegnet ved, at de enkelte vinkler mellem tilstødende mikrofoner er fra 110 til 130 grader, idet summen af de tre vinkler er 360 grader.

6. Mikrofonaggregatet i patentkrav 5, kendetegnet ved, at stråleformerenheden (32) er konfigureret til at danne 12 akustiske stråler (1a-6a, 1b-6b), kendetegnet ved, at stråleformerenheden (32) er konfigureret til at bruge *Delay-and-Sum* stråleformning med signaler fra mikrofonpar (20, 21, 22) til at danne en første del (1b-6b) af de akustiske stråler og til at bruge stråleformning via en vægtet kombination af signaler fra alle mikrofonerne til at danne en anden del (1a-6a) af de akustiske stråler, kendetegnet ved, at hver af de akustiske stråler (1b-6b) i den første del af de akustiske stråler ligger parallelt med en af siderne på den trekant (24), der dannes af mikrofonerne (20, 21, 22), kendetegnet ved, at de akustiske stråler i første del ligger parvist antiparallelt med hinanden, kendetegnet ved, at hver af de akustiske stråler (1a-6a) i den anden del af de akustiske stråler ligger parallelt med en af medianerne på den trekant (24), der dannes af mikrofonerne (20, 21, 22), og kendetegnet ved, at de akustiske stråler i anden del ligger parvist antiparallelt med hinanden.

7. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at hver af mikrofonerne (20, 21, 22) er en rundstrålende mikrofon og/eller kendetegnet ved, at accelereringssensoren (30) er et accelerometer med tre akser.

8. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at enheden til estimering af talekvalitet (38) er konfigureret til at estimere signal-til-støj forholdet på hver kanal (36A, 36B) som den estimerede talekvalitet, kendetegnet ved, at enheden til estimering af talekvalitet (38) er konfigureret til at beregne den øjeblikkelige bredbåndsenergi på hver kanal (36A, 36B) i det logaritmiske domæne, kendetegnet ved, at enheden til estimering af talekvalitet (38) er konfigureret til at beregne et

førstegangsgennemsnit af ovennævnte øjeblikkelige bredbåndsenergi ved hjælp af tidskonstanter, som sikrer at førstegangsgennemsnittet er repræsentativt for taleindholdet i kanalen (36A, 36B), idet frigivelsestiden er længere end angrebstiden med en faktor på mindst 2, og til at beregne et andengangsgennemsnit af ovennævnte øjeblikkelige bredbåndsenergi ved hjælp af tidskonstanter, som sikrer at andengangsgennemsnittet er repræsentativt for støjindholdet i kanalen, idet angrebstiden er længere end frigivelsestiden med en faktor på mindst 10, og til at bruge, i et logaritmisk domæne, forskellen mellem førstegangsgennemsnittet og andengangsgennemsnittet til estimering af signal-til-støj forholdet.

9. Mikrofonaggregatet i et af patentkrav 1 til 7, kendetegnet ved, at enheden til estimering af talekvalitet (38) er konfigureret til at estimere en score for taleforståelighed på hver kanal (36A, 36B) som den estimerede talekvalitet.

10. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at udgangsenheden (40) er konfigureret til at tage gennemsnittet af lydsignalets estimerede talekvalitet på hver kanal (36A, 36B) ved valg af den kanal, der har den højeste estimerede talekvalitet, og kendetegnet ved, at udgangsenheden (40) er konfigureret til at bruge en signalgennemsnitlig tidskonstant på fra 1 sek. til 10 sek.

11. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at udgangsenheden (40) er konfigureret til at vurdere en vægt på 100% i udgangssignalet til den kanal (36A, 36B), der har den højeste estimerede talekvalitet, bortset fra skifteperioder, under hvilke udgangssignalet skifter fra en tidligere valgt kanal til en nyvalgt kanal, kendetegnet ved, at udgangsenheden (40) er konfigureret til at vurdere, under skifteperioder, en tidsvariabelvægtning af den tidligere valgte kanal (36A, 36B) og af den nyvalgte kanal (36B, 36A) på en sådan måde, at den tidligere valgte kanal fades ud og den nyvalgte kanal fades ind, og kendetegnet ved, at udgangsenheden er konfigureret til at fade den nyvalgte kanal (36A, 36B) ind hurtigere, end den tidligere valgte kanal (36B, 36A) fades ud.

12. Mikrofonaggregatet i et af ovenstående patentkrav, kendetegnet ved, at udgangsenheden (40) er konfigureret til at suspendere kanalvalg på tidspunkter, hvor variationen i lydsignalets energiniveau ligger over en forudbestemt grænse og/eller på tidspunkter, hvor lydsignalets taleniveau ligger under en forudbestemt grænse, og/eller kendetegnet ved, at lydsignalbehandlingsenheden (36) er konfigureret til at anvende adaptiv stråleformning på hver kanal (36A, 36B), såsom ved at kombinere modsatte cardioider langs retningsaksen for den pågældende akustiske stråle.

13. Et system til fremskaffelse af lyd til mindst en bruger, som omfatter mikrofonaggregatet (10) i et af ovenstående patentkrav, hvor mikrofonaggregatet er designet til at være en sendeenhed for lydsignaler, hvor lydsignalerne sendes via et trådløst link (60), mindst en modtagerenhed (62) til modtagelse af lydsignaler fra

sendeenheden via det trådløse link; og en enhed (64) til stimulering af brugerens/brugernes hørelse i overensstemmelse med et lydsignal, der fremskaffes fra modtagerenheden.

14. Et system til taleforbedring i et lokale, der omfatter mikrofonaggregatet (10) fra et af patentkrav 1 til 12, hvor mikrofonaggregatet er designet til at være en sendeenhed for lydsignaler, der sendes via et trådløst link (60), mindst en modtagerenhed (62) til modtagelse af lydsignaler fra sendeenheden via det trådløse link, og et højttalerarrangement (98) til generering af lyd i overensstemmelse med et lydsignal, der fremskaffes fra modtagerenheden.

15. En metode til generering af et udgangsslydsignal (42) fra en brugers stemme ved hjælp af et mikrofonaggregat (10), der omfatter en fastgørelsesmekanisme, mindst tre mikrofoner (20, 21, 22) der definerer et mikrofonplan, en accelereringsensor (30), og en lydbehandlingsfacilitet (32, 34, 36, 38, 40), hvor metoden omfatter:

fastgørelse af mikrofonaggregatet med fastgørelsesmekanismen på brugerens tøj;

detektering, via accelereringssensoren, af tyngdeacceleration i mindst to ortogonale dimensioner, så tyngderetningen (G_{xy}) kan fastslås;

opsamling af lydsignaler fra brugerens stemme via mikrofonerne,

behandling af de opsamlede lydsignaler på en måde, så der dannes en pluralitet af N akustiske stråler (1a-6a, 1b-6b) med retninger fordelt over mikrofonplanet;

udvælgelse af en undergruppe af M akustiske stråler fra de N akustiske stråler, kendetegnet ved, at de M akustiske stråler er de af de N akustiske stråler, hvis retning ligger tættest på den retning (26), der er antiparallell med den fastslåede tyngderetning;

kendetegnet ved

behandling af lydsignaler i M uafhængige kanaler (36A, 36B), en for hver af the M akustiske stråler fra undergruppen, til at producere et udgangsslydsignal for hver af de M akustiske stråler;

estimering af lydsignalets talekvalitet på hver af kanalerne; og

udvælgelse af lydsignalet på den kanal med den højeste estimerede talekvalitet som udgangssignal for mikrofonaggregatet.

DRAWINGS

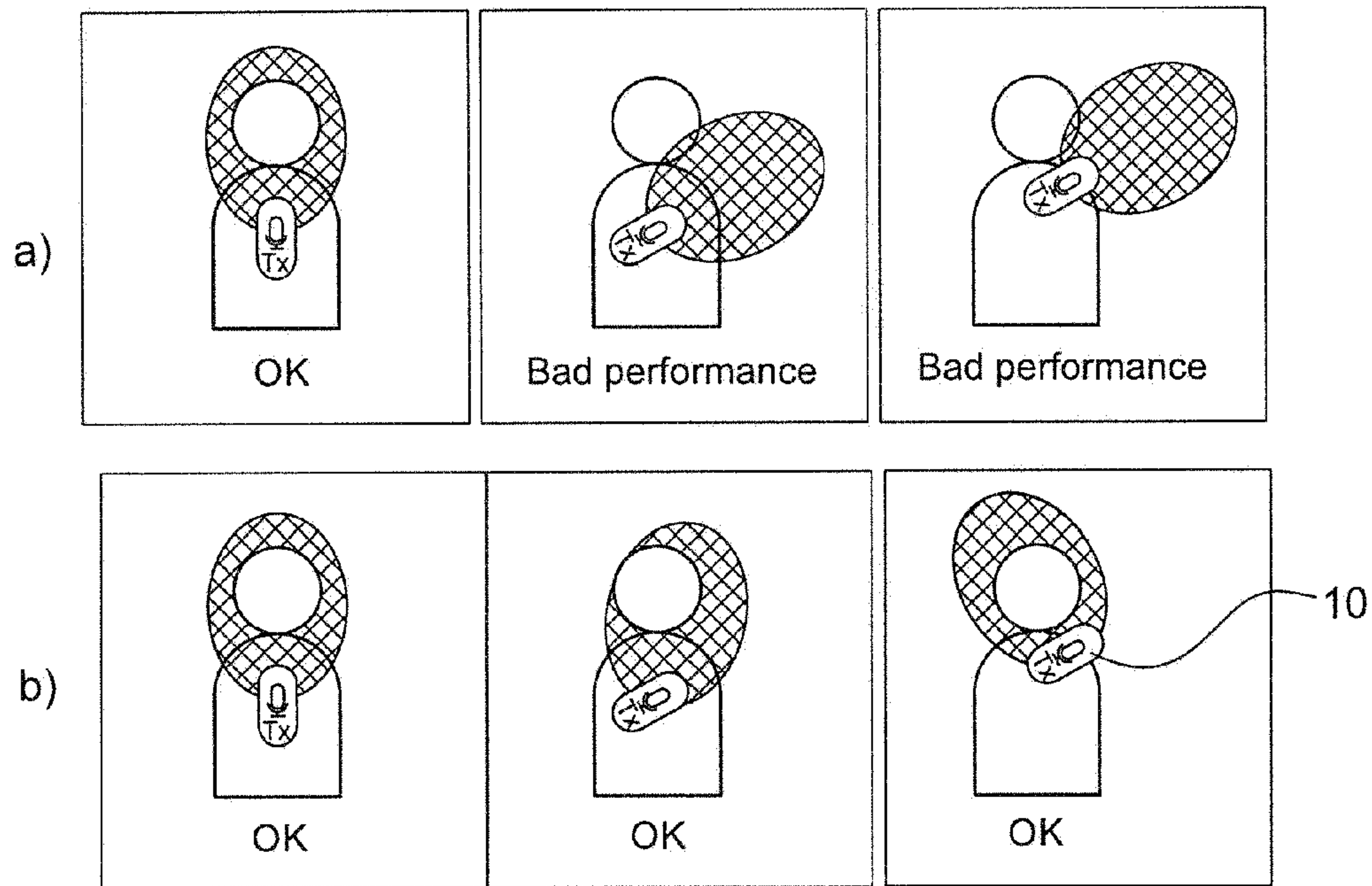


FIG. 1

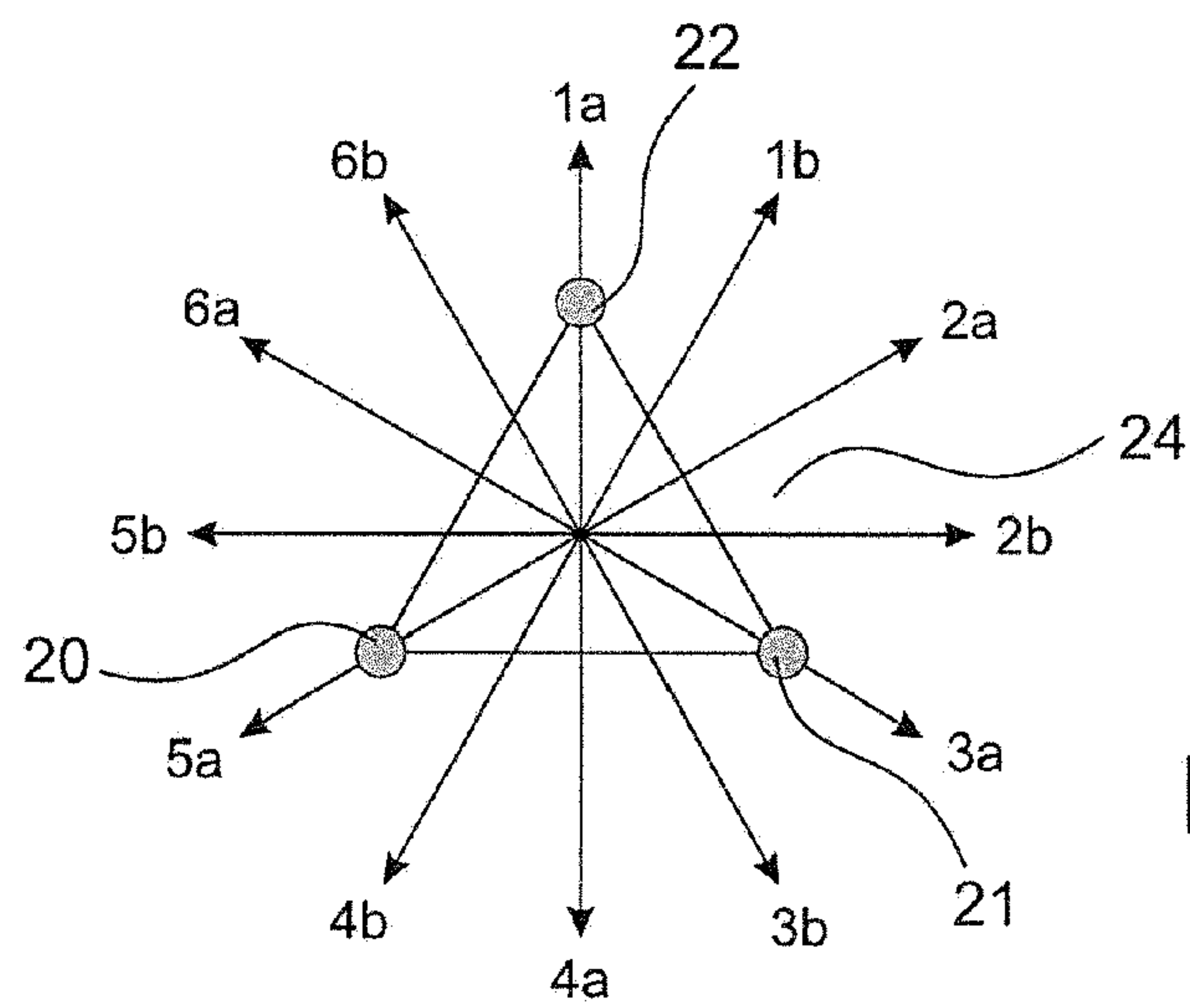
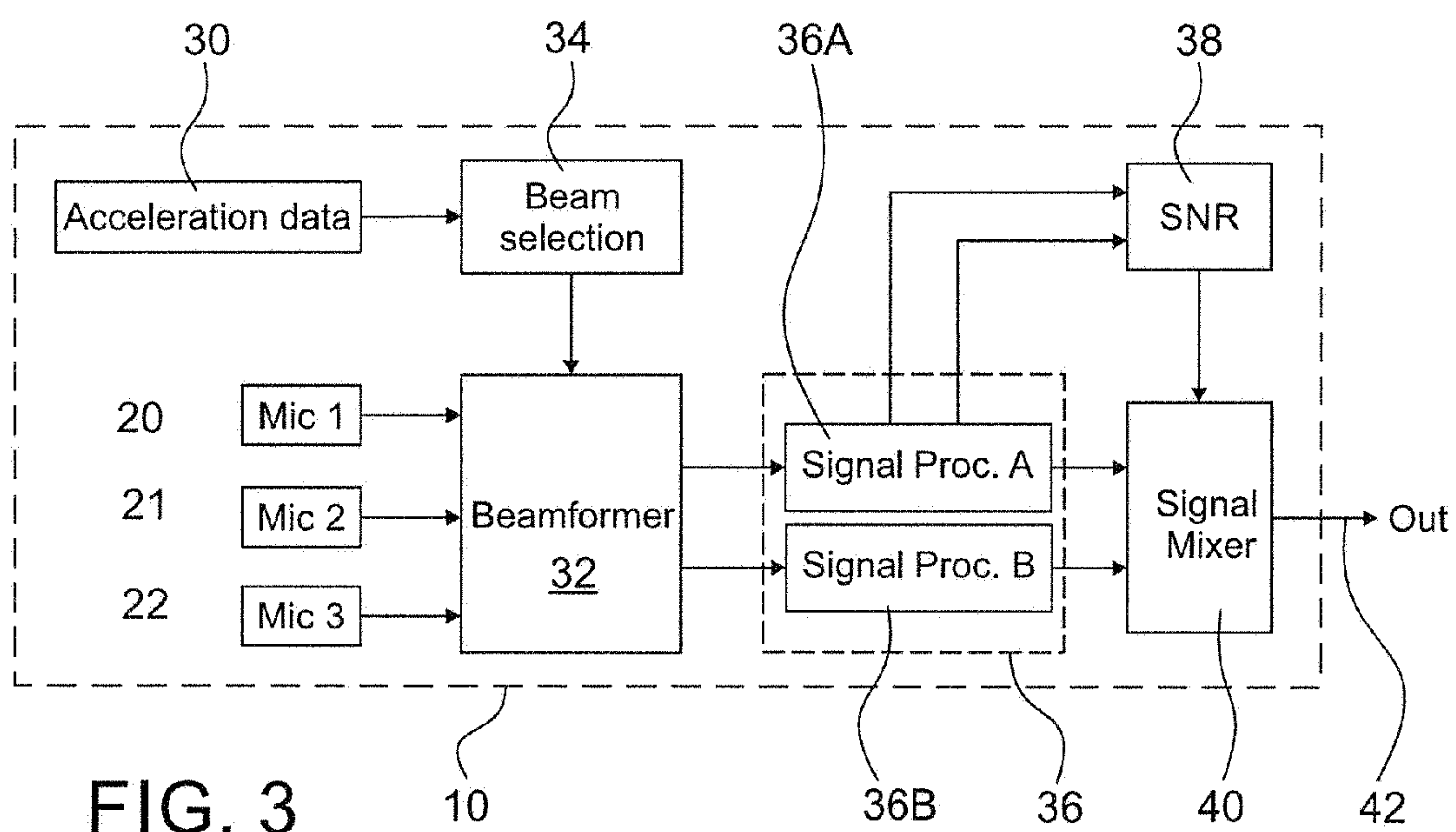
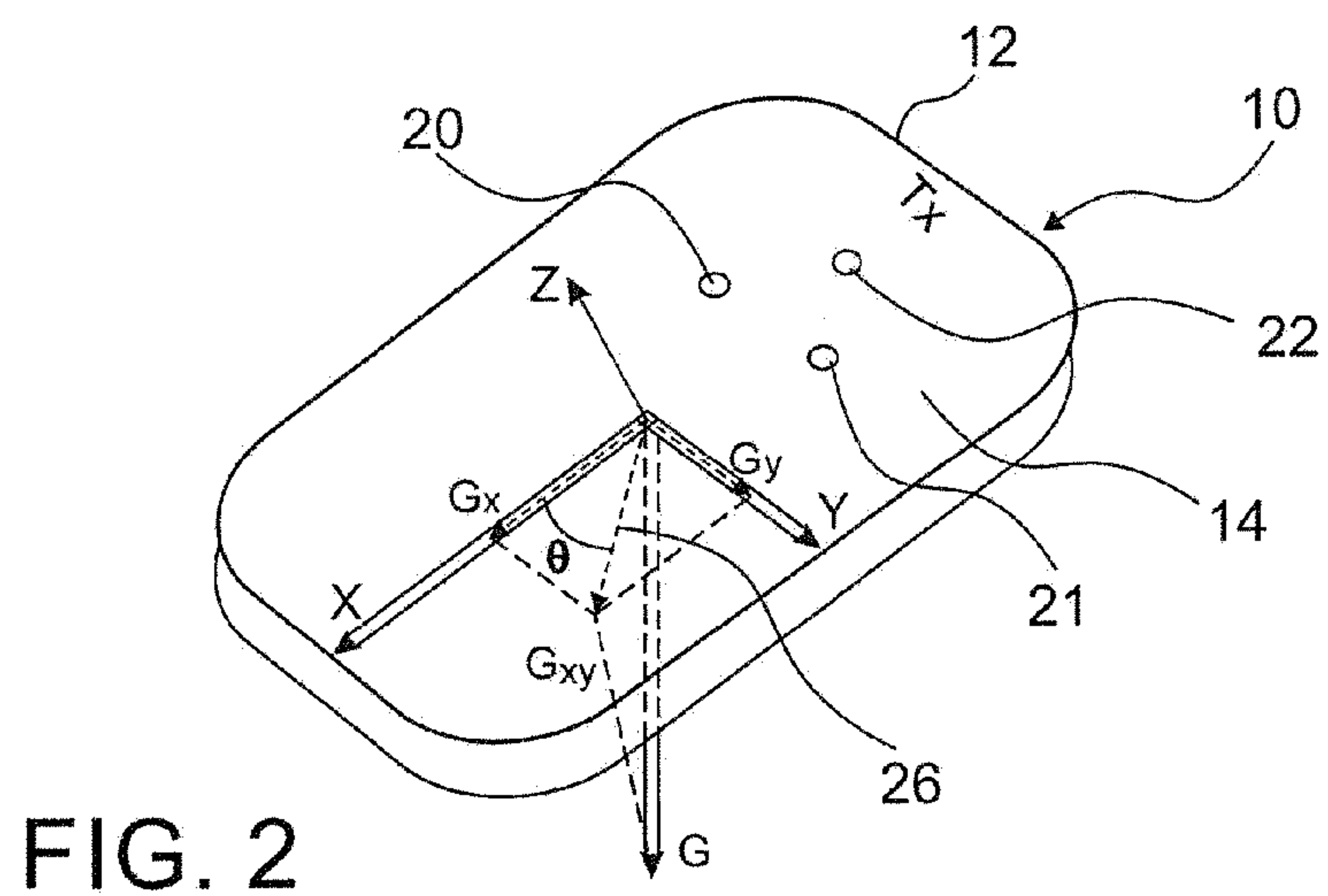
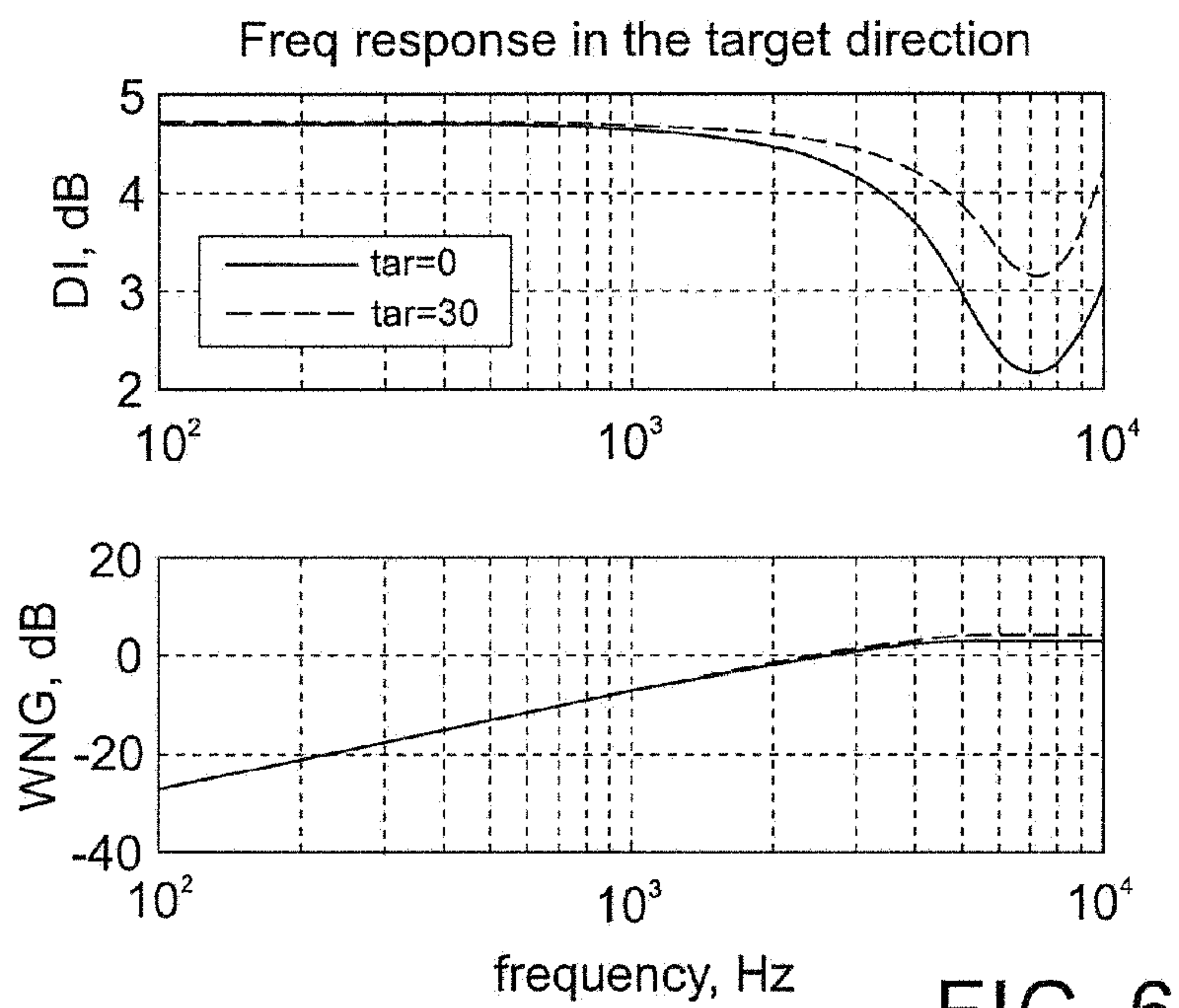
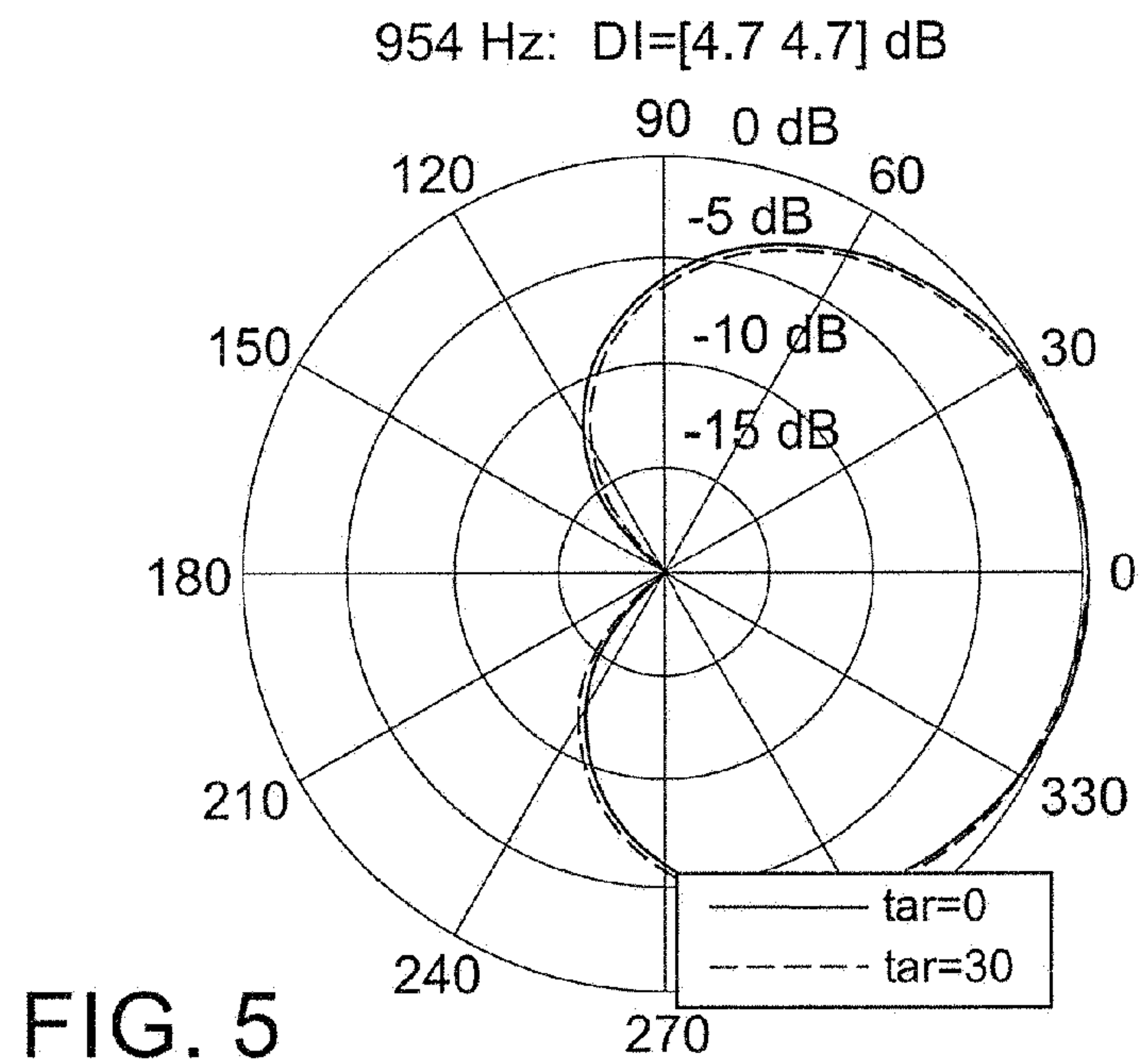


FIG. 4





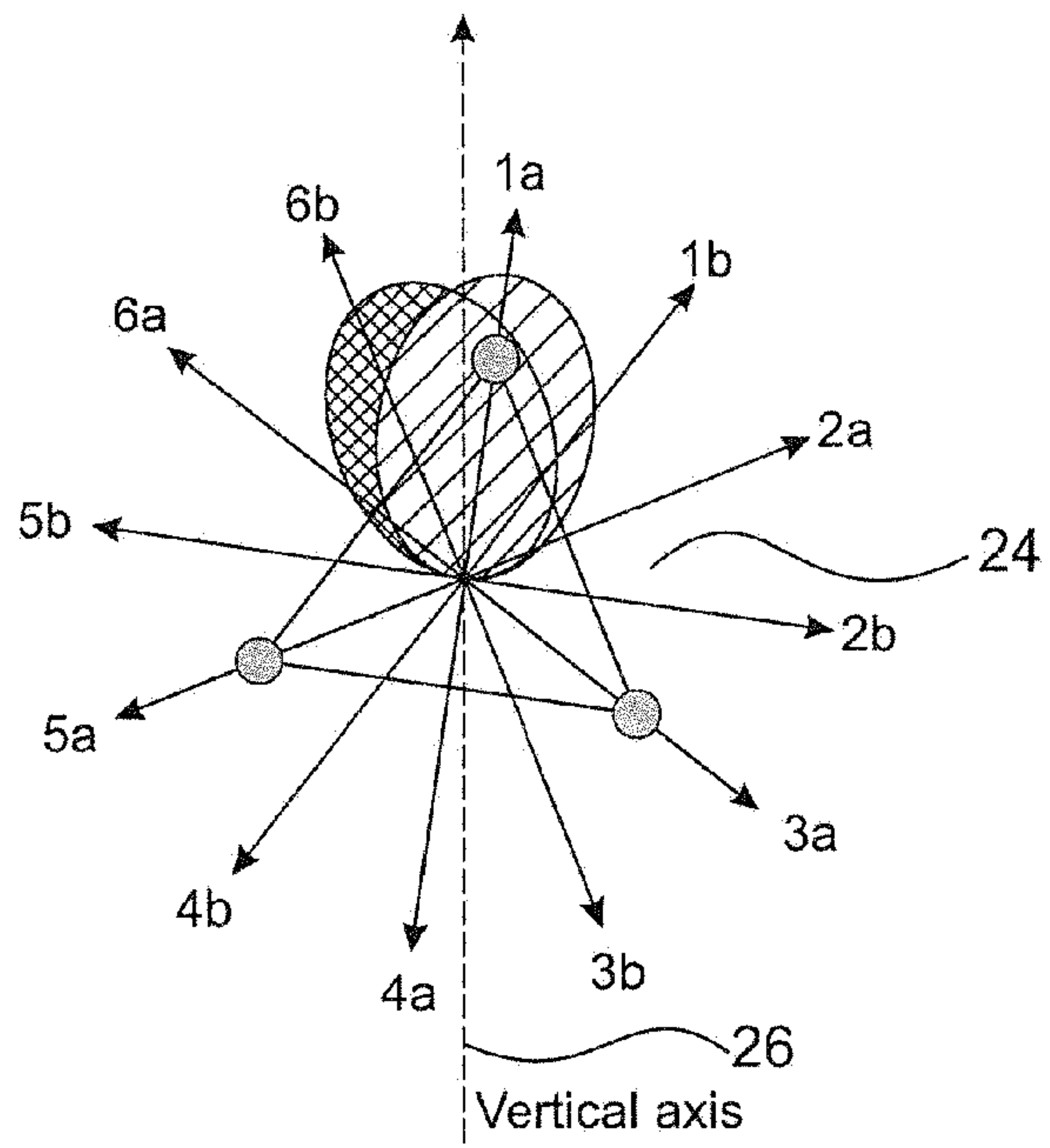


FIG. 7

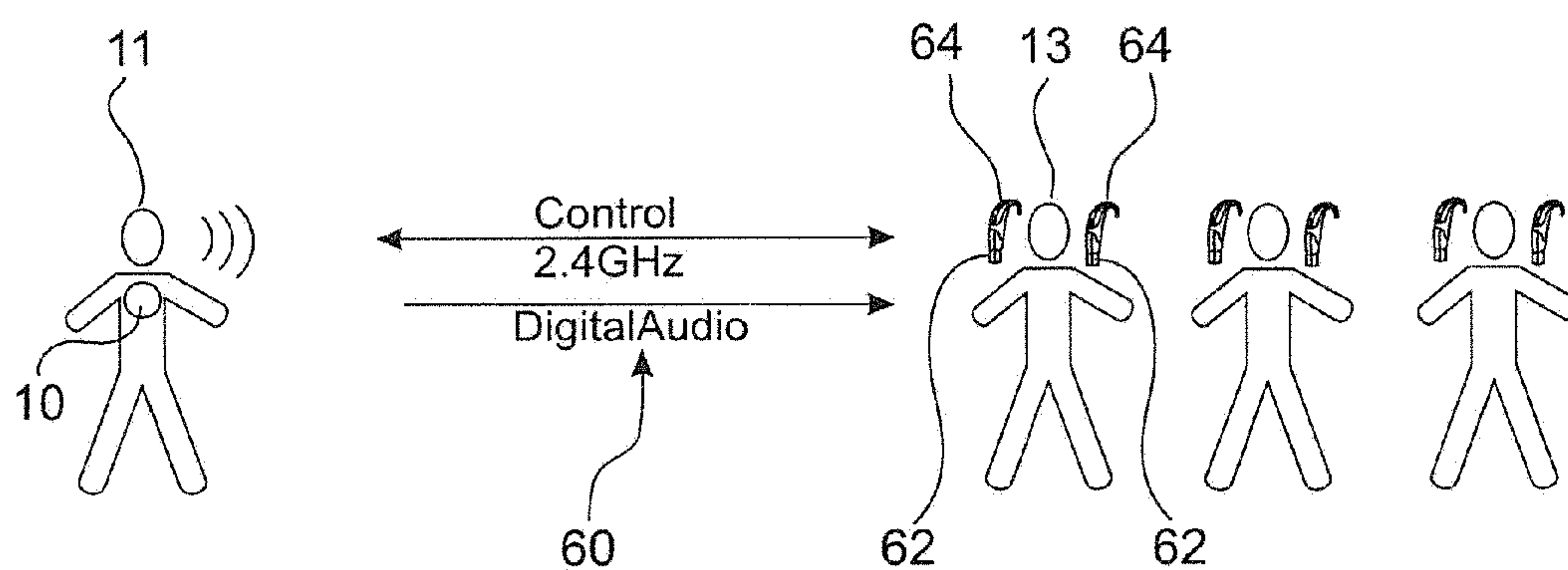


FIG. 8

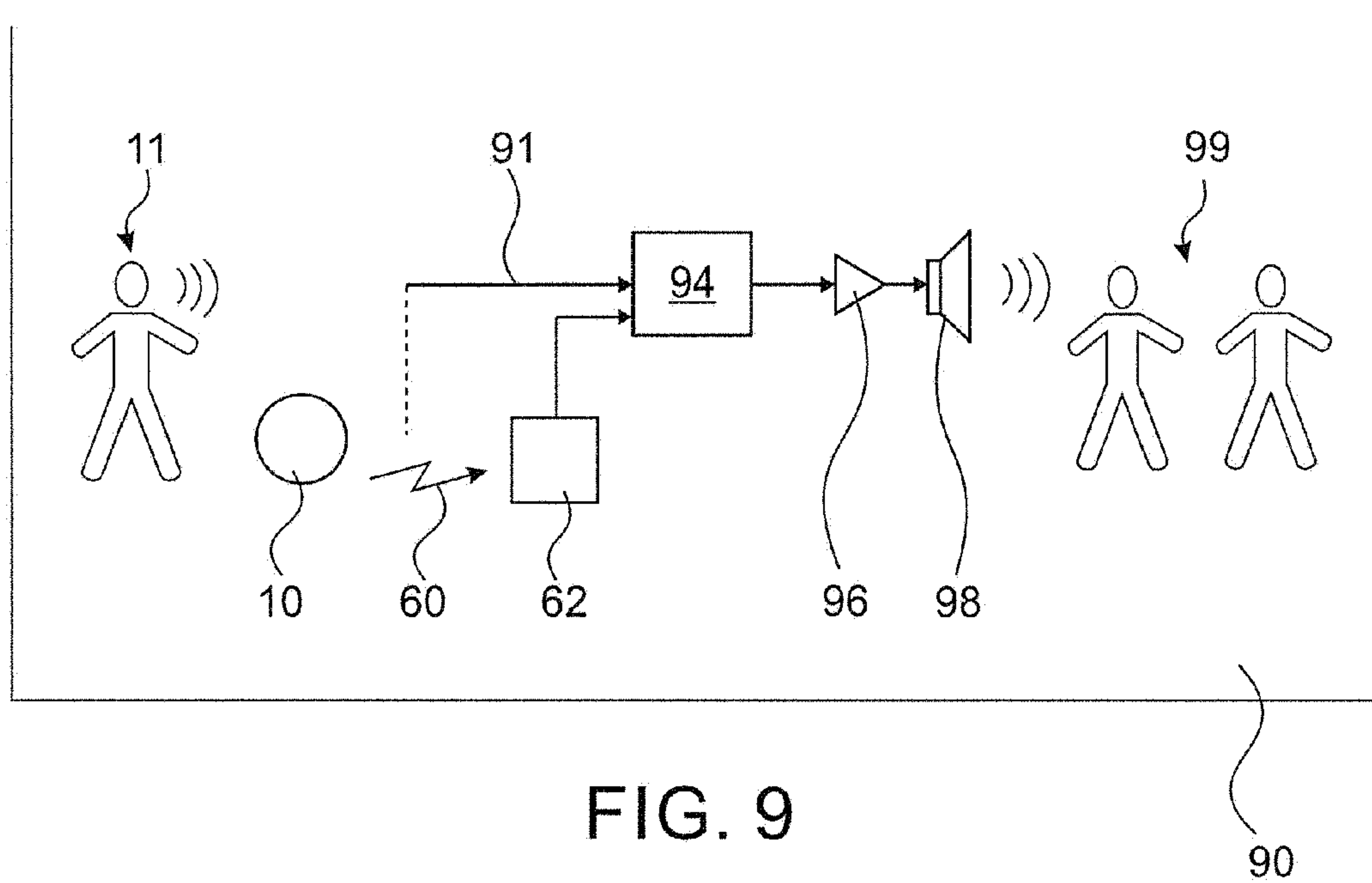


FIG. 9